Web Based Distributed Computing Environment for Nanotechnology

D. Chen[*], A. Raghunathan[**], J. Mashl, S. Chiu, S. Parker, N. Aluru and E. Jakobsson[***]

[*]University of Illinois at Urbana-Champaign, NCSA
4017 BI, 405 N Mathews, Urbana, IL, USA, chendairui@yahoo.com
[**]University of Illinois at Urbana-Champaign, Urbana, IL, USA, araghuna@uiuc.edu
[***]University of Illinois at Urbana-Champaign, Urbana, IL, USA, jake@ncsa.uiuc.edu

ABSTRACT

One promising path of computational nanoscience is the integration of device design tools and computational chemistry tools to design molecular device components. In order to improve the efficiency of the design process and also let the device designer who is not necessarily a computational scientist use these device design tools easily, we have created a web based distributed computing system to connect client layer, web server layer and parallel computing server layer to link the jobs submission, computation, result retrieval and visualization together, and run the jobs automatically. This web based distributed computing enterprise, NanoGromacs, employs the open source molecular dynamics package Gromacs for simulation, and the open source visualization program PyMOL for visualization. NanoGromacs is freely accessible to general computational nanoscience community.

***Keywords*:** distributed computing, computational, nanotechnology, nanofluidics

1 INTRODUCTION

The field of computational nanoscience and engineering has grown rapidly in the last decade. One promising direction is to integrate device design tools and computational chemistry tools to design molecular device components [1]. Up to now the work in this aspect has been done with text based command line input specific to the UNIX/LINUX environment. Unfortunately, a lot of parameters are generally required for designing a computation task, which must be edited manually and repeatedly if a command line input interface is used. This not only decreases the efficiency of the simulation research, but also limits the tool's accessibility to scientists who are not well familiar with command line interface. In order to overcome the shortcomings of the previous design process and also let the device designer who is not necessarily a computational scientist use these device design tools easily, we have created a web based distributed computing system NanoGromacs to connect client, web server and parallel computing server to link the jobs together which can be executed automatically. NanoGromacs provides client machine the ability to communicate with the server interactively during job running, which is especially convenient for those jobs in which parameters must be provided based on intermediate results. The web server layer of NanoGromacs is a machine in our laboratory called Peptide and the parallel computing server layer is the NCSA Xeon cluster Tungsten which is composed of 1280 computer nodes.

NanoGromacs is a distributed computing environment which allows users to run the Gromacs simulation suite on remote supercomputer. Users can set up system for simulation, upload any necessary files and set run-time parameters using forms on the client computer. Next the uploaded files and forms will be preprocessed on the server computer Peptide. Then the job is submitted to supercomputer Tungsten. When the job is done, the output files may be downloaded to the local machine for visualization.

Nanofluidics is another application of NanoGromacs, which is used for simulation of the bulk-channel model of nanoscale device. It has two more features during the job running. One feature is that a UNIX console is embedded in the static HTML page for the user to interactively communicate with the server machine in case some parameters have to be entered based on intermediate results. The other feature is that there is a dynamic HTML page on the interface automatically showing the current running status of the job.

2 NANOGROMACS DEVELOPMENT

As a web based distributed computing environment, NanoGromacs is a three tier system, client - web server - parallel computing server (see Fig. 1). In the particular implementation, first, on the client side, the NanoGromacs interface helps the user to design the simulation system which is composed of two steps, system topology and molecular topology. In the system topology, the user provides the user-defined molecules and/or molecules from Gromacs library. In the molecular topology, the user uploads his own molecule files (*.itp file) and/or chooses molecule files from Gromacs library.

Second the selected molecules and the selected molecule files will be reordered to compose a topology file (*.top file) for the simulation. The users have the option to check and

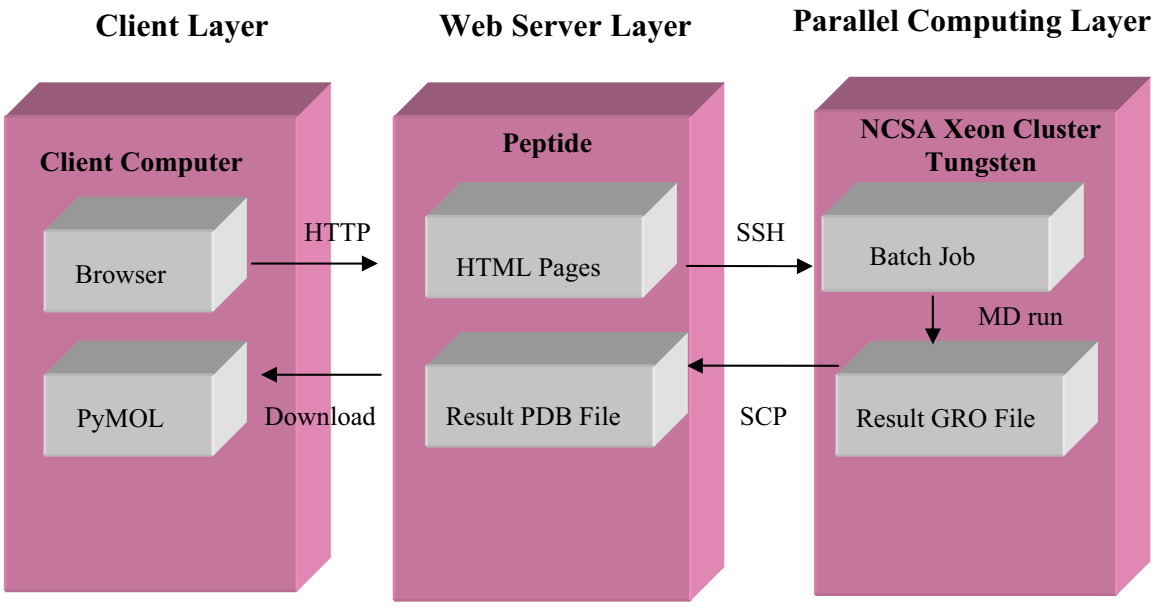

Figure 1: Structure for three tier distributed computing enterprise.

modify the amount and type of the molecules and type of the molecule files before simulation.

Then the user may choose the condition for the simulation by entering the parameters which are recorded in a special parameter file (*.mdp file).

The topology file and the parameter file will be preprocessed on the web server side Peptide to produce a run input file (*.tpr file) which serves as input for the *mdrun* program on supercomputing server side Tungsten.

Standard forms are used to process the above jobs on Peptide and the run input file will be automatically submitted to Tungsten. After the batch job is completed on Tungsten, the result coordinate file (*.gro file) is automatically converted to a PDB file and retrieved back to Peptide and/or the client computer as desired. PyMOL is automatically activated for visualization of the simulation result. Here is an example shown a simulation system (see Fig. 2).

3 NANOFLUIDICS DEVELOPMENT

Nanofluidics is a branch of NanoGromacs, which is used for designing nanoscale device. There are three steps involved, preprocessing, *mdrun*, and post processing. First, a preliminary simulation study is carried out on the simulation box. This is done using the Lennard potentials for wall atoms and corresponding changes on ions, water (Simple Point Charge/E model) in the preprocessing file. Second is the simulation of MD system using Gromacs. The shape of the channel, surface, and the number of ions are the basic parameters used in the simulation. Finally post processing involves extracting mechanical quantities such as self diffusion coefficient, density, velocity profile and flux through the channel from the MD simulation.

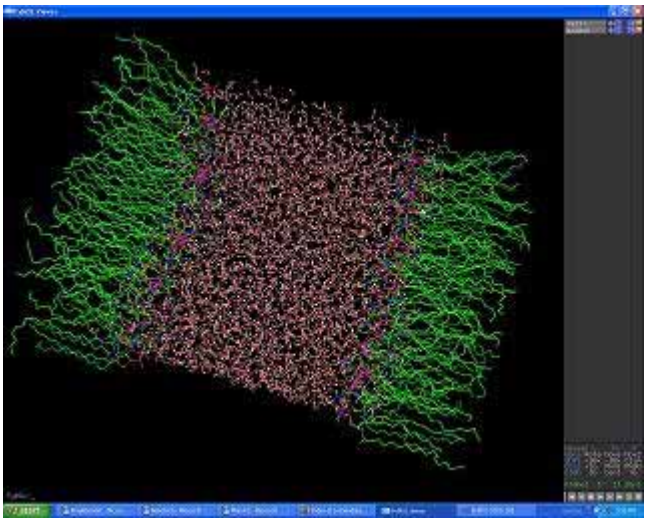

Figure 2: Rendering of an air-water interface showing explicit lipid and water molecules.

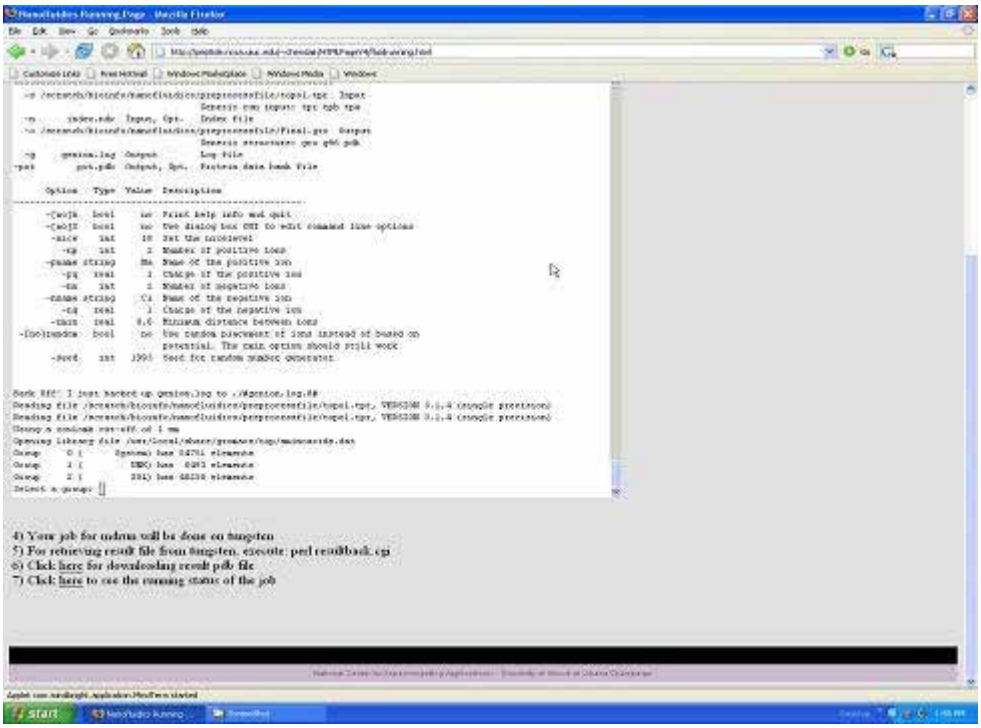

Figure 3: A UNIX console embedded in the HTML page for the user to interactively communicate with the web server.

Figure 4: Running status of Nanofluidics.

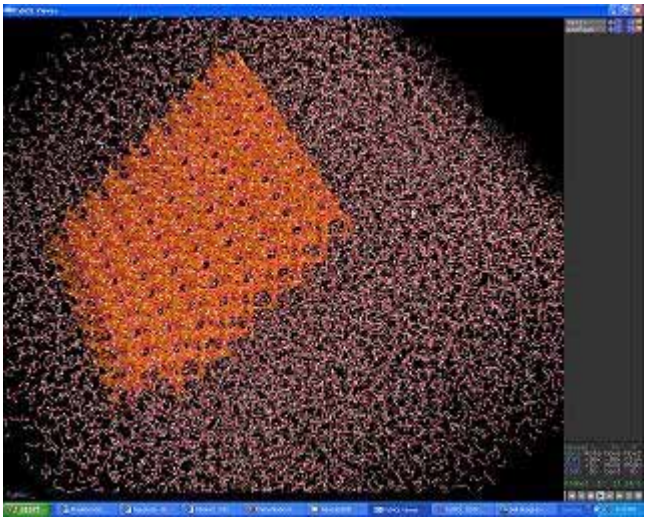

Figure 5: An example of Nanofluidics simulation.

Nanofluidics interface also provides an interactive communication method during the job running. This is especially useful if parameters must be determined based on intermediate results. In this case a java applet opens a UNIX command line window within the web interface, in the context of specific instructions for navigating the command line. An example of structure and dynamics of fluid within a nanoscale channel to illustrate the use of this feature is shown here (see Fig. 3).

The process of the implementation of the simulation includes 6 statuses, file upload, submit, queue, running, finish, retrieve, which are shown in a dynamic status page of the current running status. It is being updated every 10 seconds automatically (see Fig. 4).

After the result has been retrieved back, the structure of the system can be viewed using PyMOL. An example of Nanofluidics simulation is shown here (see Fig. 5).

4 CONCLUSION

The integration of computational chemistry tools and web based applications for nanotechnology will advance the application of computational chemistry tools in research. NanoGromacs provides a user friendly platform for communication among users, web server, and supercomputing server. It improves the efficiency for nanoscale device design for both first users and professional users.

REFERENCES

[1] N. Aluru, D. Chen, E. Jakobsson, et al., "Integration of Computational Biochemistry Tools and Device Design Tools to Design Molecular Device Components for Nanotechnology," BISTI 2003 Symposium, "Digital Biology: The Emerging Paradigm," Bethesda, Maryland, USA, November 2003.