Abstract

The hybrid applications are needed in dental research because the result of their operation could be predicted by both modeling of an oral tract and more than two kind of physical simulations. However, it takes long time to perform the hybrid applications, cost a lot of money and needs some experience to deal with them. This paper proposes a new execution procedure of two applications which can reduce the implementation time by considering parallel efficiency of those applications. It can determine which attitude should be taken, sequential or separate execution of two different applications related with each other. Following the execution procedure, the dental hybrid application was performed. As a result of that, desirable pairs of the number of CPUs allocated to each simulation could be found. This means that the adequate procedure should be considered before execution of the hybrid applications, in order to play a compute power producer for dental clinics and hospitals.

1 Introduction

The subject of this paper is efficient distribution of computational resources to perform two kinds of simulations on dental treatments. In order to supply computational resources enough to perform those simulations on dental treatments, a combination among more than two kinds of simulations is needed, such as a computational fluid dynamics (CFD) and Computational Aero Acoustics (CAA). When the hybrid simulation is performed, a total computation time should be concerned because it is the key to distribute the dental applications to each local dental office. In terms of a total cost to perform the dental application, a load average of whole CPUs should be kept at high level. In order to integrate those applications, job submissions, streaming file transfers among several applications, parameter deliveries, scheduling [4] are required. GridLab [1] might be the suitable framework to realize the integration way of some applications, which can wrap some legacy applications with Application Program Interface (API).

In the near future, dentists and patients will require not only a single simulation, such as a biomechanical simulation, but also the hybrid simulation which is combined among CFD, a biomechanical simulation and CAA. However, it is clear that the hybrid application needs a large amount of computational resources. One of the solutions about the resource requests is to gather or allocate CPUs required, but the total amount of CPUs is limited. Another solution about them is to make
the allocation of CPUs to each application balanced according to the parallel efficiency.

This paper reports on the design, implementation and performance evaluation of the optimal CPU resource allocation according to the total computation time on the dental application combined CFD and CAA as a part of DentGrid system.

2 Overview of DentGrid system

Currently, a conventional way of dental treatment tends to shift its criterion from experience to evidence. Evidence based dentistry is the term that specify this trend clearly. Patient’s medical and dental records, stored and retrieved in standardized digitized forms independently, are valuable data for the patients themselves and precious resources for evidence based dentistry, equivalently. For example, in some university dental hospital, electronic patient record systems have already been deployed and deployed with the medical markup language to exchange patient information between medical institutions. In these hospitals, they have a data ware-house (DWH) independent form the hospital information system for basic hospital practices such as an insurance claim system, an order entry system and PCAS (Picture Archive and Carry System). The data in DWH supports the decision in these hospitals both clinically and economically. Hence, the DWH thought to be an important solution to decrease the total cost of health care.

DentGrid has been promoted as a next generation health care system [2] based in the concept that the system will supply computational power to each dental hospital and clinic to perform hybrid applications efficiently. On DentGrid, CFD and CAA already had performed as a hybrid application [6] that could predict the pronunciation changes after insertion of dental prosthetics or orthodontic treatments. As a single application, the CFD simulation on a super computer SX-5 (NEC Corp.) revealed the sound source distribution of sibilant /s/ [6]. As for CAA simulation, pressure fluctuation at distant points could be required from air flow near the teeth by Lighthill’s theory [5].

The grid technologies is the key to acquisition and reallocation of patient data in dental hospitals and offices. On the other hand, it is difficult to expect for dentists to send jobs to the super computer, receive the results in thier office and visualize them in order to use them for thier patients. In a grid environment, it might be easier to perform these producers through web based applications that are accessible from the grid portal. The goal of DentGrid is that the dentists, can logged through grid portal, treat the complex network of super computers as their handy tool and complete a ton of their jobs without tears. In this study, a computational resource balancer is going to be deployed in DentGrid.

3 Design and Implementation

As an input data to the CFD simulation, a patient’s oral truct model data taken by Magnetic Resonance Imaging (MRI) was used. Air flow of the model was acquired by the CFD simulation in chronological order. Every results of the CFD simulation was input to the CAA simulation. Sound propagation was simulated by the CAA simulation. In our study, the hybrid application which consist of the CFD and the CAA was designed to be performed by an execution sequence, which might enable the total computation time of CFD and CAA simulation to be shortened. Two kind of execution sequences was considered; 1) the CAA simulation starts after all results of the CFD simulation have been acquired, 2) the CAA simulation starts before all results of the CFD simulation have not been acquired yet.
NaSt3DGP [3] was used as a CFD application in this paper, which is C++ implementation of a solver for the incompressible, time-dependent Navier-Stokes equations in three dimensions,

\[
\frac{\partial u}{\partial t} + u \cdot \nabla u = \frac{g}{Fr} - \nabla p + \frac{1}{Re} \Delta u, \quad x \in \Omega, \quad 0 \leq t \leq T_{\text{fin}}
\]

(1)

\[
\nabla \cdot u = 0,
\]

(2)

where \( Re \) is the dimensionless Reynolds-number, \( Fr \) modified Froude-number, \( u \) velocity vector, \( t \) time, \( g \) force, \( p \) pressure, \( x \) field and \( \Omega \) control volume. NaSt3DGP can be executed parallel with based on the message passing interface MPI. In this paper, it was compiled on the Redhat Linux 9.0 with libc 2.3.2, gcc 3.2, SCore 5.4.

CAA application code was implemented by S.Maeda(2005). The sound radiation from a fluid flow called as Aerodynamic sound firstly described by M.J.Lighthill [5]. The CAA code was based on the lighthill’s theory. When \( \rho \) is the density, \( p \) the pressure, \( v_i \) the velocity in the \( x_i \) direction, \( c_0 \) the speed of sound in the uniform medium, the Lighthill’s equation is

\[
\left( \frac{1}{c_0^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \left[ c_0^2 (\rho - \rho_0) \right] = \frac{\partial^2 T_{ij}}{\partial x_i \partial x_j},
\]

(3)

where \( T_{ij} \) is called the Lighthill stress tensor. This is Lighthill’s acoustic analogy. \( T_{ij} \) can be simplified in nonviscous fluid and thermal insulation as follows,

\[
T_{ij} = \rho v_i v_j.
\]

(4)

When the sources of matter are not so concentrated, the density field is given by the volume integral of terms as follows,

\[
\rho(x, t) - \rho_0 = -\frac{1}{4\pi c_0^2} \frac{\partial}{\partial x_i} \int \frac{1}{r} X_i(y, t_r) d^3 y + \frac{1}{4\pi c_0^2} \frac{\partial^2}{\partial x_i \partial x_j} \int \frac{1}{r} T_{ij}(u, t_r) d^3 y
\]

(5)

, where \( r \) is the distance from the source, and \( X_i(y, t_r) \) per unit volume per unit time is introduced at \( x \) at time \( t \). However, \( r = |x - y| \), \( t_r = t - \frac{|x - y|}{c_0} \), where \( x \) represents an observation point, \( y \) sound source point. The code of CAA simulation is written in fortran77, which can be performed by using the parallel message passing interface MPI. CAA needs the every result data from CFD simulation, because the lighthill’s equation needs \( p_0 \) and \( X_i \) at \( t \). CAA calculation can be executed if one time step data could be obtained from CAA.

It was examined which strategy should be taken, first, separately execution of CFD and CAA, and second, sequential execution of them (Fig.1). In the separately execution, total calculation time is represented as,

\[
T_{\text{ALL}} = (T_{\text{CFD}} + T_{\text{CAA}}) \times n,
\]

(6)

where \( T_{\text{ALL}} \) is the total calculation time, \( T_{\text{CFD}} \) calculation time cost a time step of CFD, \( T_{\text{CAA}} \) calculation time cost a time of CAA, and \( n \) the number of time steps needed for the hybrid application. In the sequential execution, when \( T_{\text{CFD}} > T_{\text{CAA}} \),

\[
T_{\text{ALL}} = T_{\text{CFD}} \times n + T_s + T_{\text{CAA}}
\]

(7)

and when \( T_{\text{CFD}} < T_{\text{CAA}} \),

\[
T_{\text{ALL}} = T_{\text{CFD}} + T_s + T_{\text{CAA}} \times n
\]

(8)
Figure 1. Separate and sequential type of calculation procedure.

where $T_s$ is time cost to send from the result data of CFD nodes to CAA nodes.

The input data which was made from MRI data was converted to nav file by NaSt3DGP, which can be described precise conditions, such as boundary conditions, mesh size, solver type, and so on. In this examination, time steps was set at 20,000, max of $\Delta t$ 0.00003 sec, inflow 1500 Pascal at a larynx, Reynolds-number 2500.

It was assumed that the space computed by the CFD is sound source. The observation point of sound propagation was set at 1 meter distance from sound source. Input data was obtained from CFD nodes as pressure and velocity field after CAA application read a nav-file of NaSt3DGP. The CAA code outputs density and pressure of the observation point.

Biogrid Base System was used as test bed of DentGrid project. Grid system 2 which is one of the biogrid base system has 78 nodes(Express 5800/BladeServer,NEC), Pentium III-S 1.4GHz and 1GB main memory. Total storage is 12TB.

In order to clock the calculation time which depends on the number of CPUs, five time steps calculations was performed repeatedly with adding a CPU by using Score5.8. As a result of those trials, $T_{CFD}$ and $T_{CAA}$ could be obtained.

4 Results

In order to examine optimal approach to perform the hybrid application, the effectiveness with parallel computations of both CFD and CAA was evaluated as a first phase (Fig.2). Next, $T_{CFD}$, $T_s$ and $T_{CAA}$ were obtained by the result of five steps trials, which was performed with the number of CPUs(1-11) increased (Fig.2), where the average value of $T_s$ was 3.974 second. The $T_{ALL}$ could be calculated according to the calculation formula of (6),(7) and (8), where $n$ was set at 20,000, because of the assumption of actual simulations (Fig3,4).

5 Discussions and Conclusions

Those results shows that it is effective to take parallel efficiency of each application into consideration so as not to waste compute power resources. This present study shows that there is the effective theory for the allocation of the number of CPUs to each application and the differ-
Figure 2. Compute time of CFD and CAA depending on the number of CPUs

Figure 3. Separate execution of CFD and CAA with the number of CPUs increased

Figure 4. Sequential execution of both CFD and CAA with the number of CPUs increased
ent execution sequence of several applications. However, the choice of parallel method is different according to the characteristic of the application. The different kind of the execution pattern on several applications still has problems, which can be summarized for two things. One of the problems is time spent by data transfer between a font node and slave node. The other problem is the time spent by I/O access, such as storages. Especially, taking the data transfer on CFD applications into consideration, much lower network latency must be needed, because of the high frequency message passing among nodes by an increased number of CPUs. It seems to be a good answer to make the nodes and clusters with low latency network connected a group on CFD application. However, CAA application doesn’t need such tight connected nodes and clusters. This is why the hybrid application needs the optimal allocation of CPUs in this present study, which should be useful for hybrid applications to be performed in heterogeneous environment, such as grid environment.

On the other hand, a session manager as a status watcher of both CFD and CAA applications should be needed. GAT [1] is useful for this purpose. In order to utilize the compute and storage resources on super computing center, users should log onto the portal system. The portal system knows which users have the authority to deal with applications, and also knows which nodes can be grouped together with the latency kept low. As a next step of this study, this portal system will be developed.

DentGrid aims at playing compute and storage power supplier for dental clinics and hospitals. Hence, DentGrid needs useful applications performed with combined each other, so that can be able to support dentist to predict the change after their operations. For example, the change of sibilant /s/ by mounting oral appliances can be predicted and detected which parts of morphological features of oral tract cause the sound change by performing the hybrid application of both CFD and CAA applications.

This study was supported by IT program of MEXT(Ministry of Education, Culture, Sports, Science and Technology) in Japan.

References


