Performance Enhancement Architecture for Radiation Dose Computation in Integrated Decommission Process Assessment System†

BYUNG-SEON CHOI[1], JEIKWON MOON[1], SANG-WHA LEE[2] and HAE-SANG SONG[2,*]

[1]Decontamination & Decommissioning Research Division, Korea Atomic Energy Research Institute, 989-111 Daedeok-daero, Yuseong-gu, Daejeon 305-353, Republic of Korea
[2]Faculty of Information Technology, Seowon University,377-3 Musimseoro, Heungdeok-gu, Cheongju, Chungbuk 361-742, Republic of Korea

*Corresponding author: Fax: +82 43 2998710; E-mail: hssong@seowon.ac.kr

This paper proposes a computing architecture suitable for an integrated decommission process assessment system under development by Korea Atomic Energy Research Institute, which provides innovative simulation tools for the dismantling process of main components of nuclear power plants, including features of planning and simulation of dismantling processes to assess the task feasibility, cost estimation of processes, radiation dose estimation of workers. Among those, it is most safety-critical and time consuming to conduct radiation dose estimation of chemically contaminated and highly radioactive components during nuclear facility decommissioning, for usually it uses Monte Carlo simulation methods in order to get accurate results scarifying computing costs. The estimation time further increases as we need to iterate such time-consuming simulation runs for a lot of dismantling scenarios to get the total exposure of moving workers and even changing radioactive sources. Thus, to cope with this situation, we propose a hybrid cluster computing architecture exploiting commodity hardware for cost efficiency and maintainability. We also propose software architecture to effectively enhance the performance of the assessment in both real-time mode and batch mode with fault-tolerance. We implemented a pilot 8-node cluster as proposed and with experiments for a few cases we illustrated that the proposed architecture has the extensibility for the integrated decommission process assessment system.

Keywords: Radiation doseestimation, Nuclear power plant, Decommissioning assessment, Performance enhancement, Cluster computing.

INTRODUCTION

Korea Atomic Energy Research Institute (KAERI) has been developing an integrated decommission process assessment system, called MIDAS, to provide innovative simulation tools for the dismantling process of main components in nuclear power plants, which includes planning and simulation of dismantling processes to assess the task feasibility, cost estimation of processes, radiation dose estimation of workers, safety of workers by human error, waste management and cost estimation[1]. Inherently this is a kind of alternatives supporting tools for analysis to find an optimal dismantling process among a lot of dismantling scenarios. Among various estimations of the assessment system, the radiation dose estimation of chemically contaminated and highly radioactive components during nuclear facility of decommissioning is most safety-critical and time consuming for two reasons: It commonly uses Monte-Carlo (MC) method which is inherently most time-consuming task and there are a lot of scenarios to be estimated[2].

Thus, in this paper, we propose a computing architecture in hardware and software to enhance the performance for radiation dose computation for the assessment system to achieve two goals: to reduce a run time per simulation to execute as many simulation runs as possible in a batch mode to get a good through put in easy and fault-tolerant ways.

Most acknowledged radiation dose estimation tools, such as MCNP[3] and Geant 4[4], use Monte-Carlo particle simulation to get as accurate results as possible even paying high computing costs. Generally such Monte-Carlo simulation technology consists of three parts: sampling of random variables, computation and statistics of the results as shown in Fig. 1. In cases that the samples of random variables are considered independent and identically distributed, researches on parallelizing simulation tools have been done using various parallel computing technologies. Nevertheless, we still need to iterate many times for simulations since we have many different scenarios to be estimated in an easy and fault-tolerant way. Despite, an attempt like this has not been made so far. In this
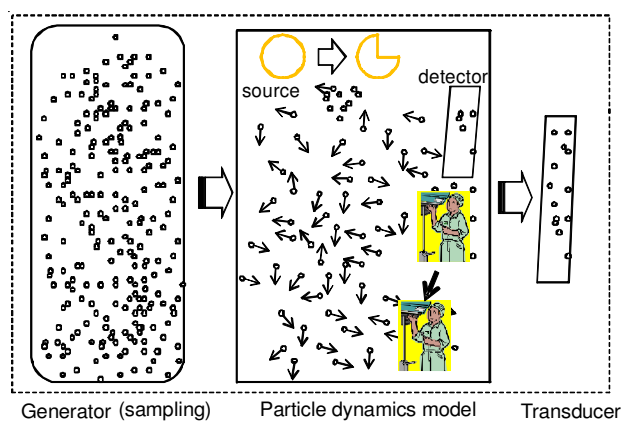
Fig. 1. Monte Carlo particle simulation with source and detector dynamically changing

paper, we propose a dedicated computing architecture optimized for our own application.

We clarify the problems, assumptions and design, followed by a proposed computing architecture satisfying the objectives. Then the suitability of the proposed architecture is illustrated from experimental results.

EXPERIMENTAL

Does estimation in the integrated decommission process assessment system: The target system with a new computing architecture includes dose estimation of exerted workers moving as planned in a process, who cut and transport highly radioactive source materials in the nuclear power plant to be dismantled. This dose estimation should be 'dynamic' in that the worker (detector) and the configuration of the radioactive materials (sources) are not in constant states. Fig. 1 shows the concept of our application using Monte Carlo simulation, where a source is represented by circle and detector trapezoid. Both of them are moving as a dismantling process undergoes. Moreover we have many alternatives of dismantling processes to be assessed. It is assumed that once a dismantling process is decided as a candidate, necessary information of the process is stored in a scenario database, which specifies way points of the workers for doing jobs and the configuration changes of the radioactive source due to the result of the jobs with time elapse.

Design objectives: The aim of the paper is to propose a hardware-software architecture suitable for the needs with the following design criteria.

Extensibility: Computing power could be easily increased as needed with a reasonable cost and acquisition feasibility to meet the needs.

Maintainability: The system could be supported for long-lasting maintenance.

Portability: Software could be easily ported with little modification as possible.

Cost: The lower the cost of hardware and software, the better.

Current technologies to enhance performance: State-of-the-art dose estimation software relies on the simulation of random walking particles, such as proton, photon, electron and so on, interacting with surrounding materials including

workers for instance. Specifically it simulates geometry of facility, source particles generation, tracking of them, interaction with intermediate materials and finally detector absorption. As a dose estimation software basically depends on Monte-Carlo method, it is relatively adequate to be parallelized to speed up simulation using divide-and-concur approach. Much effort has been done using up-to-date parallel computing technologies surveyed in Table-1[5,6].

TABLE-1
STATE-OF-THE-ART PARALLEL COMPUTING TECHNOLOGIES
AND THEIR APPLICABILITY TO MC SIMULATION
(○: PORTED, △: PARTIALLY PORTED)

Hardware	Architecture	MCNP	Geant4	Software
Multi core CPU	Multi-threads	○	△	Open MP
GPU	SIMD kernel	△	△	CUDA, OpenCL
Cluster	Master-slaves	○	○	MPI

Most of the technologies maximally utilize hardware, multi-core CPU, GPU (Graphic Processing Unit) and cluster of computer nodes connected by high speed network. As shown in Table-1, MCNP code has been utilized in most of the current parallel computing technologies, while Geant 4 is known to be rather tardy. To speed-up of a single simulation run can be solved by the above technologies. Thus, in this paper, we focus on enhancing the performance of multiple scenario evaluation by introducing a new technology, called MapReduce[7].

Proposed computing architecture

Multiple-scenario computing model: The objective of this computing model is to get maximum throughput for a bunch of simulation jobs and the throughput should be enhanced as the number of node is extended. To clarify the notion, a scene is defined as a configuration of workers and radiation sources on a fixed position. A scene is changing whenever any worker changes his position or any part of a radiation source is cut and moved. A scenario is a sequence of such scenes. Fig. 2 depicts the computing tasks for multiple scenes. Assuming that a dismantling scenario is stored by the assessment system (MIDAS) in the scenario database, first we pick a scenario to make it a series of scenes whenever important configuration changes take place due to either workers' movement or source shape transformation by cut and transportation, which will be stored in the scene database. A scene contains a set of sources and detectors placed in the work area for a simulation. Each scene is processed by independently running Monte-Carlo simulation process, whose result is sent to the merger that collects dose data of workers from all the scene simulations.

Proposed hardware architecture: Taking into account the design objectives and state-of-the-art parallel computing technologies surveyed before, we establish the hardware architecture to be an extendable cluster that is composed of nodes connected by network switch(es), each node of which has multi-core processor(s) and GPU(s) with storage(s) to exploit all the parallelizing technologies, OpenMP, OpenCL, MPI. Unlike dedicated supercomputers, all the parts of a node consist of commodity components. Thus we conclude that cluster is the best suitable hardware architecture for our application because it satisfies all the design objectives *i.e.*, extendibility, portability, maintainability and cost.
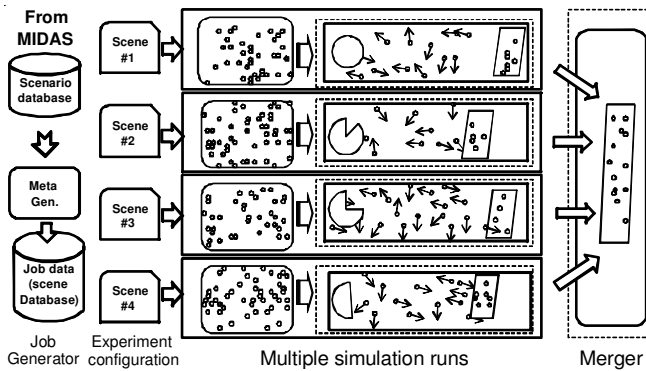
Fig. 2. Batch multiple-scenario computing model for multiple scenes

Hadoop map reduces treaming framework: Map reduce (MR) is a simple computing method for processing big data of <key, value> pairs. A map function takes a key-value pair as input and processes the pair to produce a resulting key-value pair as in equation (1). A reduce function is similar except that the input is a key-value list (2), where K_i and V_i are data types of keys and values, respectively, [×] represents a list.

$$\text{map}: K_1 \times V_1 \rightarrow K_2 \times V_2 \qquad (1)$$

$$\text{reduce}: K_2 \times [V_2] \rightarrow K_3 \times V_3 \qquad (2)$$

A map reduce computing framework is an enabling technology for the map reduce computing method. It allows user to plug-in map and reduce functions, respectively, providing a big distributed storage system. Hadoop is one of such framework, released by Apache Foundation for free, that consists of two parts: map reduce computing framework and big data storage called HDFS (Hadoop Distributed File System)[7]. Specifically, we are interested in Hadoop streamingthat comes with the Hadoop distribution, which allows you to create and run map reduce jobs with any executable or script that comes with standard input and output as depicted in Fig. 3.
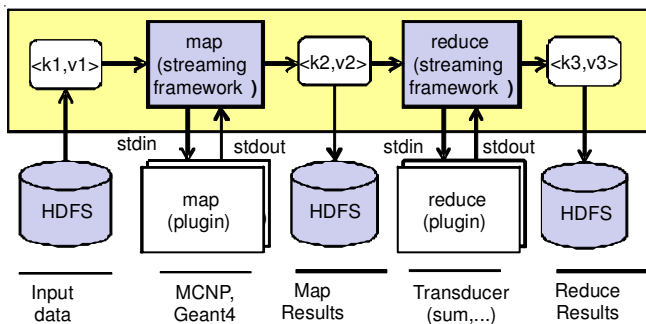

Fig. 3. Hadoop map reduce streaming framework

Adoptionof Hadoop map reduce streaming framework: Using Hadoop streaming framework, multiple simulation runs of scenes of a scenario in Fig 2 are executed by multiple mapper tasks as shown in Fig. 3, where an input key-value of a mapper, stored in HDFS, is <scene id, configuration data > while the output key-values are <scene id, detector id, dose values> that are collected by the Merger executed by a reducer task. The mapper tasks can be executed concurrently on distributed nodes in the cluster. Using the streaming framework, any single threaded Monte Carlo simulation code could have benefit by the architecture with little modificationonly

in terms of input and output aspects to make the outputs stored in HDFS instead of local file system.

RESULTS AND DISCUSSION

To show the feasibility of the idea, first we construct to experiment with all the features of parallel and distributed computing architecture a 8-node pilot cluster rack with each node having i7 quad-core CPU, NVIDIA GPUs and 16GB RAM clustered by a Gigabit Ethernet switch, which could be extended easily. Installed in the cluster are Ubuntu Linux OS, MPI, OpenMP, OpenCL programming environment and Hadoopframework. MCNP6 released by RISCC is also installed.

The MCNP simulation model in the experiment is RPV (Reactor Pressure Vessel) of Kori-I nuclear power plant in Korea, planned to retire in a decade. The number of particles to be generated is 10^6, with the simulation time of about tens of seconds. For comparison, we first run a stand-alone MCNP6 simulation. Then, issuing jobs with increasing number of tasks to the Hadoop framework, we observe the elapsed time and total CPU time of each job.

The experiment is designed to see if we can use Hadoop MapReduce framework for a bunch of MCNP simulation runs, though other parallel versions of simulation could be executed on the cluster. To fit into the framework, MCNP6 is plugged in asa mapper on the Hadoop framework as shown in Fig. 3. To simplify the experiment we use map-only streaming framework, where the result of MCNP6 can be directly stored in HDFS. Maximum number of map tasks is set to eight to see the tendency of the performance enhancement.

The experimental results shown in Fig. 4 illustratethat the elapsed timesequal up to eight tasks, for each task is distributed and run on different node of the cluster independently. However if the number of tasks exceed the number, the framework queues and executes a bunch of eight tasks at a time.We can see the elapsed time of Hadoop job is slightly higher than the pure batch job of MCNP issued manually, which is due to the overhead of the Hadoop framework that is negligible compared to the run time of the mappers. As expected, total CPU time is increasing in proportion to the number of tasks as expected. We can easily know that the results shows that if we extend the number of computing nodes of the cluster, the rate of increment of the elapsed time will belowered, while it could gain all the benefits Hadoop framework provides: load balancing, fault tolerance, big data storage, compared to manual batch management.

Conclusion

This paper deals with an adequate computing architecture for dose estimation of chemicallycontaminated and highly radioactive components during nuclear facility decommissioning suitable for integrated decommission assessment system, which has a lot of scenarios to be estimated. To cope with the situation, we proposes a computing architecture in hardware and software to enhance the performance for the radiation dose computation. To execute many simulation runs like MCNP to get a good throughput in easy and fault-tolerant ways, we adopted map reduce framework, especially Hadoop streaming framework. We implemented a pilot cluster with commodity
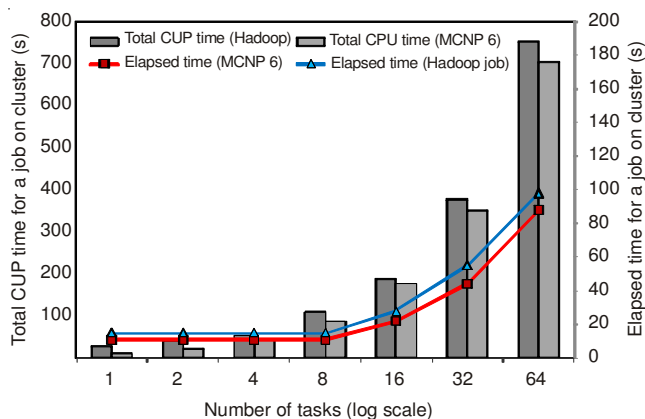
Fig. 4. Comparison of computing time with varying number of tasks

hardware and constructed relevant software for the experiment. As the result, we shows that the proposed architecture has the feasibility and extensibility to enhance the performance of dose estimation of the decommission system. For future work, we need to devise the interface between MIDAS and the proposed computing architecture.

REFERENCES

1. H.S. Song, D. Hyun, G.H. Kim, B.S. Choi and J. Moon, *Asian J. Chem.*, **25**, 7041 (2013).
2. N. Metropolis and S.M. Ulam, *J. Am. Stat. Assoc.*, **44**, 335 (1949).
3. J.T. Goorley, M.R. James, T.E. Booth, F.B. Brown, J.S. Bull, L.J. Cox, J.W. Durkee Jr, J.S. Elson, M.L. Fensin and R.A. Forster, Initial MCNP6 Release Overview-MCNP6 version 1.0 Technical Report No. LA-UR-13-22934, Los Alamos National Laboratory (LANL) (2013).
4. S. Agostinelli, J. Allison, K. Amako, J. Apostolakis, H. Araujo, P. Arce, M. Asai, D. Axen, S. Banerjee, G. Barrand, F. Behner, L. Bellagamba, J. Boudreau, L. Broglia, A. Brunengo, H. Burkhardt, S. Chauvie, J. Chuma, R. Chytracek, G. Cooperman, G. Cosmo, P. Degtyarenko, A. Dell'Acqua, G. Depaola, D. Dietrich, R. Enami, A. Feliciello, C. Ferguson, H. Fesefeldt, G. Folger, F. Foppiano, A. Forti, S. Garelli, S. Giani, R. Giannitrapani, D. Gibin, J.J. Gómez Cadenas, I. González, G. Gracia Abril, G. Greeniaus, W. Greiner, V. Grichine, A. Grossheim, S. Guatelli, P. Gumplinger, R. Hamatsu, K. Hashimoto, H. Hasui, A. Heikkinen, A. Howard, V. Ivanchenko, A. Johnson, F.W. Jones, J. Kallenbach, N. Kanaya, M. Kawabata, Y. Kawabata, M. Kawaguti, S. Kelner, P. Kent, A. Kimura, T. Kodama, R. Kokoulin, M. Kossov, H. Kurashige, E. Lamanna, T. Lampén, V. Lara, V. Lefebure, F. Lei, M. Liendl, W. Lockman, F. Longo, S. Magni, M. Maire, E. Medernach, K. Minamimoto, P. Mora de Freitas, Y. Morita, K. Murakami, M. Nagamatu, R. Nartallo, P. Nieminen, T. Nishimura, K. Ohtsubo, M. Okamura, S. O'Neale, Y. Oohata, K. Paech, J. Perl, A. Pfeiffer, M.G. Pia, F. Ranjard, A. Rybin, S. Sadilov, E. Di Salvo, G. Santin, T. Sasaki, N. Savvas, Y. Sawada, S. Scherer, S. Sei, V. Sirotenko, D. Smith, N. Starkov, H. Stoecker, J. Sulkimo, M. Takahata, S. Tanaka, E. Tcherniaev, E. Safai Tehrani, M. Tropeano, P. Truscott, H. Uno, L. Urban, P. Urban, M. Verderi, A. Walkden, W. Wander, H. Weber, J.P. Wellisch, T. Wenaus, D.C. Williams, D. Wright, T. Yamada, H. Yoshida and D. Zschiesche, *Nucl. Instrum. Methods Phys. Res. A*, **506**, 250 (2003).
5. B. Yang, K. Lu, J. Liu, X.P. Wang and C.Y. Gong, GPU Accelerated Monte Carlo Simulation of Deep Penetration Neutron Transport, 2nd IEEE International Conference on Parallel Distributed and Grid Computing (PDGC), Dec. 2012, Solan, pp. 899-904 (2012).
6. G. Cooperman, V.H. Nguyen and I. Malioutov, Parallelization of Geant4 Using TOP-C and Marshalgen, Fifth IEEE International Symposium on Network Computing and Applications, July 2006, Cambridge, MA, pp. 48-55 (2006).
7. http://hadoop.apache.org/docs/r1.1.2/streaming.html, retrieved in (2013).