NAMD Benchmarking on Publicly Available Philippine Computational Resources

Ronny Cheng, Ren Tristan Dela Cruz, Francoise Neil Dacanay, Gil Claudio, Ricky Nellas

Abstract—NAMD benchmarks were done on five different proteins with varying system sizes (anoplin, kalata B1, North-Atlantic ocean pout antifreeze protein, Pseudomonas aeruginosa PA01 lipase and octopamine receptor in mushroom bodies, OAMB) solvated with TIP3P water through four different publicly available computer resources in the Philippines. Our results show that the high-end desktop generated the most ns/day for small and medium-sized systems (e.g. anoplin, kalata B1, and antifreeze protein) while BlueGene/P generated the most ns/day for larger system sizes (e.g. lipase and octopamine receptor). Although these computing resources are capable of exploring protein behavior through molecular dynamics (MD) simulations for small to medium-sized systems, dealing with large systems require tremendous computational resources. This benchmark highlights the importance of intercommunication in NAMD. Moreover, our results showed the advantage of using GPU-accelerated desktops for certain MD simulations. However, the poor scalability of the high-end desktop does not make it viable for simulating large systems. Improvements in Philippine computing infrastructure and protocol is highly recommended to keep up with advances in high performance computing globally.

This work is supported in part by the Newton Agham Programme (Project Number FP160010), the Office of the Vice Chancellor for Research and Development (Grant Number PNE151512), the Natural Sciences Research Institute, and the Office of the Vice President for Academic Affairs of the University of the Philippines Diliman under the Emerging Interdisciplinary Research (EIDR) Program. Acknowledgment is also made to ASTI and the Computing and Archiving Research Environment (CoARE) of the Department of Science and Technology, Philippines, Computational Science Research Center and Philippine Genome Center - Core Facility for Bioinformatics of the University of the Philippines for the allocation of computing resources required for this study.

Ronny L. Cheng is affiliated with the Institute of Chemistry, University of the Philippines, Diliman, Quezon City, Philippines (e-mail: richcheng@up.edu.ph).

Ren Tristan A. dela Cruz is affiliated with the Department of Computer Science, University of the Philippines, Diliman, Quezon City, Philippines (e-mail: rentristandela cruz@gmail.com).

Francoise Neil D. Dacanay is affiliated with the Institute of Chemistry, University of the Philippines, Diliman, Quezon City, Philippines (e-mail: fddacanay@upd.edu.ph).

Gil C. Claudio is affiliated with the Institute of Chemistry, University of the Philippines, Diliman, Quezon City, Philippines (e-mail: gcc.upd@gmail.com).

Ricky B. Nellas is affiliated with the Institute of Chemistry, University of the Philippines, Diliman, Quezon City, Philippines (e-mail: ricky.nellas@upd.edu.ph).

Index Terms—NAMD, benchmarking, Philippine computational resources, BlueGene/P, GPU-accelerated desktops

I. INTRODUCTION

Molecular dynamics (MD) simulation is a technique that computes the classical many-body problem involving the interactions of large systems at the molecular level, i.e. biomolecules and materials. [1], [2] Its development has been crucial in the molecular level understanding of the complex motions of proteins with important biological functions. 3 Using MD, we can infer the relation between the structure and the motion of biomolecules, and its resulting structure-function relationship. [4] Understanding the motions of these biomolecules have significant importance in protein engineering and drug design. [5] MD simulations are also used in ab initio folding of proteins for predicting the 3D structures of proteins and to understand enzyme kinetics and its underlying mechanism. [6], [7], [8] It is also used to provide insight in nucleation phenomena and to design and optimize molecularly imprinted polymers. [9], [10]

Earliest implementations of MD include the simulation of hard sphere systems, [1] fluid dynamics of liquid water and argon, [11], [12], [13] calculation of thermodynamic properties of binary liquid mixtures, [14] and protein folding of crambin (46 amino acid residues), [15] among others. Most of these studies are limited only to less than a thousand atoms because of the existing computer capabilities at that time. [16] With the advent of faster computers and smarter algorithms, simulations of ten thousands up to millions of atoms are now possible with timescales ranging from picoseconds to milliseconds. [17] Simulations at these magnitudes require computational capabilities greater than a commercial desktop computer can handle. [2] To provide a solution to this problem, parallel processing is utilized to perform the simulations that would otherwise be very time consuming and impractical to execute on commercial off-the-shelf computers. [17] The simulation benefits from numerous microprocessors that are configured to perform parallel computations that significantly reduce simulation time. [17]

Another advancement to high performance computing is the use of graphic processing units (GPU).[18] Initially used as
A graphics processor, it could also be used in tandem with central processing units (CPU) to further increase computing capabilities. [19] In this set-up, GPU handles the parallel portions of a code while the CPU handles the serial portions, providing an optimized performance compared to using CPU cores alone. [19] Use of GPU-accelerated computer systems are gaining popularity due to its application as cost-effective alternatives to HPC, along with its potential to conserve power and space. [18] Desktops with Intel processors also have the capacity to use hyper-threading, wherein virtual cores are available after utilizing all physical cores. [20] Further parallelization occurs through sharing of resources between physical and virtual cores. [20]

Nanoscale Molecular Dynamics (NAMD) is an MD software able to simulate biological systems in realistic environments by taking advantage of parallel computing machines to handle the computational complexity of large molecules, by using spatial and force decomposition. [21], [22] It is available on multiple platforms, including parallel implemented clusters, desktops and laptops. [22] NAMD is highly applicable for multimillion atom systems, capable of simulating up to 2.64 million atoms and can run simulations at a femtosecond time scale. [17], [21]

The implementation of parallel computation in NAMD allows for the algorithm to be scalable, and parallel efficiency can be related to factors such as system size, number of processors and intercommunication infrastructure. [21], [23] A source of NAMD’s scalability is the overlap of calculations of non-bonded forces and communication-intensive Particle Mesh Ewald method. [24] Scalability is influenced by the amount of atoms in the system and the number of processors used. [22] Parallel efficiency decreases due to increased communication time between processors, especially for relatively small molecules wherein communication time is greater than computing time. [17] Thus, parallel efficiency is determined by the communication-to-computation ratio. [21] The maximum speed up of an algorithm is the ratio of parallel simulation time over serial simulation time. This is theoretically obtained based on the proportion of parallel components of the algorithm compared to the serial. As stipulated from the Amdahl’s law, the speedup of a program is limited by the serial components of the task. [25]

NAMD benchmarking is being done to compare simulation time of the algorithm on different processors since it is greatly affected by processor speed. [22] It is also affected by computing time and intercommunication efficiency between processors. [26] Some benchmarks for NAMD found a turnover point wherein for systems with number of atoms above this point, reduced computing time is expected for increased number of cores, enabling extrapolation of computing time. [17] Using BlueGene/P, increased system size generally results to a linear increase in calculation speed. [27] In comparison to CPU-only version of NAMD, GPU-accelerated NAMD runs 7.1 times faster and is also 2.73 times more power efficient. [18]

This study aims to benchmark NAMD using different publicly available computer resources located in the Philippines by simulating five different biochemical system sizes. Results here may be used to determine the amount of nanoseconds per day that each resource can simulate using NAMD. Also, benchmark results here may be used in the future as a guide to create a Philippine roadmap for development of biomolecular computations and related researches.

II. SYSTEMS AND METHODS

Proteins used in this study were anoplin (UniProt ID:P0C005), [28] kalata B1 protein (PDB ID:1NB1), [29] North Atlantic ocean pout antifreeze protein (PDB ID:1KDF), [30] Pseudomonas aeruginosa lipase (PDB ID:1EX9), [31] and octopamine receptor in mushroom bodies OAMB homology model (UniProt ID:Q7JQF1). [32], [33] whose cartoon representations are shown in Figs. 1a to 1e, respectively. [34], [35] For charged proteins, ions were introduced to the system to neutralize charge. The CHARMM force field was applied to every protein and all systems were solvated with TIP3P water box spanning 15 Å from the protein. [36] The number of atoms in each solvated system is shown at Table 1.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Total number of atoms</th>
<th>Number of protein atoms</th>
<th>Number of solvent atoms</th>
<th>Number of ions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anoplin</td>
<td>10 615</td>
<td>187</td>
<td>10 425</td>
<td>3</td>
</tr>
<tr>
<td>Kalata B1</td>
<td>13 828</td>
<td>379</td>
<td>13 449</td>
<td>0</td>
</tr>
<tr>
<td>N. Atlantic ocean pout antifreeze</td>
<td>21 920</td>
<td>991</td>
<td>20 929</td>
<td>0</td>
</tr>
<tr>
<td>P. aeruginosa lipase</td>
<td>52 546</td>
<td>196</td>
<td>48 345</td>
<td>5</td>
</tr>
<tr>
<td>Octopamine receptor OAMB</td>
<td>112 856</td>
<td>159</td>
<td>103 692</td>
<td>5</td>
</tr>
</tbody>
</table>

![Fig. 1: Cartoon representations of benchmarked systems: (a) Anoplin protein (UniProt ID: P0C005, 11 amino acid residues), (b) Kalata B1 protein (PDB ID: 1NB1, 29 amino acid residues), (c) North Atlantic ocean pout antifreeze protein (PDB ID: 1KDF, 70 amino acid residues), (d) Pseudomonas aeruginosa PAO1 lipase (1EX9, 285 amino acid residues), and (e) Octopamine receptor in mushroom bodies OAMB (UniProt ID: Q7JQF1, 645 amino acid residues).]
Fig. 2: NAMD benchmark results for the ASTI HPC. (a) The NAMD performance against the number of processors for the ASTI HPC. (b) NAMD speed up against the number of processors for the ASTI HPC.

Fig. 3: NAMD benchmark results for BlueGene/P. (a) The NAMD performance against the number of processors for the BlueGene/P. (b) NAMD speed up against the number of processors for the BlueGene/P.

The simulation was done at 300 K applying Langevin dynamics to regulate the simulation temperature and pressure, while the Particle Mesh Ewald (PME) method was used to calculate long range interactions. [37], [38], [39] The Shake algorithm was also used to constrain water bond geometries. [40] The cutoff for Van der Waals and electrostatic interactions were set at 10 Å. A smooth switching function for both electrostatic and van der Waals interactions is applied for 8 Å interatomic distances.

NAMD simulations were done using four different computer systems: (1) high performance computing cluster (HPC) located at the Advanced Science and Technology Institute (ASTI) (48 × Intel Xeon CPU E5-2697 v2 @ 2.70 GHz), (2) BlueGene/P located at the Philippine Genome Center (PGC) (1 rack, 1024 × 4-core IBM PowerPC 450 @ 850 MHz), (3) HPC located at the Computational Science Research Center (CSRC) (2 × 4-core Intel Xeon CPU E5405 @ 2.0 GHz) and (4) high-end desktop computer located at the good ViBES laboratory, Research Building, Institute of Chemistry, University of the Philippines Diliman (4 × Intel Core i7-4790 with 4 virtual cores @ 3.60 GHz accelerated with NVIDIA Geforce Jetstream GTX970). Version 2.12 of NAMD was ran for ASTI and CSRC HPCs while NAMD version 2.7b1 was utilized for the BlueGene/P platform. For the high-end desktop, CUDA-build NAMD was used.

III. RESULTS AND DISCUSSION

A. Computing resource performance at different system sizes

1) ASTI HPC: ASTI HPC is capable of simulating between ~4.58 ns/day for OAMB with 44 processors and ~32.98 ns/day for anoplin with 24 processors, with a linear decrease in ns/day for increasing system size (Fig. 2a). A positive correlation between relative speed up ratio and processors used was observed for 1EX9 and OAMB, while breakdown was
observed after 24-32 processors for other systems, indicating scaling breakdown (Fig. 2b). The scaling breakdown can be attributed to the high latency of Gigabit Ethernet network compared to other intercommunication networks, which have relatively high CPU overload when sending data, leading to longer idle time for processors. Thus, good scalability in ASTI HPC is only achieved at larger system sizes.

An advantage of ASTI HPC is its accessibility to Philippine researchers. However, a repercussion of this is that multiple jobs requiring HPC capabilities (not limited to NAMD) are shared in the same nodes, which causes interference greatly affecting parallel performance. Thus, there is a limit in the number of cores and processed jobs allotted to each user. In addition, a seven day job time limit is imposed and jobs surpassing this are automatically terminated.

2) BlueGene/P: The BlueGene/P is capable of generating between ~ 6.05 ns/day for OAMB using 256 processors and ~ 30.03 ns/day for anoplina using 192 processors (Fig. 3a). A negative correlation was observed between system size and generated ns/day. As shown in Fig. 3b, increase in processors used led to an increase in relative speed up ratio, indicating that the algorithm is scalable using the BlueGene/P. The slopes indicate that parallel efficiency has not been maximized yet. The lack of plateau also indicates that further increase of processors can be done without compromise of speed up. This information is vital, especially for BlueGene/P, since only 1 rack composed of 256 cores is available at the Philippine Genome Center, whereas a maximum of 72 can be utilized.

The scalability in BlueGene/P can be attributed to its intercommunication infrastructure, which utilizes five different networks to facilitate intercommunication. Of the five networks, 3-D torus is the main network used for message passing wherein each node is connected to six nearby nodes. In effect, the average path length between nodes and required bandwidth decreases, thus reducing latency.
3) CSRC HPC: The CSRC HPC on the other hand, is capable of simulating between ~0.95 ns/day for OAMB and ~11.19 ns/day for anoplin, both using 8 processors (Fig. 4a). A decrease in ns/day generated is observed for increasing system size. As a consequence, relative speed up ratio also increases, indicating good scalability (Fig. 4b). While CSRC HPC also uses Gigabit Ethernet, an explanation to its scalability is due to the lower computational capabilities of its processors (2 x 4-core Intel Xeon CPU E5405 @ 2.0 GHz), wherein computing within processors becomes the bottleneck.

The insignificant difference in slope between OAMB and 1EX9 indicate that the parallel efficiency has already been maximized, which means that if larger system sizes are simulated, a lower relative speed up ratio is to be expected.

4) High-end desktop computers: Benchmarks using the high-end desktop computer shows a subsequent decrease in generated ns/day for increasing system size. It is able to simulate between ~4.05 ns/day for OAMB and ~44.64 ns/day for anoplin, both using four processors (Fig. 5a). An increase in relative speed up ratio was observed for increasing processors, followed by a decrease after four processors (Fig. 5b). This indicates that hyper-threading for high-end desktop computer is not beneficial to NAMD, and reduced performance may be attributed to inferior specifications such as synchronous dynamic random-access memory (SDRAM) interface. The high-end desktop cluster utilizes DDR3 SDRAM (1666/1333/1066 MHz) while its next generation, DDR4 SDRAM is widely available capable of having higher transfer rates. [44] This indicates poor scalability on NAMD due to dependency on intercommunication.

Since the GPUs utilize a single instruction multiple data (SIMD) organization, lower throughput compared to using CPUs alone is expected due to minimal overlap. [45] This is a result of prioritizing high aggregate performance over optimized performance of threads. [45] In CUDA-accelerated NAMD, the GPU handles short-range nonbonded forces while the CPU keeps the atom coordinates and calculates long-range electrostatic and bonded forces. [24]

Communication comes from data transfer of coordinates and calculated nonbonded forces. [24] Since intercommunication is processed before starting the processing of calculations per node in every timestep, idle time is aggregated causing the limiting factor for high-end desktop computers to be the intercommunication. [18], [46]

B. Which computing resource to use?

1) Small system size: Comparison for small system sizes (anoplin in water system), is shown in Fig. 6. The high-end desktop computer provided the most ns/day, followed by ASTI HPC, BlueGene/P and CSRC HPC (Fig. 6a). For small system sizes, wherein intercommunication dependence is less due to minimal load balancing, the performance of the high-end desktop computer is probably due to the clock rate and bandwidth of the processors, which is the highest among computing systems used. This indicates that high-end desktop computer has the best computational capabilities among benchmarked resources.

The relative speed up ratio for small system sizes suggest that increasing the number of processors for CSRC HPC and BlueGene/P will further increase the ns/day generated for both resources (Fig. 6b). Sublinear speedup is observed for all computing resources, which is in agreement with how NAMD works in parallel systems due to the dependence in intercommunication. The relative speed up ratio value suggests that the parallel capabilities of NAMD have not yet been maximized.

2) Medium system size: Simulations for medium sized systems (1KDF in water system) show that the high-end desktop computer and BlueGene/P generated nearly identical values in terms of ns/day, followed by ASTI HPC and CSRC HPC (Fig. 7a). The benchmark shows that intercommunication becomes more important for medium system sizes.

Based on the relative speed up ratios for medium system sizes, CSRC HPC and BlueGene/P will benefit from addition of processors wherein more ns/day can be generated (Fig. 7b).
Sublinear speedup is also observed for all computing resources. The parallelization of NAMD is not yet maximized for medium sizes, as exhibited by the relative speed up ratios.

3) Large system size: For large system sizes (OAMB in water system), the BlueGene/P was able to generate the most ns/day, followed by the high-end desktop computer, ASTI HPC, CSRC HPC (Fig. 8a).

For increasing system size, it can be seen that BlueGene/P becomes a more suitable computing resource for NAMD. The results show that processor speed is not the only factor in simulation time because in terms of processor speed, BlueGene/P processors have lower processing capability compared to the high-end desktop computer processors due to lower clock rate, which is done to keep BlueGene/P power efficient. [47] Intercommunication becomes a very important factor in NAMD performance for larger system sizes.

Relative speed up ratios for all computing resources used for large system sizes show that further speed up is done by increasing the number of processors for ASTI HPC, CSRC HPC and BlueGene/P (Fig. 8b). Sublinear speedup is observed and has the highest ratio values for BlueGene/P and ASTI HPC, hence, parallel performance of NAMD is not yet maximized.

4) Synthesis: Benchmarks show that GPU-accelerated desktop cluster systems provide an alternative to other publicly available computing systems in the Philippines to solve problems using NAMD, especially for system sizes lower than 20 000 atoms. However, for larger system sizes, BlueGene/P provide the most ns/day. This signifies the importance of intercommunication in NAMD simulations, especially considering that the BlueGene/P processors have the lowest clock rate among all computing resources used, in response to being comparatively power efficient compared to other supercomputers. [47] Despite the high-end desktop computer having superior computational capabilities, bottleneck due to
intercommunication diminishes its computational advantage especially at higher system sizes where intercommunication is vital.

Our computing resources are capable of performing simulations for applications such as considering solvent and membrane environment in protein dynamics for small compounds. [2] However, simulating certain problems like protein function elucidation and biological processes still require better HPC specifications that is capable of simulating up to the microsecond timescale.[48]

To be able to increase output of local computational research, enhancement of local computing resources should be done. Better computational resources will allow us to simulate other biomolecules of interest with much larger sizes such as ribosomes, viruses, cellulose, among others.[49] Despite GPU-accelerated desktops being more cost effective, its poor scalability shows that adding cores do not provide the solution to simulate larger system sizes. Instead, systems with better scalability such as BlueGene/P should be invested upon, wherein the computational resources and time required to complete a simulation of NAMD given a certain system size can be estimated. Addition of BlueGene/P racks can be done, as previous NAMD benchmarks using BlueGene/P indicated that a further decrease in simulation time is expected when 8192 cores are used instead of 4096. [27] The successor of BlueGene/P, the BlueGene/Q system, may also be used. Benchmarks show improvements of BlueGene/Q based on power efficiency and computational capability compared to BlueGene/P. [50]

IV. CONCLUDING REMARKS

In this study, NAMD benchmarks on four different publicly available Philippine computing resources were done for different system sizes. For all computer systems, increased system size resulted in a decrease in ns/day generated. Nearly similar times were found for ASTI HPC and BlueGene/P. However, accessibility issues hamper performance for both computing resources. CSRC HPC, while no accessibility issues were found, had the slowest simulation time among all computing systems due to inferior specifications. Although the high-end desktop computer had the best hardware specifications, high intercommunication cost affects its NAMD performance, with optimal performance using four cores. Hyper-threading did not optimize NAMD simulation for the high-end desktop computer but instead contributed to the bottleneck in intercommunication.

For small and medium system sizes, the high-end desktop computer generated the most ns/day. However, for large system sizes such as OAMB, more ns/day was simulated using BlueGene/P due to its optimized intercommunication network. The benchmark shows the significance of intercommunication in NAMD simulation, especially for larger system sizes. The benchmarks also show the feasibility of using GPU-accelerated desktops as alternative to non-GPU-accelerated HPCs in NAMD for certain system sizes. However, considering that global computing resources can generate hundreds of nanoseconds per day, the performance of publicly available Philippine computing resources lag in comparison to international counterparts. The protocols regarding HPC use also needs to be varied to accommodate simulations that may take several weeks or months to process.

The Philippines is capable of producing computational research despite challenges in resources. However, to keep up with international advances in computational science and applications, the Philippines should invest in latest scalable HPC resources such as BlueGene/Q. Investment to better computing resources provides not only increased computational capabilities to solve larger problems but also improved safety and power costs, which may not only be applicable in computational chemistry but in other applications such as particle physics and weather forecasting, among others.

ACKNOWLEDGMENT

This work is supported in part by the Newton Agham Programme (Project Number FP160010), the Office of the Vice Chancellor for Research and Development (Grant Number PNE151512), the Natural Sciences Research Institute, and the Office of the Vice President for Academic Affairs of the University of the Philippines Diliman under the Emerging Interdisciplinary Research (EIDR) Program. Acknowledgment is also made to ASTI and the Computing and Archiving Research Environment (CoARE) of the Department of Science and Technology, Philippines, Computational Science Research Center and Philippine Genome Center - Core Facility for Bioinformatics of the University of the Philippines for the allocation of computing resources required for this study.

REFERENCES

[8] I. Feierberg and J. Äqvist, “Computational modeling of enzymatic keto-


