Common platform of Monte Carlo dose calculation on universal grid interface with Geant4 based particle therapy simulation framework

G Iwai¹, W Takase¹, T Aso¹,², Y Watase¹, T Sasaki¹, T Akagi³, T Yamashita³, Y Maeda⁴ and T Nishio⁵

¹ Computing Research Center, High Energy Accelerator Organization, 1-1 Oho, Tsukuba, Ibaraki 305-0801, Japan
² Department of Electrical Engineering, Nagaoka University of Technology, 1603-1 Kamitomioka, Nagaoka, Niigata 940-2188, Japan
³ Hyogo Ion Beam Medical Center, 1-2-1 Koto, Shingu-cho, Tatsuno, Hyogo 679-5165, Japan
⁴ Fukui Prefectural Hospital Proton Therapy Center, 2-8-1 Yotsui, Fukui, Fukui 910-8526, Japan
⁵ National Cancer Center Hospital East, 6-5-1 Kashiwanoha, Kashiwa, Chiba 277-8577, Japan
E-mail: aso@nagaokaut.ac.jp

Abstract. While Monte Carlo (MC) simulation is believed to be the most reliable method of dose calculation in particle therapy, the simulation time is critical in attaining sufficient statistical accuracy for clinical applications. Therefore, parallelization of simulations is essential. This paper describes a common platform of MC dose calculation in grid-distributed computing environments. The platform is flexible and effective for dose calculation in both clinical and research applications for particle therapy. The platform consists of the universal grid interface (UGI) and the Geant4-based particle therapy simulation framework (PTSIM). The UGI, written in Python, provides a command-line interface for job submission, file manipulation, and monitoring in multiple-grid middleware environments. The PTSIM is a single software application for modeling a treatment port with patient data obtained from CT images. The common platform was constructed in grid computing environments using the computing resources in five institutions. The platform utilized these resources through the NAREGI grid middleware under UGI to provide stable computing resources and a common environment for MC dose calculation in particle therapy.

1. Introduction

Geant4 [1, 2] is a software toolkit for simulating the interaction of particles in matter. Monte Carlo (MC) simulation using Geant4 takes into account the physics of particle interactions by using theoretical models or experimental cross-section data. Therefore, MC simulation is believed to be the most reliable and useful application in the design of particle treatment facilities and the quality assurance of treatment plans. However, the simulation time is critical in performing dose calculation with sufficient statistical accuracy [3]. There are many techniques to reduce simulation time such as variance reduction techniques [4, 5] and parallelization using grid computing environments or computer graphics processing units (GPUs) [6, 7, 8]. The parallelization of simulations in grid computing environments is a straightforward way to gain...
computation speed with validated physics processes according to fundamental physics principles. Therefore, parallelization of simulations with high computing resources is essential.

Many reports about medical grid projects have been reported over the past few years [9, 10, 11]. These projects developed regulated grid environments for sharing bioinformatics resources and analyzing standardized data such as CT and MRI images for patient diagnostics. The development of such environments was boosted because the data format is well standardized. In contrast to these grid environments, the development of a common platform for MC dose calculation has had the problems of high simulation costs and nonstandard data formats dependent on machine parameters. Furthermore, the confidential information of the machine parameters belongs to the vendor, and so the parameters are not open to the public. Therefore, a common platform is desirable to provide high-performance computing resources as well as a policy of data management.

In this paper, we describe our strategy for a common platform of MC dose calculation in multiple-grid environments. The platform should be flexible and effective for dose calculation in clinical applications, research, and education. We discuss the operational strategy in terms of simulation time and data resource sharing between members in order to facilitate the widespread use of MC dose calculation.

2. Materials and methods

2.1. Particle therapy simulation framework

The Geant4 based particle therapy simulation framework (PTSIM) [3, 12, 13] is a single software application program for modeling a treatment port with patient data obtained from CT images. Online commands are provided for modifying the configuration parameters of the treatment port, so that users do not need to develop a programming code.

In hospital information systems, CT images and treatment plans are exchanged in the Digital Imaging and Communications in Medicine (DICOM) format [14] and the extension of DICOM format for radiotherapy (DICOM-RT). The dataset of a treatment plan is called RTIonPlan and includes the configuration parameters of the treatment machine. The PTSIM has a DICOM interface for CT images and an RTIonPlan for coordinating with hospital information systems. Therefore, the PTSIM performs dose calculation by using the CT images and the RTIonPlan data to ensure the same conditions in the treatment planning system and the treatment machine. Although DICOM-RT is well established, the use of RTIonPlan is still limited because the parameters might not have the same meaning in different treatment facilities. The current PTSIM provides an online command for switching the RTIonPlan interface in accordance with the particle therapy facility. Therefore, a single PTSIM program is used among users in different treatment facilities.

Additionally the PTSIM supports an interface for reading and writing track information in the IAEA phase space data format [15]. The data include the particle type, kinetic energy, position, and direction of each track at a certain plane in the simulation volume. The data are prepared in advance by simulating particles from the beam extract window to the isocenter in the treatment port. The track information of photons, electrons, positrons, neutrons and protons are recorded. The simulation time is reduced by starting simulations from the phase space data recorded in front of the patient area. All particles in the data are generated for the simulation.

2.2. Universal grid interface

The universal grid interface (UGI) [13] provides a platform for sharing affordable CPU power through multiple-grid environments. It provides a command-line interface and Web interface for job submission, file manipulation, job monitoring, and a simple tool for result checking. The UGI is developed on top of Python binding of the Simple API for Grid Applications (SAGA)
[16, 17] framework, the Resource Namespace Service (RNS) [18], and the XML RPC [19]. By using plug-in adapters in the UGI, connections to multiple-grid environments are established by various grid middleware packages such as NAREGI [20] and gLite [21]. Figure 1 illustrates a conceptual diagram of the multiple-grid computing environment using UGI. We assumed two usage scenarios of the UGI environment for the PTSIM. In the first scenario, the PTSIM is preinstalled in each grid environment by administrators, so that users run the PTSIM in the UGI environment. In the second scenario, the UGI handles the transfer of the PTSIM code and data from the UGI portal site, and then builds and runs the PTSIM. The first scenario is applied as a regular operation for PTSIM application users to reduce the maintenance costs, whereas the second scenario is for PTSIM developers who have to examine the modified PTSIM code before release. UGI environments have been examined for both scenarios.

2.3. Common platform and operational strategy

The common platform consists of the PTSIM and the UGI. It utilizes the multiple-grid computing environment to provide high computing resources and to share data resources with collaborating members.

For clinical applications, we mask private information with DICOM data to assure a patient’s anonymity. Access to DICOM data is managed according to the permission flag of the file under a set of resource-sharing rules and conditions.

The PTSIM needs the machine parameters of the treatment port to ensure the same condition for the treatment machine. However, the machine parameters have to be kept confidential except for authorized members because the information belongs to the vendor of the treatment port. This restriction limits the widespread use of the PTSIM for research and educations.

We propose to construct a database of particle phase space data in the IAEA format for each configuration of validated beam delivery systems. The database allows hospital members, vendors, and research institutes to share the simulation of a beam delivery system for quality assurance, research, and education without exploring the machine parameters. Figure 2 shows the conceptual diagram of the common platform at the point of data management. For example, the phase space data recorded at the upstream of the multi-leaf collimator are used with the machine parameters for the verification of treatment plan in clinical applications by authorized members. The phase space data are also applicable without the machine parameters for researchers who are engaging to develop new beam devices in the treatment head. For the educational purpose, the phase space data are applied for simulating dosimetry experiments to optimize the experimental conditions.

Since all beam devices are referred to as beam modules in the PTSIM, only beam modules selected by users are placed in the beam line and configure a treatment port. By using the phase space data, the beam modules of the beam delivery system are removed from the beam line without modifying the PTSIM code. Therefore, the machine parameters are separated from the distribution of the PTSIM code. In addition, starting simulations by using the phase space data consolidates redundant simulations of the beam delivery system that had been performed by individual users. The validations and productions of the phase space data are performed by the authorized members such as medical physicists in treatment facilities and the PTSIM developers. Consequently, the use of phase space data reduces simulation time of database users.

3. Results and discussions

3.1. Simulation time

In our previous research [13], we examined the simulation time by simulating a 150 MeV proton therapeutic beam. The $9 \times 10^8$ primary protons were simulated from the beam extract window to the water phantom from CT images at the isocenter. The configuration parameters of the treatment port were taken from RTIonPlan data, which mimic the treatment plan for a
patient. The treatment head included a multi-leaf collimator, a bolus compensator, and a patient collimator.

In this research, we repeated simulations in the shared central computing system in KEK (KEKCC) with the same condition used in our previous research. We compared average simulation times with and without phase space data. The phase space data were recorded at the entrance of patient area, which corresponds to the upstream of the multi-leaf collimator. By separating the job into 500 tasks, the average simulation time of each task without phase space data was approximately 100 minutes, while it was 37 minutes with phase space data.

### 3.2. Data resource sharing

The phase space data need to be prepared for each set of typical treatment parameters, e.g., combination of beam energies, spread-out Bragg peak (SOBP), and range shifter. The data size was approximately 12 GB for $9 \times 10^8$ primary protons in the example treatment parameters. If we assume a treatment port in the Fukui prefectural hospital proton therapy center (FPHPTC) in Japan, the gantry treatment port has a variety of three beam energies, twelve SOBPs, and five range shifters. Currently, 146 combinations of those parameters are registered in the treatment planning system for clinical applications in the gantry treatment port. Therefore, 146 times more disk space and CPU hours are needed. This amount corresponds to a total of approximately 2 TB disk space and 120,000 CPU hours. These simulations insist large scale high computing resources for the validation of the beam delivery system with confidential machine parameters. Apparently, such preparation of phase space data by individual users is inefficient. Therefore, the common platform with the operational strategy is suitable to make and share IAEA phase space data among members for efficient usage of computing resources.

### 4. Conclusions

The common platform of MC dose calculation in multiple-grid environments provides stable computing resources and a common environment particle therapy in clinical applications, research, and education, while protecting the information of machine parameters and patient privacy.
Acknowledgments
The authors would like to thank the members of the National Institute of Informatics, and the Global Scientific Information and Computing Center of the Tokyo Institute of Technology. Development of the PTSIM was supported in part by the Japan Science and Technology Agency as part of the Core Research for Evolutional Research and Technology program. The present study was supported in part by grants from JSPS KAKENHI (No. 23540350 and No. 23390308).

References
[1] Agostinelli S et al. 2003 Geant4, a simulation toolkit, Nucl. Instr. Meth. A 506 250
[(287,258),(330,335)
[2] Allison J et al. 2006, Geant4 developments and applications IEEE Trans. Nucl. Sci. 53 No. 1 270
[3] Akagi T et al. 2011 The PTSim and TOPAS projects, bringing Geant4 to the particle therapy clinic Prog. in Nucl. Sci. Technol. 2 912
[4] Kawrakow I et al. 2000 Investigation of variance reduction techniques for Monte Carlo photon dose calculation using XVMC Phys. Med. Biol. 45 2163
[5] Ma C et al., 2002 A Monte Carlo dose calculation tool for radiotherapy treatment planning Phys. Med. Biol. 47 1671
[6] Kohno R et al. 2011 Clinical implementation of a GPU-based simplified Monte Carlo method for a treatment planning system of proton beam therapy Phys. Med. Biol. 56 N287.
[7] Hissoiny S et al. 2011 GPUMCD: a new GPU-oriented Monte Carlo dose calculation platform Med. Phys. 38 754
[8] Xun J et al. 2012 GPU-based fast Monte Carlo dose calculation for proton therapy Phys. Med. Biol. 57 7783
[9] Blanchet C et al. 2006 Integrating bioinformatics resources on the EGEE grid platform Sixth IEEE Int. Sympo. on CCGrid 48
[10] Eschenbach A C V et al. 2006 calBIG: Cancer Informatics Vision Cancer Informatics 2 22
[11] Krefting D A et al. 2009 MediGRID: Towards a user friendly secured grid infrastructure Future Generation Computer Systems 25 326
[12] Aso T et al. 2007 Geant4 based simulation framework for particle therapy system IEEE NSS/MIC Conf. Rec. NSS’07 2564
[13] Aso T et al. 2012 Extension of the particle therapy simulation framework to hospital information systems and multi-grid environments IEEE Int. Conf. on Comp. and Eng. (CSE) IEEE Comp. Soc. 229
[14] DICOM Digital Imaging and Communications in Medicine The National Electrical Manufacturers Association [Online] http://medical.nema.org/
[15] IAEA INDC (International Nuclear Data Committee) phase-space database for external beam radiotherapy 2005 INDC(NDS) 0484 1.
[16] SAGA Simple API of grid application [online] http://saga.cce.lsu.edu/
[17] OGF Open Grid Forum OGF document:OGF-R-O.90 [Online] http://www.ggf.org/
[18] RNS Resource Namespace Specification OGF document: Recommendation [Online] http://www.ggf.org/
[19] XML-RPC [Online] http://xmlrpc.scripHng.com/
[20] NAREGI National Research Grid Initiative [online] http://www.naregi.org
[21] gLite Lightweight middleware for grid computing [online] http://glite.cern.ch/