13th International Conference on Hydroscience & Engineering
Advancement of hydro-engineering for sustainable development
June 18-22, 2018, Chongqing, China.

Parallel Processing of Multi Scenario Flood Simulation Using Cloud Computing Service

Satoshi Yamaguchi[1] *and Takashi Kusuda*[2]

1: Center for Technology Innovation – Digital Technology, R&D Group, Hitachi, Ltd. (Kokubunji, Tokyo, Japan)

2: ICT Solution Dept., Hitachi Power Solutions Co., Ltd. (Chiyoda, Tokyo, Japan)

ABSTRACT: We propose a flood simulation system using cloud-computing service in order to shorten calculation time. We conducted experiments to analyze independent 76 flood scenarios concurrently using 12 virtual machines. In the experiments, the calculation shortened to 1/10. The speed-up effect was almost proportionate to the number of virtual machines. The proposed system is effective for simultaneous execution of independent scenarios in short time. The proposed system is promising to apply to flood hazard mapping.

KEY WORDS: cloud computing; flood hazard mapping; 1D-2D coupled model; urban flood resilience; Microsoft Azure; DioVISTA Flood.

INTRODUCTION

Areas protected by levees would be heavily damaged if the levee fails during a flood. The location and the length of the levee failure significantly affects the amount of the loss. Okada et al. [2011] assessed 6 different scenarios of levee failure occurred by 200-year return period flood in Tone River (Japan). The maximum insured residential loss in the 6 scenarios was 400 times larger than the minimum one.

We should consider uncertainty of flood for adequate assessment of flood hazard and flood risk. A major source of the uncertainty is assumption of levee failure. Vorogushyn et al. [2010] assessed 3000 scenarios on the Elbe River reach (Germany) by using Inundation Hazard Assessment Model (IHAM). Domeneghetti et al. [2013] did 8000 scenarios on the Po River reach (Italy) in order to generated probabilistic flood hazard maps.

Flood hazard mapping method in Japan (MLIT, 2005) explicitly reflects a concept of uncertainty of levee failure. Artcle 14 in Flood Control Act requires the river manager to delineate the Flood Inundation Risk Areas. The method to delineate the Flood Inundation Risk Areas is defined in detail by an official technical manual (MLIT, 2015). Example of defined items in the manual are; the workflow of the analysis, the considered elements in the analysis (e.g., river, floodplain, levee, embankment of highway and railway in the floodplain, underpass of the embankment, pumping stations and sluice gates), the numerical model of the analysis, data format of the results (e.g., CSV and netCDF), and the style and the legend of the map (color corresponding to flood depth). The required numerical model is; 1) 1D unsteady flow model based on Saint-Venant Equations for river channel; 2) 2D unsteady flow model for floodplain; and 3) Empirical levee failure model in which length of levee failure changes as time goes on. The levee failure position is restricted to one per scenario. Consolidating the result of many scenarios, the extent and depth of Flood Inundation Risk Areas are delineated. The extent is the envelope of the flood areas in the all scenarios (Fig. 1). The depth is the maximum depth among the all scenarios.

In a typical case, the cell size of 1D model is 200 m, and that of 2D model is 25 x 25 m or finer. The simulation period is 3 days or more in order to assess flood duration time. The number of scenario is a few to several hundreds. These requirements make the total calculation time very long.

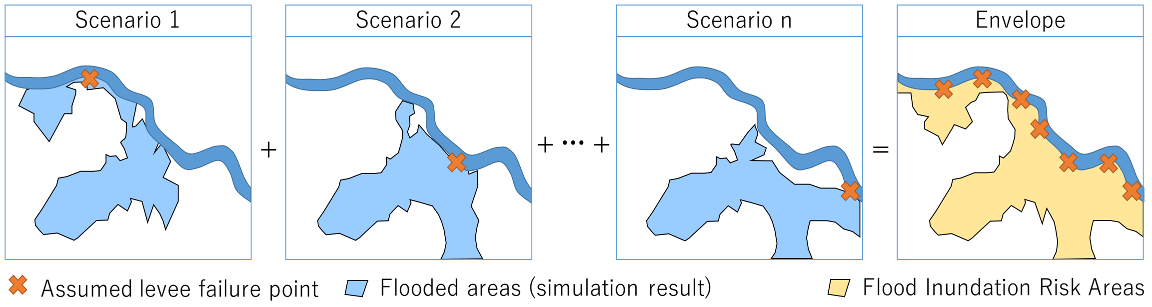

Fig. 1: Delineation method of Flood Inundation Risk Areas based on simulated flooded areas.

To shorten the calculation time, using cloud-computing service is a cost effective option. Cloud-computing service enables us to use computer resources without owing computing infrastructures. Quiroga et al. [2013] conducted uncertainty analysis of integrated flood models by using Amazon Elastic Compute Cloud (EC2) platform. Simulation of 99 scenarios by using hydraulic modeling system HEC-HMS took 22 min on single computer. The time reduced to 4 min by parallel processing on 5 virtual machines. Glenis et al. [2013] developed urban flood modeling software CityCat. They deployed the software to Amazon EC2, and conducted a parameter sweeps. By using the system, they processed 15,450 hours (21 months) of simulation time in a single calendar month. They concluded that using cloud-computing services is cost effective when the use is sporadic and with tight deadlines.

To shorten the calculation time, we have been proposed two methods. First, we proposed dynamic domain defining method (Dynamic DDM, Yamaguchi and Iwamura, 2007), which makes areas around flooded areas automatically included in the calculation area, and non-flooded areas are automatically excluded from calculation area during the calculation. Second, we proposed vectorization by using CPU extension instructions (Streaming SIMD Extensions) and multi-core simultaneous execution processing by using Microsoft Parallel Patterns Library (Yamaguchi and Yamaho, 2016). In this study, we propose the third method – using cloud-computing service.

PROPOSED METHOD

Our system configuration is shown in Fig. 2. The proposed system consists of one virtual machine (VM) for management (Manager), 12 VMs for calculation (Worker), and one shared disk. It is deployed to Microsoft Azure. The Manager distributes tasks to Workers. Manager VM controls Worker VMs' disk access so that only one Worker VM can access to the shared disk at any moment. The number of CPU cores per Worker VM is 16, and the total number of CPU cores is 192. We used our simulation software DioVISTA Flood. We used distributed runoff model and 1-D/2-D coupled model.

A workflow of a simulation on the system is as follows. Step 1: user uploads simulation scenario files to the Cloud. Simulation scenario file is an xml file that includes levee failure position and width information. Step 2: Manager VM distribute simulation scenario files to Worker VMs. Step 3: Worker VM runs the simulation. The result is written in a local storage of each VM. Step 4: Worker VM transfers the result file to the shared disk. Step 5: Manager VM notices to the user that all the scenarios are completed. Step 6: The user downloads the result files.

EXPERIMENTS

We simulated 76 scenarios of large-scale flood disaster in Yodo River basin (Japan, Fig. 3). Its basin area is 4,392 km^2. Because its Flood Inundation Risk Area includes major cities (Osaka and Kyoto), significant flood risk exists.

We built up a Yodo River model (Table 1). The runoff and river models were calibrated with historical flood. Simulated peak river water level of 2004 October flood was 9.6 cm higher than the observation, and was 10 min earlier than the observation (Fig. 4). We concluded the model accuracy is enough for the following simulation. We assumed twice strength of largest recorded rainfall (Typhoon Man-yi, T1318). We assumed 75 levee failure points (39 points on left bank and 36 points on right bank). We assumed that the levee failure point is restricted to one per scenario. Thus, we generated 75 levee failure scenarios and 1 overflowing scenario (no-levee-failure scenario).

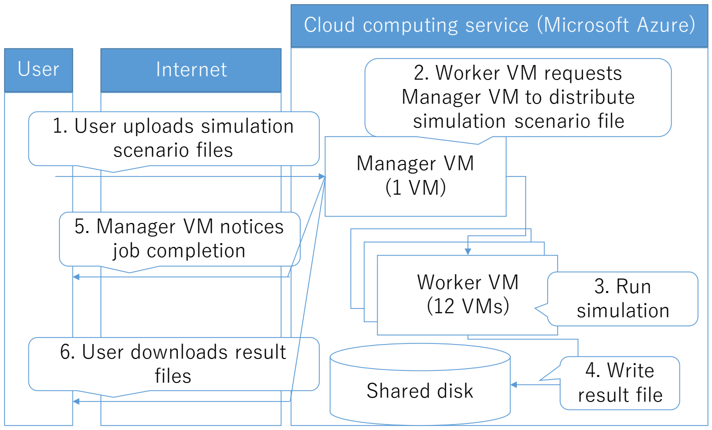

Fig. 2: System configuration of proposed system

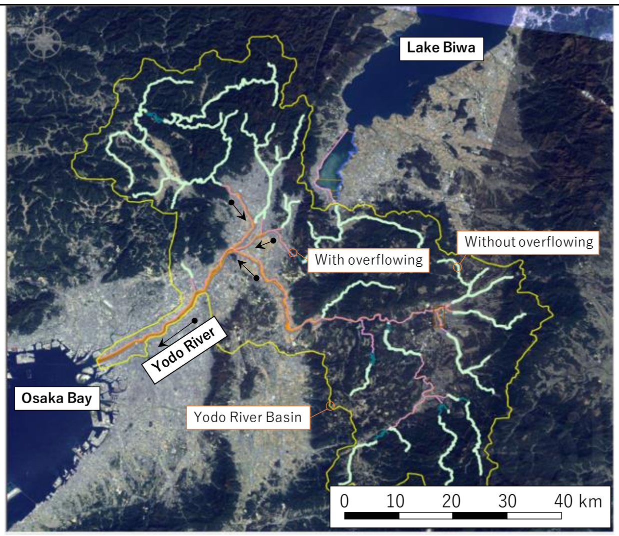

Fig. 3: Yodo River basin

We simulated the 76 scenarios by using the model. The analysis period was 72 hours and the cell size of 2D model was 25 m x 25 m. Flood areas were ranged between to 119 to 222 km^2 (average 150 km^2, Fig. 5). The scenario that resulted in the largest flooded area is shown in Fig. 6.

RESULTS

We analyzed 76 scenarios on the proposed system. Four VMs out of the 12 VMs analyzed 7 scenarios, and 8 VMs analyzed 6 scenarios (Fig. 7). All results were obtained after 47,660 seconds (13 hours) from the start of calculation. The total calculation time was 450,286 seconds (5 days 5 hours). That is, the calculation time was shortened to 1/10.4 compared with the case of executing on one PC. When we reduced the number of VM to half (6 VMs), the calculation period increased 1.9 times. It turned out that the number of VM proportionally contributed to shortening the calculation period.

Histograms of flooded area, calculation time, and result file size are shown in Fig. 5, Fig. 8, and Fig. 9, respectively. The correlation between flooded area and result file size is R^2=0.99. The correlation between flooded area and calculation time is R^2=0.68. These results indicate that required computer resources were almost proportionally corresponded to flooded area. This nature makes easy to estimate required computational time and disc space beforehand the simulation.

CONCLUSIONS

The proposed system using the cloud computing service contributed to shortening the analysis time. Based on the experiments on Yodo River basin, independent 76 flood scenarios were concurrently processed using 12 virtual machines. The calculation shortened to 1/10. The speed-up effect was almost proportionate to the number of virtual machines. The proposed method is effective for simultaneous execution of independent scenarios. The proposed system is promising to apply to flood hazard mapping.

ACKNOWLEDGEMENTS

The application shown in this paper is DioVISTA Flood Version 3.0 by Hitachi Power Solutions Co., Ltd. In this work, we used a part of result in the task of building river models for Tokio Marine & Nichido Risk Consulting Co., Ltd.

REFERENCES

Domeneghetti, A., Vorogushyn, S., Castellarin, A., Merz, B., and Brath, A. (2013). Probabilistic flood hazard mapping: effects of uncertain boundary conditions, Hydrol. Earth Syst. Sci., 17, 3127-3140. https://doi.org/10.5194/hess-17-3127-2013

Glenis, V., McGough, A. S., Kutija, V., Kilsby, C. and Woodman, S. (2013). Flood modelling for cities using Cloud computing, J. of Cloud Computing, 2 (7). https://doi.org/10.1186/2192-113X-2-7

Ministry of Land, Infrastructure and Transport (MLIT). (2005). Flood Hazard Mapping Manual in Japan. http://www.icharm.pwri.go.jp/publication/pdf/2005/flood_hazard_mapping_manual.pdf

Ministry of Land, Infrastructure and Transport (MLIT). (2015). Manual of delineating Flood Inundation Risk Area, 4th Edition (in Japanese). http://www.mlit.go.jp/river/shishin_guideline/pdf/manual_kouzuishinsui_1507.pdf

Moya Quiroga, V., Popescu, I., Solomatine, D. P., Bociort, L. (2013). Cloud and cluster computing in uncertainty analysis of integrated flood models, 15 (1), 55-70. https://doi.org/10.2166/hydro.2012.017

Okada, T., McAneney, K. J., and Chen, K. (2011). Estimating insured residential losses from large flood scenarios on the Tone River, Japan – a data integration approach, Nat. Hazards Earth Syst. Sci., 11, 3373-3382. https://doi.org/10.5194/nhess-11-3373-2011

Vorogushyn, S., Merz, B., Lindenschmidt, K. E., and Apel H. (2010), A new methodology for flood hazard assessment considering dike breaches, Water Resour. Res., 46, W08541. https://doi.org/10.1029/2009WR008475

Yamaguchi, S. and Iwamura, K. (2007). Fast flood simulation method using dynamic DDM. IPSJ Transaction on Mathematical modeling and its application, 48 (SIG 6 TOM 17), 92-103. (in Japanese)

Yamaguchi, S. and Yamaho, S. (2016). Optimizing two-dimensional flood model with SSE and concurrent processing, Proc. of ICHE 2016, pp. 08-0001.

Table 1: Specifications of Yodo River model

Target river	Yodo River
Runoff model	Basin area: 4,392 km2 (Lake Biwa basin is excluded), Distributed, cell size = 100 × 100 m
River model	1 Mainstream, 28 tributaries 1D unsteady flow, cell size = 50 m
Inundation model	2D unsteady flow, cell size = 25 × 25 m
Dam model	7 dams, water volume-water level relation model, rule based gate control
Input boundary conditions	Precipitation data (radar observation, 1 km mesh, update every 30 min), Water level at river mouth (hourly)
Rainfall event for model calibration	2013-09-14 to 17 (Typhoon Man-yi, T1318)
Rainfall event for model validation	2015-07-16 to 19 (Typhoon Nangka, T1511) 2004-10-19 to 22 (Typhoon Tokage, T0423) Water level was compared at 5 gauging stations
Target period	2013-09-14 to 17 (72 hours)
Rainfall scenario	Twice strength of largest recorded rainfall (Typhoon Man-yi, T1318)
Levee failure scenario	36 points on right bank and 39 points on left bank (75 point in total)
Num. of scenarios	76 (75 levee failure + 1 overflowing)

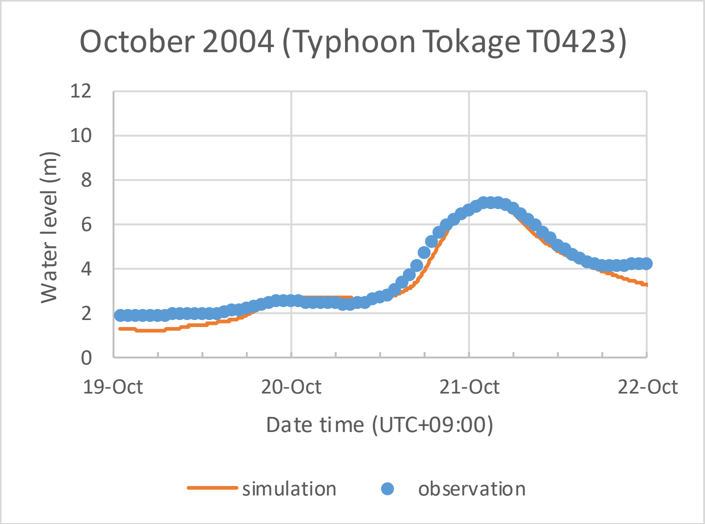

Fig. 4: River water level at Hirakata gauging station.

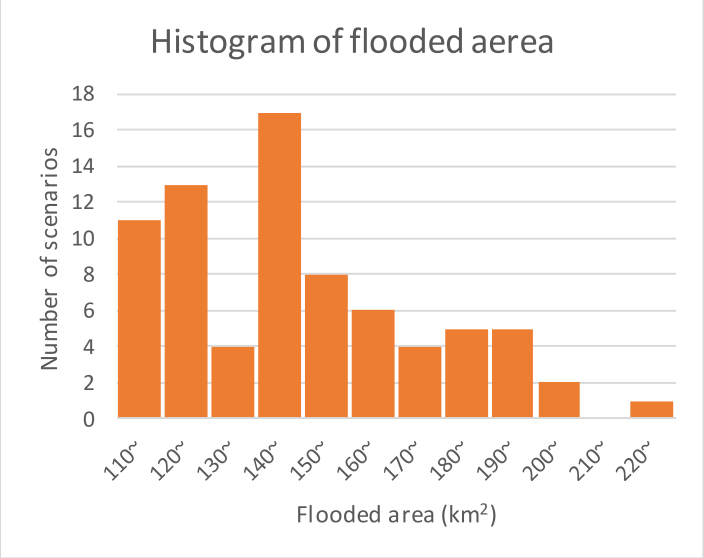

Fig. 5: Histogram of flooded area in 76 scenarios

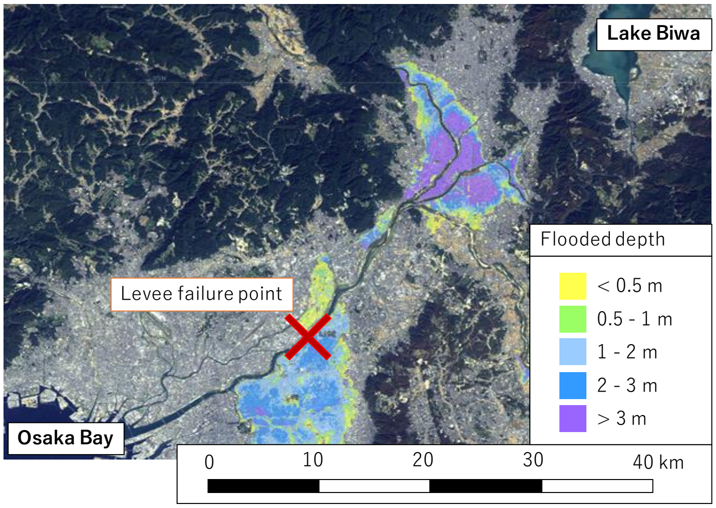

Fig. 6: Scenario with largest flooded area (levee failure at left bank, 19.2 km from river mouth, flooded area is 222 km^2)

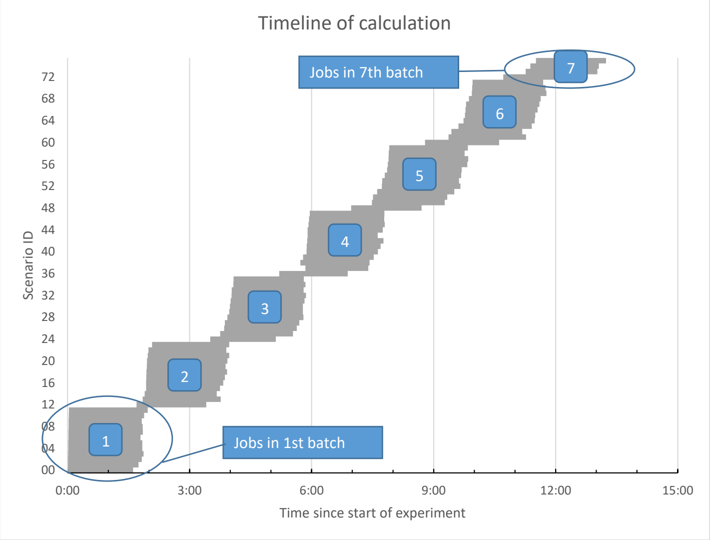

Fig. 7: Timeline of calculation process in Cloud.

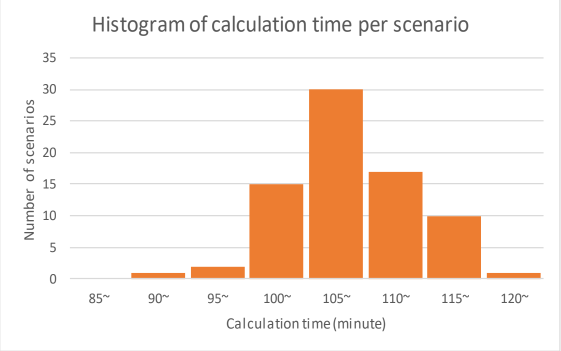

Fig. 8: Histogram of calculation time in 76 scenarios

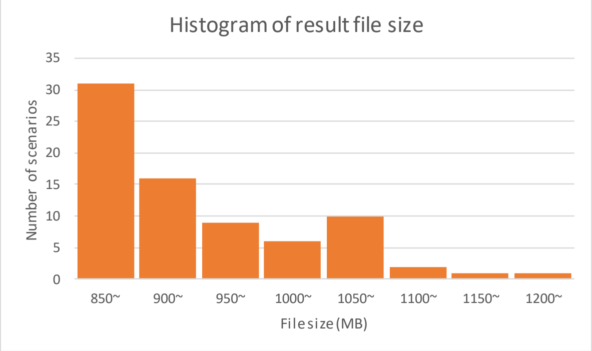

Fig. 9: Histogram of result file size in 76 scenarios