Fast scalable implicit solver with convergence of physics-based simulation & data-driven learning: toward high-fidelity simulation with digital twin city

Tsuyoshi Ichimura[1,2,3], Kohei Fujita[1,3], Kentaro Koyama[4], Ryota Kusakabe[1],
Kazuo Minami[3], Hikaru Inoue[4], Seiya Nishizawa[3], Miwako Tsuji[3], Tatsuo Nishiki[4],
Muneo Hori[5], Lalith Maddegedara[1,3], Naonori Ueda[2]
[1]Earthquake Research Institute & Department of Civil Engineering, The University of Tokyo
[2]Center for Advanced Intelligence Project, RIKEN, [3]Center for Computational Science, RIKEN
[4]Fujitsu Limited, [5]Japan Agency for Marine-Earth Science and Technology

I. INTRODUCTION

In the with-corona and post-corona era, there are great expectations for supercomputing that enables unexplored analysis as guardians of the cities and societies. Conventionally, optimization of logistics and energy for smart cities were considered as one of the frontiers of supercomputing (e.g. [1]), but contributions are also expected in evaluating the safety and robustness against threats to cities and society such as the Covid-19 pandemic. For example, not only the pandemic itself but also complex and deadly threats such as the combination of pandemic and natural disasters is feared to occur; thus, there is a demand from society for more accurate estimation of the state of the city and response for mitigation and control.

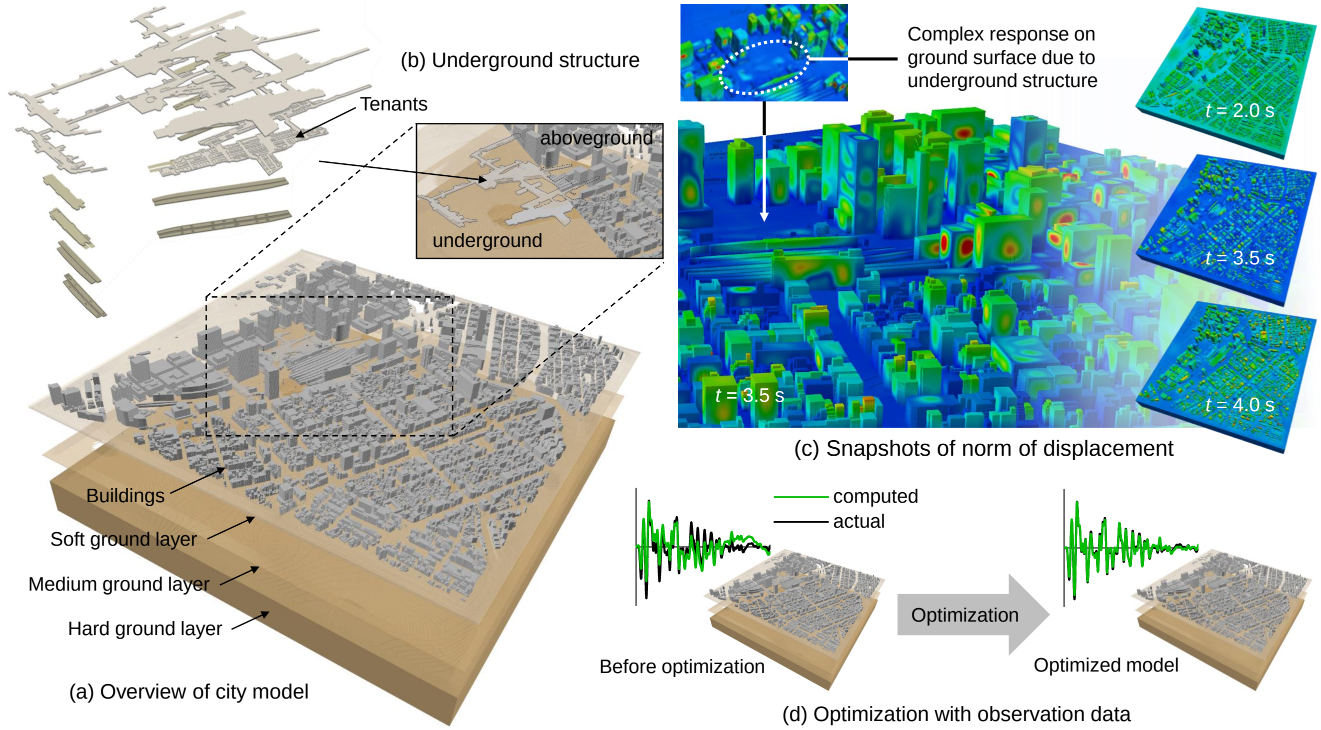
One of the ways to improve the accuracy and resolution of city state estimation is generating high-fidelity digital twin of the city, updating it in real time by assimilation with a large amount of measurement data of the city, and through performing high-resolution simulations based on the Big Data & Extreme Computing (BDEC) (e.g., [2]) approach. However, since the cities are huge and complicated, the analysis cost of highly-detailed simulations will be huge; so, there is a strong demand for reduction of analysis cost for enabling the above.

In this research, we attempt to reduce the analysis cost focusing on earthquake disasters, which is one of the disasters that could be a fatal threat to lives and cities. The urban earthquake problem is a nonlinear time evolution problem from a mathematical viewpoint, and from a computational science/computer science viewpoint, the random memory access dominated low-order unstructured finite element analysis used in solving this problem is challenging for attaining performance; thus, great leap in methodology development becomes necessary. In this paper, a new approach is developed in merging high-performance computing (HPC)-enhanced physics-based simulations with data-driven learning. Here we propose a new, fast HPC-based scalable implicit solver $IRIS$ that uses data generated while conducting physics-based modeling for data-driven learning. This solver will accelerate the convergence process by making constraints in the solution space for local errors. For the challenging urban earthquake problem, we showed that the developed method significantly reduces analysis cost compared to conventional solvers: The scalable implicit solver $IRIS$ achieved 15.2-fold speedup over the conventional method $PCGE$ with 96.4% size-up scalability up to 24,576 nodes (1,179,648 cores) of Arm SVE CPU-based Fugaku. $IRIS$ increased the capability of solving extremely large-scale problems within a short period of time. The performance of $IRIS$ was 12.7% of FP64 peak on 48 nodes of Fugaku and 12.8% of FP64 peak on 48 nodes of an Intel Xeon-based (Cascade Lake) supercomputer Oakbridge-CX, which is very high for the low-order unstructured finite-element method involving random access-dominated sparse-matrix operations and global communication. Furthermore, $IRIS$ attained a 10.3-fold speedup over the current state of the art solver $GAMERA$ [3] (SC14 Gordon Bell Prize finalist) on a super-high resolution analysis of urban models which were considered too costly to conduct in the past (see Fig. 1; see poster for details).

Through HPC-based Extreme Computing, which combines physics-based modeling and data-driven learning, added value is expected to be generated from Big Data of physical space/cyber space/analysis space. By expanding the frontier to urban analysis, this research is expected to become one of the innovations that further enhance the capability of supercomputing so that it helps to defend and save cities and societies.

II. OVERVIEW OF INNOVATION REALIZED

This section outlines the innovation realized (see poster "Innovation realized" for details). To evaluate the urban behavior under large earthquakes with high reliability, it is necessary to analyze the behavior of the digital twin of the city in high details. This leads to solving nonlinear time evolution problems on a wide area with locally detailed and complex structures. The size of the target problem is $10^3 \times 10^3 \times 10^{1\sim 2}$ m with structures that have very complex geometry in a resolution of $10^{-2\sim -1}$ m. This requires solving nonlinear time evolution continuum mechanics problems with three-dimensional finite-element methods using low-order solid elements (i.e., second-order unstructured tetrahedral elements), which can analytically satisfy the traction-free boundary conditions at the surface. Here, linear system of equations $\mathbf{A}^n \delta \mathbf{u}^n = \mathbf{f}^n$ is solved in each of the $10^{3\sim 4}$ time steps with double precision. The degrees-of-freedom of $\delta \mathbf{u}$ is over 10^{10} and also computation associated with the sparse matrix $\mathbf{A}^n$ becomes random access cost dominant as unstructured low-ordered finite-elements are used, which is not straightforward to attain performance on recent computer architecture. Based on the above, a surrogate model is generated based on a data-driven approach learning the convergence characteristics up to high-frequency modes in solutions of past time steps, and used it as a preconditioner to



(a) Overview of city model

(b) Underground structure

(c) Snapshots of norm of displacement

(d) Optimization with observation data

Fig. 1. Computation example of 11,002,859,706 degrees of freedom; 2,593,545,132 element model of fully coupled ground and aboveground/underground structure earthquake shaking analysis using $IRIS$ on 98,304 CPU cores (2,048 nodes) of Fugaku: (a) The 2 × 2 km domain was modeled with 0.5-m tetrahedral elements with three soil layers; 9323 buildings; (b) an underground complex with 645 tenants; (c) The complex nonlinear seismic response was computed as a result of the three-dimensional ground structure and building configuration; and (d) Fast analysis by $IRIS$ enables the optimization of city models and is expected to be the key technology for realizing a highly reliable digital twin.

improve convergence characteristics for solving $\mathbf{A}^n \delta \mathbf{u}^n = \mathbf{f}^n$. In this algorithm, not only the convergence characteristic is improved, but it transfers the random access computation of $\mathbf{A}^n$ to sequential access based data-driven computation; thus, it is suitable for recent computer architecture. This algorithm also can be applied to general nonlinear time evolution problems with bad convergence characteristics. Although most of the computational cost is transferred to the surrogate model, kernel of multiplying a vector to $\mathbf{A}^n$ remains as a costly kernel; thus, tuning of SpMV for the newest architecture is important. As Arm SVE aware tuning, we developed a SIMD buffering method and a coloring method that enables efficient use of the Fugaku system. Combining with the surrogate modeling with the tuning for the newest CPU architecture led to the significant speedup from the current state of the art.

III. IMPLICATIONS

By extending the convergence of physics-based simulation and data-driven learning in the HPC resources, we designed an algorithm that benefits from the strengths of both physics-based simulation and data-driven learning to accelerate solvers. By considering the local characteristics of time space and error space in the solving process, and by obtaining low-frequency components using physics equation-based modeling in a coarse model on the one hand and obtaining high-frequency components using data-driven learning on data of past computation results on the other, we designed a method that can improve the estimation accuracy for all low- to high-frequency components of the solution, and, thus, improve the convergence characteristics of the scalable implicit solver. This can be considered as change in the standard solver algorithm to an algorithm that is suitable for recent computer architecture. The physics-based simulation methods (with random memory access-dominant low-order unstructured finite-element computation) was shifted to data-driven learning (with sequential memory access-dominant vector computation). Together with the development of physics-based simulation techniques suitable for the compute architecture led to a high-performance solver. Although the difference in accuracy and characteristics of the physics-based simulation and data-driven learning methods made it seem difficult to integrate these methods, we managed to facilitate the connection between these methods by designing a suitable interface, which solved the problems that were difficult to solve either by physics-based simulation or by data-driven learning. We can expect constructing more accurate data-rich data-driven models by utilizing last-level storage in the developed approach. This approach can be expected as a way to utilize the diverse storage hierarchies with different bandwidths and capacities, which are difficult to utilized in terms of accelerating physics-based simulations.

Acknowledgment: Results were obtained using Fugaku at Center for Computational Science, RIKEN, and Oakbridge-CX at Information Technology Center, The University of Tokyo. This work was supported by MEXT as “Program for Promoting Researches on the Supercomputer Fugaku” (Large-scale numerical simulation of earthquake generation, wave propagation and soil amplification: hp200126). This work was supported by Post K computer project (Priority Issue 3 - Development of integrated simulation systems for hazards and disasters induced by earthquakes and tsunamis: hp190177) and JSPS KAKENHI Grant Numbers JP18H05239, JP18H03795, JP17K14719.

REFERENCES

[1] HPC Connects Plenary: The Century of the City, SC17: International Conference for High Performance Computing, Networking, Storage and Analysis, 2017.

[2] Big Data and Extreme-scale Computing. [Online]. https://www.exascale.org/bdec/

[3] T. Ichimura, K. Fujita, S. Tanaka, M. Hori, M. Lalith, Y. Shizawa, and H. Kobayashi, "Physics-based urban earthquake simulation enhanced by 10.7 BlnDOF x 30 K time-step unstructured FE non-linear seismic wave simulation," in Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis (SC'14), IEEE Press, 2014, pp. 15–26.