

Article

Molecular Dynamics Simulations of Clathrate Hydrates on Specialised Hardware Platforms

Nicola Varini ^{1,†}, Niall J. English ^{2,*} and Christian R. Trott ³

¹ Irish Centre for High-End Computing, Trinity Enterprise Tower, Pearse St., Dublin 2, Ireland; E-Mail: nicola@ivec.org

² The SEC Strategic Research Cluster and the Centre for Synthesis and Chemical Biology, School of Chemical and Bioprocess Engineering, University College Dublin, Belfield, Dublin 4, Ireland

³ Scalable Algorithms, Sandia National Laboratories, Albuquerque, NM 87185-1322, USA; E-Mail: ceearm@gmx.de

[†] Present Address: Research and Development, Curtin University, GPO Box U 1987, Perth, WA 6845, Australia and iVEC, 26 Dick Perry Ave, Kensington WA 6151, Australia.

* Author to whom correspondence should be addressed; E-Mail: niall.english@ucd.ie; Tel.: +353-1-7161646; Fax: +353-1-7161177.

Received: 29 June 2012; in revised form: 28 August 2012 / Accepted: 29 August 2012 /

Published: 12 September 2012

Abstract: Classical equilibrium molecular dynamics (MD) simulations have been performed to investigate the computational performance of the Simple Point Charge (SPC) and TIP4P water models applied to simulation of methane hydrates, and also of liquid water, on a variety of specialised hardware platforms, in addition to estimation of various equilibrium properties of clathrate hydrates. The FPGA-based accelerator MD-GRAPE 3 was used to accelerate substantially the computation of non-bonded forces, while GPU-based platforms were also used in conjunction with CUDA-enabled versions of the LAMMPS MD software packages to reduce computational time dramatically. The dependence of molecular system size and scaling with number of processors was also investigated. Considering performance relative to power consumption, it is seen that GPU-based computing is quite attractive.

Keywords: hydrate; GPU; FPGA; molecular dynamics

1. Introduction

Clathrate hydrates are non-stoichiometric crystalline inclusion compounds in which a water host lattice encages small guest atoms or molecules in cavities; the empty lattice is thermodynamically unstable, and its existence is due to hydrogen bond stabilization resulting from the enclathration of the trapped solutes in its cages [1,2]. There are three known common hydrate structures: (s)I, II and H. In type I hydrate, the unit cell is formed from two small 5^{12} pentagonal dodecahedral cavities and six slightly larger tetrakaidecahedral $5^{12}6^2$ cages, with 46 water molecules [1,2]. Methane hydrates are the most widespread type of clathrate, and are thought to exist in nature primarily as type I in the permafrost and deep ocean regions; this has driven significant interest in their use as a resource, as a potential geohazard, their role in the global carbon cycle, and potential influence on climate change [1–4]. Progress in research in these areas depends fundamentally on the availability of high-quality property data and an increased understanding of underlying physics and chemistry governing clathrate hydrates in nature. Molecular dynamics (MD) simulation offers a detailed understanding of underlying molecular mechanisms of hydrate behaviour, and the use of specialised hardware platforms offers obvious potential for the longer timescales required for modelling, e.g., kinetic properties. The kinetic mechanisms of clathrate hydrate crystallisation and dissociation, especially at the molecular level, are understood rather poorly. Theoretical predictions of these rates can differ from experimentally measured rates by at least an order of magnitude [1–4]; improving the understanding of the underlying mechanisms of hydrate kinetics is desirable in terms of enhancing the viability of large-scale methane production from hydrates [3,4]. Equally pertinently, the use of specialised hardware for MD modelling of dynamical properties of clathrates has the implication that once relatively time-consuming calculations can now be done very rapidly as a matter of routine. One of the goals of this article is to highlight the accelerations possible for calculations on moderate- to medium-scale systems of hydrates, and also water, and demonstrate the utility of this for determination of hydrate properties in a routine and power-efficient manner.

In parallel to developments to our understanding of hydrates, the application of specialised hardware platforms to the acceleration of molecular dynamics (MD) has seen a marked increase in activity in the past decade or so, together with the mapping of computational algorithms thereon, e.g., for non-bonded interactions, such as electrostatics. The use of Field Programmable Gate Array (FPGA) architectures has been led by the various MD-GRAPE versions [5–7], which allows for the computationally demanding calculation of non-bonded interactions [8] on the FPGA while the “mother” CPU-based (parallel, e.g., MPI/Open MP) process(es) handle the calculation of any bonded interactions, holonomic constraints and propagation of positions, velocities, *etc.* More recently, what may be described as a major “paradigm shift” in the speed of MD towards the routine sampling of millisecond timescales in a matter of only weeks, has taken place with the development of the Application Specific Integrated Circuit (ASIC)-based ANTON platform of D.E. Shaw, targeted specifically for the acceleration of MD in conjunction with the Desmond MD software suite [9,10], encompassing many state-of-the-art algorithmic advances in parallelisation and electrostatics (e.g., neutral-territory [11] and Gaussian Split Ewald [12], respectively). However, recent parallel developments in more commodity-based Graphical Processing Unit (GPU)-type HPC platforms, originating from the computer games industry, offer the prospect of relatively substantial accelerations

in MD (albeit more modest than those offered by more dedicated and specific FPGA or ASIC architectures), but that are nonetheless attractive vis-à-vis conventional CPU-based HPC architectures, especially as a hybrid CPU/GPU ensemble, in view of lower capital cost and power consumption. In such cases, CUDA is used for the mapping of non-bonded interactions onto GPU kernels, with the mother CPU-based processes handling the less computationally demanding aspects [13].

A goal of this study was to compare and contrast the performances of popular rigid water models, namely Simple Point Charge (SPC) and TIP4P [14], on the MD GRAPE 3 and a range of GPU platforms, applied to clathrate hydrates and liquid water. Another objective was to compare the performance of the potentials relative to experimental data for selected physical properties, and to demonstrate that estimation of clathrate equilibrium properties is now largely routine on specialised hardware platforms, in terms of time requirements and power consumption.

2. Simulation Methodology

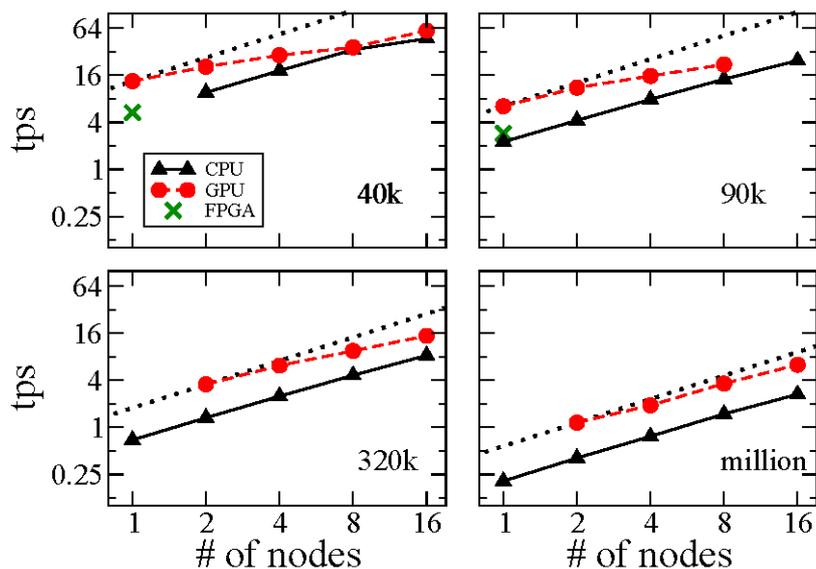
Classical MD was carried out on MD-GRAPE 3 and the M2090 NVidia Tesla double-precision GPU platforms, with two connected to each “mother” dual-CPU, quad-core nodes. The cut-off radius for Lennard-Jones (LJ) interactions was set at 8.5 Å, and the time step used was 1 fs. The Verlet method was used for propagating positions. The Particle Mesh Ewald (PME) technique [15] was used to treat long-range electrostatic interactions on MD-GRAPE-3, whilst Particle-Particle Particle-Mesh Ewald (PPPM) was used on pure-CPU and CPU-GPU [16]. A relative error tolerance of 10^{-3} was used for PME electrostatics in conjunction with MD-GRAPE-3 (with the reciprocal-space part of the computation implemented on the CPU cores), while this was 10^{-4} for PPPM. For benchmarks on liquid water, simulations were carried out in the NVT ensemble [17] at 298 K initially, with the thermostat period set to 0.2 ps, to allow for relaxation, prior to NPT simulation at 298 K and 1 bar [16], with thermostat and barostat periods of 0.2 and 0.5 ps, respectively. As a comparative test of scaling on GPU deployment *versus* a CPU implementation, four different system sizes of liquid SPC water were used for performance benchmarks, comprising 39,788, 91,430, 318,304 and 1,074,276 molecules (labelled herein as “40 k”, “90 k”, “320 k” and “million”). These were simulated for 50 ps in the NVT ensemble at 298 K, followed by around 100 ps in the NPT ensemble at 298 K and 1 bar in order to assess systematically speed and scaling with differing numbers of GPUs, CPUs and system size. The approximate dimensions of these cubic systems were, respectively, 106, 139, 211 and 322 Å.

The GPU-enabled MD code LAMMPS was used for pure CPU runs. For GPU-based simulations its so-called USER-CUDA package was employed [18]. In both cases the long-range part of the Coulombic interactions was treated with the PPPM algorithm, using its OpenMP variant from the USER-OMP package for GPU deployment [19]. The latter allows for an overlap of calculating pairwise forces on the GPU with solving long-range electrostatics on the CPUs. Since only one MPI rank per GPU is used with the USER-CUDA package, the OpenMP variant of PPPM must be utilised in order to employ all CPU cores. The simulations on the FPGA accelerators were run with the code MD-GRAPE 3. In contrast to LAMMPS, the now-typical domain decomposition strategy for parallelisation was not used, but rather the Brode-Ahrlrichs approach [20]. Where possible, estimates were made of power consumption. In addition to scaling and performance testing, the configurational energy and pair distribution functions were computed [8].

Following performance evaluation on specialised platforms, selected properties are presented for TIP4P and SPC-modelled clathrates. Methane was represented by a five-site rigid model, comprising a Lennard-Jones (LJ) 12–6 interaction site on the carbon atom and partial charges on the carbon and hydrogen atoms [21]. The C-H bond length and H-C-H bond angles were constrained at 1.09 Å and 109.47°, respectively, for this potential. In all cases, the Lorentz-Berthelot mixing rules [8] were applied for LJ interactions between different types of LJ sites. The starting configuration was based on the x-ray diffraction analysis of structure I ethylene oxide hydrate by McMullan and Jeffrey [22], which provides the positions of the oxygen atoms and the centres of mass of the methane molecules. The initial orientations of the water molecules were selected in a random manner to conform to the Bernal-Fowler rules [23] and to have a vanishingly small total dipole moment. The cubic simulation box length (for the $3 \times 3 \times 3$ unit cells) was 36.09 Å, consistent with the unit cell length of 12.03 Å at -25 °C [22]. These were simulated for 50 ps in the NVT ensemble at 250 K, followed by around 100 ps in the NPT ensemble at 250 K and 50 bar.

3. Results and Discussion

The simulation speeds in timesteps per second (tps) are plotted in Figure 1 for both CPU and GPU modes of deployment as well as for the FPGA runs on one node only. The CPU and CPU-GPU times per execution of a time step are specified in Table 1. It was found that the evaluation of the short-ranged pairwise interactions is almost completely overlapped by the calculation of long-range Coulombic interactions via PPPM for the GPU arrangement. Although the reciprocal-space calculations using the OpenMP version of PPPM are mostly slower than using the pure MPI version, this effect is smaller on a larger number of nodes. The reason is the decreased communication involved with a lower MPI rank count for OpenMP-based PPPM. When using 16 nodes, the OpenMP variant of PPPM is faster than the pure MPI variant for the 40 k- and million-molecule simulation. The implementation of holonomic constraints was also faster on a CPU-GPU basis. In comparison, one Hapertown-Stoakley node employing MD-GRAPE 3 for non-bonded force evaluation via Brode-Alrichs achieved 5.3 and 2.9 tps for the “40 k”- and “90 k”-molecule systems, respectively. The fact that this is less than half the speed of the single-node CPU-GPU arrangement is due to the less efficient CPU-implementation of reciprocal-space PME in conjunction with pairwise-force evaluation on MD-GRAPE 3 as compared to the OpenMP variant of PPPM in LAMMPS. As an aside, using a simple Coulombic cut-off leads to faster performance by MD-GRAPE 3 on a single-node arrangement than a single-node CPU-GPU implementation (by a factor of approximately 1.5). From inspection of Figure 1, it is readily evident that at larger numbers of “hybrid” CPU-GPU nodes, there is a fall-off in parallel efficiency (as can be seen by the corresponding 100%-ideal scaling lines), particularly with small system sizes. This is hardly surprising, given that the smaller molecular systems, especially “40 k”, simply do not provide enough work to keep a large number of GPU nodes busy. As has shown in previous reports, e.g., reference [18], a certain system size per GPU node is needed to provide the necessary workload for a GPU to enable it to hide memory and register latencies. If this is not ensured, performance drops substantially. Communication between GPUs on the other hand was no major bottleneck, consuming a maximum of 21% of the total runtime in the worst-case scenario of running the 40 k system on 16 nodes.

Figure 1. Time steps per second (t.p.s.) CPU, GPU and FPGA arrangements.**Table 1.** Clock times per time step (s) for various CPU and CPU-GPU runs.

No. of Molecules	No. of Nodes	CPU	CPU-GPU
39,788	1		0.0744
39,788	2	0.104	0.0485
39,788	4	0.0545	0.0346
39,788	8	0.0294	0.0273
39,788	16	0.0212	0.0167
91,430	1	0.443	0.156
91,430	2	0.235	0.0898
91,430	4	0.127	0.0637
91,430	8	0.0699	0.0453
91,430	16	0.0405	
318,304	1	1.449	
318,304	2	0.755	0.283
318,304	4	0.400	0.162
318,304	8	0.217	0.105
318,304	16	0.120	0.0673
1,074,276	1	4.863	
1,074,276	2	2.471	0.869
1,074,276	4	1.305	0.524
1,074,276	8	0.675	0.275
1,074,276	16	0.379	0.158

The oxygen-oxygen pair distribution function was computed for both SPC and TIP4P liquid water on the hybrid CPU-GPU arrangement, and agreed for either CPU or GPU evaluation. This was in excellent agreement with previous results [24], and in fair accord with experimental data [24]. The configurational energy of around -9.9 kcal/mol was achieved for both models at 298 K, also in good accord with previous results [24].

It was found that the configurational energies of the SPC- and TIP4P-modelled hydrates were -10.31 ± 0.03 and -10.42 ± 0.03 kcal/mol, in reasonable accord with previous simulations with the same potentials, although at different temperature/pressure state points and treated with Lekner electrostatics [25]. The densities were found to be 931 ± 2 and 943 ± 2 kg/m³, respectively, again consistent with previous simulation results [25].

The cage radii were determined. The radius $r_i(t)$ for cage i was defined as the averaged distance of each constituent water molecules' centre-of-mass from the cavity's hypothetical "centre-point" $r_c(t)$, with the central point computed from the sum of the cage's water molecules' centres-of-mass, *i.e.*:

$$\mathbf{r}_{c,i}(t) = \sum_{j=1}^{N_{T,i}} \mathbf{r}_j(t) \quad (1)$$

where $N_{T,i}$ denotes the number of constituent water molecules in the cage i of type T (*i.e.*, 20 for small cavities and 28 for large cages). The cages were found to be stable in structure throughout, and the identity of their constituent water molecules did not vary. The time-averaged 5¹² cage radii were found to be 3.98 ± 0.02 and 3.95 ± 0.02 Å for SPC- and TIP4P-modelled hydrate, while the corresponding respective results for 5¹²6² cages were 4.40 ± 0.03 and 4.36 ± 0.03 Å. The cage radii are in good agreement with respective experimental data at 250 K for small and large cage radii of 3.90 and 4.33 Å [2] and 3.91 and 4.34 Å [26].

4. Conclusions

It has been found that the GPU computing may be used with some profit for a variety of differently-sized water and clathrate hydrate systems, with an attractive speed-up with respect to pure-CPU deployment, provided care is taken with minimal system size for effective parallelisation. Focussing on hybrid CPU-GPU architectures, given that the power draw of two M2090 cards is about 450 W per "mother" CPU node, whilst that of the Intel Xeon node itself is around 400 W, it can be seen that the *de facto* doubling of power consumption leads to a more than proportionate increase in speed. In this sense, GPU computing is therefore attractive vis-à-vis power consumption for acceleration of MD. However, the lower power consumption of the MD-GRAPE 3 card [1–3] tends to make this more competitive in terms of MD performance per unit of power consumed for (short-ranged) pairwise potentials; in this respect, although native double-precision GPU technology with CUDA is maturing, FPGA approaches are still very attractive for MD where short-ranged potentials are appropriate. However, the need to implement reciprocal-space Ewald computations on the CPU for MD-GRAPE 3 does limit this to some extent, as can be seen by the FPGA results here for water simulation with long-range (reciprocal-space) electrostatics computations. For short-ranged potentials (in the sense that no reciprocal-space electrostatics computations are necessary), then the acceleration afforded by FPGA and GPU architectures is especially attractive. Naturally, the ASIC approach of ANTON is still 2–3 orders of magnitude faster in terms of absolute performance than the acceleration in absolute performance described here, although the development of MD-GRAPE 4's ASIC approach will serve as a possible alternative to FPGA- and GPU-acceleration of MD. It is expected that the deployment of specialised-platform MD towards the simulation of clathrate hydrates will not only render the calculation-speed of equilibrium properties a rather trivial exercise, but will

also be expected to have a dramatic impact on simulation of hydrate kinetics, where millisecond timescales may become routinely accessible by the community within the next decade or so.

Acknowledgements

We acknowledge useful discussions with Gilles Civario, Christos Kartsaklis, Tetsu Narumi, Kholmirzo Kholmurodov, J.C. Baratault, Ritwik Kavathekar, Ron Dror, Ilian Todorov and Mike Ashworth. The authors thank Science Foundation Ireland (grant code: 10 RFP-MTR2868), the Irish Centre for High-End Computing and the PRACE preparatory access initiative for provision of, and access to, computational resources.

References

1. Makogon, Y.F. *Hydrates of Hydrocarbons*; PennWell Books: Tulsa, OK, USA, 1997.
2. Sloan, E.D.; Koh, C.A. *Clathrate Hydrates of Natural Gases*, 3rd ed.; CRC Press, Taylor & Francis: Boca Raton, FL, USA, 2007.
3. Waite, W.F.; Stern, L.A.; Kirby, S.H.; Winters, W.J.; Mason, D.H. Simultaneous determination of thermal conductivity, thermal diffusivity and specific heat in Si methane hydrate. *Geophys. J. Int.* **2007**, *169*, 767–774.
4. Boswell, R. Resource potential of methane hydrate coming into focus. *J. Petrol. Sci. Eng.* **2007**, *56*, 9–13.
5. Narumi, T.; Ohno, Y.; Okimoto, N.; Suenaga, A.; Yanai, R.; Taiji, M. A high-speed special-purpose computer for molecular dynamics simulations: MDGRAPE-3. *John Von Neumann Inst. Comput.* **2006**, *34*, 29–36.
6. Taiji, M. MDGRAPE-3 Chip: A 165-Gflops Application-Specific LSI for Molecular Dynamics Simulations. In *Proceedings of the 16th IEEE Hot Chips Symposium*, Stanford, CA, USA, 22–24 August 2004.
7. Kikugawa, G.; Apostolov, R.; Kamiya, N.; Taiji, M.; Himeno, R.; Nakamura, H.; Yonezawa, Y. Application of MDGRAPE-3, a special purpose board for molecular dynamics simulations, to periodic biomolecular systems. *J. Comput. Chem.* **2009**, *30*, 110.
8. Allen, M.P.; Tildesley, D.J. *Computer Simulation of Liquids*; Clarendon Press: Oxford, UK, 1987.
9. Shaw, D.E.; Dror, R.O.; Salmon, J.K.; Grossman, J.P.; Mackenzie, K.M.; Bank, J.A.; Young, C.; Deneroff, M.M.; Batson, B.; Bowers, K.J.; *et al.* Millisecond-Scale Molecular Dynamics Simulations on Anton. In *Proceedings the Conference on High Performance Computing Networking, Storage and Analysis*, Portland, OR, USA, 14–20 November 2009; pp. 1–11.
10. Shaw, D.E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R.O.; Eastwood, M.P.; Bank, J.A.; Jumper, J.M.; Salmon, J.K.; Shan, Y. Atomic-Level characterization of the structural dynamics of proteins. *Science* **2010**, *330*, 341–346.
11. Shaw, D.E. A fast, scalable method for the parallel evaluation of distance-limited pairwise particle interactions. *J. Comput. Chem.* **2005**, *26*, 1318.
12. Shan, Y.; Klepeis, J.L.; Eastwood, M.P.; Dror, R.O.; Shaw, D.E. Gaussian split Ewald: A fast Ewald mesh method for molecular simulation. *J. Chem. Phys.* **2005**, *122*, doi:10.1063/1.1839571.

13. Stone, J.E.; Phillips, J.C.; Freddolino, P.L.; Hardy, D.J.; Trabuco, L.G.; Schulten, K. Accelerating molecular modeling applications with graphics processors. *J. Comput. Chem.* **2007**, *28*, 2618–2640.
14. Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Impey, R.W.; Klein, M.L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935.
15. Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N-log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
16. Brown, M.W.; Kohlmeyer, A.; Plimpton, S.J.; Tharrington, A.N. Implementing molecular dynamics on hybrid high performance computers—Particle–particle particle-mesh. *Comp. Phys. Commun.* **2012**, *183*, 449–459.
17. Martyna, G.J.; Tuckerman, M.E.; Tobias, D.J.; Klein, M.L. Explicit reversible integrators for extended systems dynamics. *Mol. Phys.* **1996**, *87*, 1117–1157.
18. Trott, C.R. LAMMPSCUDA—A new GPU accelerated Molecular Dynamics Simulations Package and Its Application to Ion-Conducting Glasses. Ph.D. Thesis, University of Technology Ilmenau, Germany, 2011.
19. Kohlmayer, A. Implementation of Multi-level parallelism in LAMMPS for improved scaling on petaflops supercomputers. Presented at 2011 TMS Annual Meeting, San Diego, CA, USA, 27 February–3 March 2011.
20. Brode, S.; Ahlrichs, R. An optimized MD program for the vector computer cyber-205. *Comput. Phys. Commun.* **1986**, *42*, 51.
21. Tse, J.S.; Klein, M.L.; McDonald, I.R. Computer simulation studies of the structure. I. Clathrate hydrates of methane, tetrafluoromethane, cyclopropane, and ethylene oxide. *J. Chem. Phys.* **1984**, *81*, 6146–6153.
22. McMullan, R.K.; Jeffrey, G.A. Polyhedral clathrate hydrate. IX. Structure of ethylene oxide hydrate. *J. Chem. Phys.* **1965**, *42*, 2725–2732.
23. Bernal, J.D.; Fowler, R.H. A theory of water and ionic solution, with particular reference to hydrogen and hydroxyl ions. *J. Chem. Phys.* **1933**, *1*, 515–547.
24. English, N.J. Molecular dynamics simulations of liquid water using various long-range electrostatics techniques. *Mol. Phys.* **2006**, *103*, 243–253.
25. English, N.J.; MacElroy, J.M.D. Structural and dynamical properties of methane clathrate hydrates. *J. Comput. Chem.* **2003**, *24*, 1569–1581.
26. Sloan, E.D. Fundamental principles and applications of natural gas hydrates. *Nature* **2003**, *426*, 353.