FlashLite: A High Performance Machine for Data Intensive Science

David Abramson
Research Computing Centre, School of IT & EE
The University of Queensland
St Lucia, AUSTRALIA, 4072
david.abramson@uq.edu.au

***Abstract*— Data is predicted to transform the 21st century, fuelled by an exponential growth in the amount of data captured, generated and archived. Traditional high performance machines are optimized for numerical computing rather than IO performance or for supporting large memory applications. This paper discusses a new machine, called FlashLite, which addresses these challenges. The paper describes the motivation for the design, and discusses some driving application themes.**

Keywords—Data Intensive Science; High Performance Computing; IO Intensive Computing;

I. INTRODUCTION

Data is predicted to transform the 21st century, fuelled by an exponential growth in the amount of data captured, generated and archived [10][17]. Jim Gray identified data intensive science as the fourth scientific paradigm after computation as the third [17][19]. Not surprisingly, there are numerous projects targeting the challenges that underpin the exploitation and management of this data explosion, for example [5][22][12].

There has been significant progress towards addressing some of the opportunities and infrastructure challenges posed by managing such rapid increase in data volumes. Substantial investment, through the Research Data Aliance (RDA) in the US, and the Australian National Data Service (ANDS) in Australia, together with data infrastructure investment advance our ability to search, manage and store large amounts of data. In Australia alone, there has been investment in the Research Data Strategic Infrastructure (RDSI) and the National eResearch Collaboration Tools and Resources (NeCTAR) projects, both of which provide leading edge cloud infrastructure. However, whilst important, these investments do not address the escalating scientific imperative to exploit and *process* data.

This paper discusses a high-performance computing system, called FlashLite, that is designed explicitly for data intensive science and innovation. It supports applications that need very high performance secondary memory as well as large amounts of primary (main) memory, and optimises data movement within the machine. Data intensive applications are neither well served by traditional supercomputers nor by modern cloud-based data centres. Conventional supercomputers maximise Floating Point Operations per Second (FLOPS) and inter-processor communication rates through high bandwidth and low latency networks. Conversely, modern Cloud systems minimise the cost of ownership through reliance on virtual machines and shared storage; they thus utilize relatively slow processors and networks and, by and large, do not support parallel processing.

FlashLite, on the other hand, maximises Input Output Operations per Second (IOPS) while achieving competitive FLOPS ratings and high performance networking, producing a balanced system for applications that exploit parallelism, high-speed arithmetic, and high performance Input/Output (IO). The machine also incorporates novel software mechanisms that provide seamless access to data regardless of its location, making it easier to build new data-intensive applications, whilst supporting legacy codes using familiar techniques.

Whilst it is a new machine in its own right, FlashLite is inspired by Gordon [20][24][23][30][34], the US National Science Foundation (NSF) machine that has achieved impressive performance on data intensive applications at the San Diego Supercomputer Centre (SDSC). A wide range of applications, from computer science to computational biology and bioinformatics, have benefited from Gordon's unique design features.

II. MOTIVATION

The growth of digital data is leading to the so-called "data tsunami", driven by advances in digital detectors, networks and storage systems. Each of these technologies are following Moore's Law, which provides improvements in detector resolution (with associated lower costs), network transmission rates and memory densities that are doubling at least every 18 months. The growth is also driven by increased resolution of computational models, which leverages improved computational power through multi-core devices, again following Moore's Law. Critically, the gap between the arithmetic processing speeds and transfers between main and disk memory is increasing, making it more and more difficult to analyse and process increasingly large volumes of data in a timely way.

Traditional High Performance Computing systems (HPC) are designed to deliver maximum computational performance on scientific codes, but are not typically optimised for maximising memory throughput – either primary (main) memory or secondary (disk) memory.

Figure 1 summarises access times for different storage technologies in terms of clock cycles for a typical memory hierarchy (drawn from [20]). This hierarchy works only because programs exhibit both spatial and temporal locatility, meaning that when a data item has been accessed, the program is likely to both access the data in adjoining locations, and also will likely access that same location in the near future. Conventional programming lanuages abstract the top few layers of the hierarchy, and move data between registers, cache (often multiple layers of cache) and main memory, transparently. In parallel machines, access to memory in other processor nodes is either performed by the hardware, using shared memory operations, or explicitly moved using message passing primitives. Below this layer, spinning disk provides bulk sotrage, but at the cost of millions of processor cycles, and using explicit file IO calls. Importantly, there is a latency gap between remote memory and spinning disk of around three orders of magnitude.

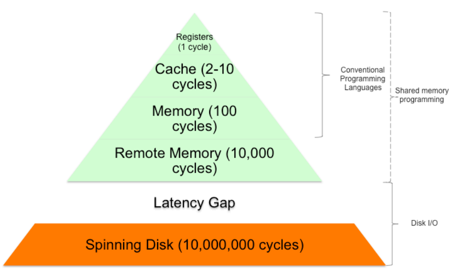

Figure 1a – Latency Gap in traditional systems [20]

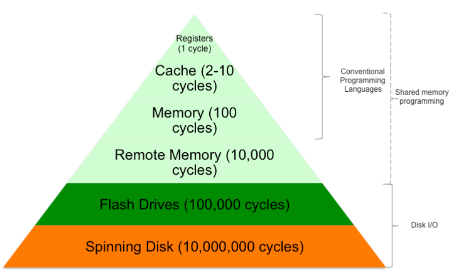

Figure 1b – The Role of Solid State Disk [20]

In order to enable HPC applications over large data volumes, it is time to rethink existing computer architectures to redress this imbalance. FlashLite is such a system, and incorporates three key innovations that are not typically embraced by traditional HPC systems:

- high throughput solid state disk (instead of spinning disk);
- large amounts of main memory; and
- software shared memory.

These innovations mean that applications have increased access to a high performance memory system, spanning high speed main memory (usually supported by DRAM technology) and low latency, high throughput, solid state disk (also known as Flash memory). A novel (commercial) software abstraction, called Virtual Shared Memory (vSMP) [14] bridges these components, thus simplifying programming. vSMP also makes it possible to aggregate nodes to provide the right amount of memory per application, rather that having to configure nodes with the maximum required by the largest application. This makes it possible to buy memory at the sweet spot in size against price rather than having to buy the most expensive memory just to achieve the maximum size.

FlashLite is not intended as a general purpose machine but is optimised for data intensive applications. In the next section we propose some typical templates for the types of applications that may benefit from this new architecture.

III. Data Intensive Themes

FlashLite provides a paradigm shift in the underlying computational techniques, in turn delivering substantial new outcomes in core scientific domains. Accordingly, we have identified four key themes that will benefit significantly from this class of machine. These are:

1) applications that directly manipulate large amounts of data;
2) applications that integrate observational data and computation;
3) applications that require large main memories to operate efficiently; and
4) applications with significant temporary storage requirements.

A. *Theme 1: Applications that Directly Manipulate Large Amounts of Data*

This theme concerns new developments in core computational techniques that address large data sets (sometimes called '***Big Data'***). FlashLite offers unique opportunities for innovation, both in the way we ***manage*** large data sets (applying database technologies to very large (in memory) databases), and the way we ***understand*** large data sets (using machine learning and classification techniques). Both of these areas are currently limited by the speed and scale of data access mechanisms. FlashLite will improve both of these significantly. Below we give some typical examples from our collaborators of such applications.

1) Large Memory Database Systems

Traditional database management systems (DBMS) focus on optimising I/O costs. Current data warehousing and data mining techniques are largely evolved from the same assumption of large-and-slow-disk and small-but-fast-memory systems. In recent years, there has been significant international R&D effort to develop new database systems that work with large SSD disks, main-memory systems and distributed and parallel systems. One of our collaborating groups at the University of Queensland is the world's first to develop such systems for dynamic spatial data [32][33].

Effective management and real-time processing of large-scale spatial data has transformative potential for existing and future location-based enterprise applications. Current spatiotemporal database systems are not ready to support location-aware real-time big data analytics, due to the lack of a computing foundation for (1) efficiently processing large-volume historical locations of a large number of objects; and (2) continuous real-time processing of high-velocity streamed location data. A traditional bottleneck is intrinsically high I/O cost for spatial data processing, and the difficulties of parallelising spatial query processing. The main innovation of this project lies in ensuring that we can push the frontier of the research in this area to the level that can support real-time analytics for very large amounts of spatial data by using modern computer architecture, exploring the power of SSD-based databases, column/row stores and multicore and cluster-based parallelism. We will contribute to the knowledge base of big moving-object data analytics through a range of innovations including: (1) new data models with compression techniques specifically designed for main-memory and SSD trajectory data representation; (2) a set of architecture-aware parallelism-by-design spatiotemporal operations; (3) an extensible suite of advanced queries that can deal with terabyte-scale compressed spatial data; and (4) a proof-of-concept prototype based on FlashLite.

2) Machine Learning and Classification

Probabilistic models, statistical learning, multi-linear factorisation models and constrained factorisation models are powerful approaches for understanding data, with applications in computer vision, Web-scale data mining and social network analysis, to name just a few. Previously, this group at the University of Technology in Sydney (UTS) has used traditional clusters to implement multi-linear factorisation models for vision-based human behaviour analysis [29][27][28]. The work yielded new results in machine learning that were reported in high tier conferences and journals. However, the scale of this work has been limited by traditionally low IO rates.

FlashLite will facilitate the execution of a large number of particles (using Markov Chain Monte Carlo techniques) over large data sets, which will speed up inference and improve confidence in the results enormously. For large-scale Web data mining, it is often required to store and analyse millions of Web pages and their link structures (such as PageRank or SimRank scores). This type of mining process is typically IO intensive, but FashLite, will achieve real-time analysis and mining which was previously impossible.

B. Applications that Integrate Observation Data and Computation

Traditionally, computational models provide predictions that are then compared with data for corroboration and validation. They are then enhanced and augmented with new mechanisms to improve their "skill". As sensor networks and instruments generate more (possibly real-time) data, it is increasingly important to integrate this data into computational models as they execute. The IO speeds of traditional supercomputers limit our ability to do this, but FlashLite's high speed IO will make it possible to integrate data as computation proceeds.

The variety of modelling techniques used in these applications varies enormously. However, the common attributes of this theme are that applications require both excellent floating point and secondary memory performance and that these are matched to optimise overall throughput. Here we describe 6 projects that fall into this class, and will benefit from FlashLite's increased IO rates.

1) Astrophysics

Cosmology has made rapid progress in the last decade with sufficient high-quality observational measurements now available to constrain a detailed cosmological model with some 12-20 physical parameters. Such models are fitted to the data sets using advanced statistical techniques, notably the Markov Chain Monte Carlo (MCMC) approach where the multi-dimensional parameter space is explored in an optimum fashion not just for efficiency, but to map out the probability distribution of the parameter values. One of the most popular packages currently used in cosmology is "CosmoMC" [11] [21].

This approach will not be possible for the next generation of cosmological models, because they will rely on high-resolution simulations instead of simple analytic functions to describe each model. The problem is that the simulations cannot be generated rapidly during the MCMC fitting, but must be generated in advance for a suitable parameter range and stored. These simulation data sets are very large as they each describe the matter distribution in a large volume of the universe over many time steps, leading to data sizes of order 10 TB.

FlashLite's (SDD) memory will enable a new paradigm for cosmological model fitting. This will be the first time that the non-linear modelling represented by large simulation data sets can be included in the cosmological models fitted with the MCMC approach, leading to significantly new cosmological predictions.

2) Healthy hearts

One of the great revolutions of the 21stcentury is the move towards personalised health care. A crucial step in this long process is the integration of patient-specific heterogeneous data sets across many different scales and modalities in order to build validated, predictive models with real power. Sudden cardiac death is one of the leading causes of mortality in the western world and one of the main factors responsible is myocardial ischaemia, the reduction of blood supply to the heart due to coronary artery occlusion, which may lead to arrhythmias and, in particular, re-entrant circuits that are initiated in the ischaemic region, such as those characteristic of ventricular tachychardia (VT) [31][4][6][1].

This project integrates a variety of data sets in order to build models of the human heart with predictive power. These data sets include those obtained via electrocardiogram (ECG), where electrodes are placed on the surface of the patient's body; optical voltage mapping based on non-contact visualisation of surface electrical activity; Magnetic resonance diffusion tensor imaging (DTI) to non-invasively assess cardiac structure by assessing the self-diffusion of water molecules; and finally histological data that provides structural information about the alignment of the myocyte fibres and sheets in cardiac tissue.

These data sets are diverse and large, including both complex temporal and spatial features, which amounts to hundreds of GBytes per simulation. We are developing new approaches to automatically extract spatial features from images but these techniques are computationally and data intensive – thus we need access to a mix of fast, large memory and substantial computing power in order to compare the data sets from patient to patient. FlashLite will provide a system that balances these for the first time, leading to dramatically improved predictive models.

3) Coastal Management

One of our partners has designed and implemented data-intensive, high-resolution numerical models for various coastal management issues, including high-resolution coupled ocean modelling and real-time storm surge forecasting [3][2]. These codes not only require high performance computers with large memory, but also require fast I/O processing and secure storage. For example, I/O processing with traditional hard disks accounts for a high-resolution coupled ocean model and consumes approximately 30% of the total execution time.

Real-time storm surge forecasting is critical for a timely emergency response, and requires the model to deliver inundation maps as quickly as possible. FlashLite's high speed IO will impact the performance of this model and as a consequence improve the research outcomes significantly, providing more accurate predictions faster than real time.

4) Climate Change

Another of our collaborators performs research into climate change adaptation and mitigation [16]. The program promotes a multidisciplinary approach to climate change research inclusive of the environmental, social, economic, and governance dimensions. The goals of this work is to provide a phase shift in the capacity of Australian researchers and their international partners to model the responses of plant and animal species to projected future climate parameters.

This group will use FlashLite to model the responses of thousands of species to hundreds of climate change scenarios, including local scales that utilize regionally down-scaled global climate change output. Modelling 19 bioclimatic variables across 18 Global Climate Models for each of 9 emission scenarios at each of 8 time series points involves processing inputs from raster coverage files, species distribution files and environmental data (3TB at 1km grid resolution, with the addition of 41TB at 250m resolution already planned), as well as species distribution input files and environmental data, requiring, in a total of 50,000 files), dynamic manipulation of terabytes of data and production of enormous volumes of output data that then needs to be down-scaled. This is currently a massively IO-intensive process, with processing times of weeks or months on conventional supercomputers. In addition, both the spatial and temporal resolution of species models is rapidly increasing, along with their complexity and use of ancillary environmental data. FlashLite's high speed IO capabilities will reduce the running time significantly and improve model "skill".

5) LIDAR processing

Another group undertakes extensive research into erosion processes in the Australian landscape, with a particularly emphasis on the land use drivers of erosion and how the eroded sediment is impacting critical assets such as the Great Barrier Reef and Moreton Bay [15]. In recent years the rapidly developing LiDAR (Light detection and ranging) technology has been a key tool employed by the group to measure changes in the land surface at high resolution and high precision. Repeat airborne LiDAR is used over large areas (100s of km2) to measure changes in the ground surface topography at a level of detail (+/- 10cm) that would have been unimaginable only a decade ago; while terrestrial LiDAR is used at more local scales (few 100m2 to several hectares) to measure changes at mm scale accuracy.

These technologies, however, generate extremely large image datasets (100s of GB per image), that require real-time processing using GIS and other image processing software. At present large amounts of time are spent segmenting these datasets to allow them to be processed on a high-end workstation. The processing of these datasets is currently a time limiting factor in our research workflow, and FlashLite will increase our efficiency dramatically by allowing much larger image segments to be processed in real time, leading to dramatically improved erosion models that can be applied in practice.

C. Applications that Require Large Main Memories

Some codes have extremely large working memory requirements, and do not execute efficiently on conventional supercomputers that often have very restricted main memories. Moreover, many of these codes do not contain large amounts of parallelism, and thus do not speed up well on traditional supercomputer clusters. FlashLite has 68 nodes containing 500 GB of main memory each, and thus it is possible to run applications that simply cannot execute on many of the existing facilities. The vSMP software included in FlashLite allows codes to access remote memory transparently, making it possible to address up to 4 TB from a single node . Gordon has already demonstrated the effectiveness of this on applications such as de novo genome assembly (using Velvet [13]). FlashLite makes improvements to Gordon's design (such as increasing the amount of main memory available) and this will allow it to run codes of this class more efficiently.

1) Genomics

Sequencing the human genome is considered to be the greatest of human achievements, and advances continue to be made in other organisms, with outcomes for improved agriculture and the environment. Genome sequencing technology continues to advance with ever increasing throughput and reducing costs. A major leap in this technology occurred over the last few years with the development of high throughput Next Generation Sequencing (NGS) methods, which differ from traditional sequencing technology in that they produce very large numbers of relatively short reads. This approach has been applied successfully for the rapid and cost effective sequencing of genomes.

The complexity and size of the majority of eukaryotic genomes makes this work computationally challenging, with genomes such as bread wheat containing 17 thousand million letters across three related genomes, and assembled using reads of only 100 letters. One of the main limitations of this work is access to very large memory computing infrastructure capable of storing “kmer” strings and “De Bruijn” graph assemblies. However, we have observed significant variability in the memory footprint, making it difficult to decide an optimal memory configuration for the nodes.

FlashLite’s large memory nodes, and novel vSMP software will underpin our continued research and allow a scale of activity not currently possible. Users can configure Super Node in increments of 512GB of main memory, up to a limit of 4TB. This allows us to perform whole genome reconstruction without locking into a particular size node.

D. *Applications with Significant Temporary Storage Requirements*

This class of codes execute on conventional HPC systems and leverage high performance floating point arithmetic units, such as those found on modern multi-core chips and, more recently, GPU accelerators. However, some of these have significant requirements for temporary (‘scratch’) storage, and thus spend a significant amount of their time reading and writing to spinning disk. By replacing this with SSD, FlashLite will improve the execute time of these codes significantly.

1) *Computational Chemistry*

Computational chemistry is often used to enhance the understanding and development of new materials for a diverse range of applications including lithium-ion batteries, fuel cells, solar hydrogen production, strengthening of metals using metal matrix composites, and new catalysts and membranes for gas separation and desalination [8][7][35][16][26][25]. These applications require a number of computer packages for calculations of electronic structure, reaction pathways and dynamics. Computational results from the groups have led to breakthroughs in the identification and understanding of properties of materials, which have been published in high profile journals. Currently we are performing high-throughput screening of advanced materials for energy storage and conversion. However, the complexity of the systems that can be studied, and the accuracy with which they can be treated, are limited by the resources currently available.

Gordon has demonstrated that Flash disk and large amounts of shared-memory can significantly enhance the performance of Gaussian 09 computations for electronic energy calculations. Furthermore, GAMESS, for electronic energy calculations, CPK2 for first principles molecular dynamics calculations have been installed on the SDSC facility, and we expect similar benefits to those demonstrated for Gaussian. FlashLite’s combination of high performance parallel computing with high speed IO will dramatically improve the speed of executing these codes, and will have significant impact on scale of the systems being studied.

IV. FLASHLITE DESIGN

FlashLite is in many respects a traditional cluster with a number of compute nodes connected by a high speed interconnect. It is optimised for data intensive computation, and 68 compute cores containg a total of:

- 1632 cores
- 34.8 TB of RAM
- 326.4 TB of SSD storage
- 65.28 Tflop/s (Rpeak).

The network is a dual rail 56Gbps Mellanox infiniband fabric, providing non-blocking within groups of 24 nodes and 2:1 blocking factor between groups of 24 nodes. The rails are dedicated so that one supports message passing traffic (MPI) whilst the other provides vSMP messaging. This means that applications can exploit both message passing and shared memory paradigms concurrently without network contention. This is important for hyrbid codes that exploit both message passing and shared memory (provided by vSMP). It also allows both message passing and file input output.

The system also has a number of software components

- ScaleMP vSMP software that aggregates multiple nodes into "super nodes" with larger memory/CPU/disk/IO than the individual nodes
- ROCKS cluster management software
- Torque + Maui batch system

FlashLite is connected via Infiniband to a high performance GPFS based file system. This provides parallel IO to all of the nodes, supprting application that have a high IO throughput to the remote file system. Note that since the Flash drives are directly connected to the PCI bus, they can be addressed directly by each node.

Compute nodes , shown below in Figure 2, consist of:

- 2 x Xeon E5-2680v3 2.5GHz 12core Haswell processors with 30MB Smart Cache
- 16 x 32GB DDR4-2133 ECC LRDIMM modules – total 512GB (256GB per socket)
- 2 x 500GB 2.5" 7.2K HDDs as RAID 1 system disk
- 3 x 1.6TB Intel P3600 2.5" NVMe (SSD) drives for local data storage
- 2 x Mellanox Connect-IB 56Gb/s FDR Single Port Infiniband PCIe3 x8 adapter.

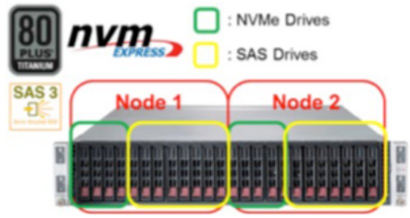

Figure 2 – Structure of a Compute Node

Compute nodes can be flexibly aggregated together into larger "supernodes" using ScaleMP's vSMP software. This software aggregates multiple physically separate servers into one single virtual high-end system. Such a "vSMP supernode" aggregates the CPUs, memory, and I/O capabilities of multiple physical hosts into one virtual machine (VM).

The upper limits to the configurations of our supernodes are:

- maximum of 4 supernodes
- maximum of 16TB aggregate RAM
- maximum of 1056 aggregate cores (88 processors).

For example, one such configuration of FlashLite might see:

- 1 x 8TB RAM supernode with 384 cores and 76.8TB of SSDs (16 physical nodes)
- 1 x 4TB RAM supernode with 192 cores and 38.4TB of SSDs (8 physical nodes)
- 2 x 2TB RAM supernodes with 96 cores and 19.2TB of SSDs (2x4 physical nodes).
- Plus the remaining physical nodes running outside of vSMP
- 36 x 0.5TB physical nodes 24 cores and 4.8TB of SSDs (36 physical nodes).

Given the network topology, supernodes of up to 24 physical compute nodes in one non-blocking group with 576 cores and 12TB RAM is optimal. It is possible, however, that some applications may be able to make effective use of a larger supernode that spans two non-blocking network groups (subject to the limits outlined above).

The SSD drives can be arranged in a number of ways. Each compute node is connected directly to 3 x 1.6 TB drives. These can be configured as a single 4.8 TB drive, or cut into three 1.6 TB partitions. Importantly, the drives can be aggregated across the nodes either using vSMP's device virtualisation, and thus each node in a supernode might see all of the drives. Alternatively, they can be mapped into the Infiniband address space and virtualised that way. At present we have not determined the optimal strategy for this, and will test a number of alternatives.

V. CONCLUSIONS AND FURTHER WORK

FlashLite was only commissioned in August 2015, and a limited number of early adopters have recently received accounts. Thus, it is far too soon to report any performance results. The appliction themes we have illucidated provide a way of categorizing which applications might benefit from FlashLite's architecture. However, it will be some time before we have hard data on this and before we can report on the lessons learned. We expect to report on these in the future.

ACKNOWLEDGMENT

FlashLite is funded by the Australian Research Council Linkage Infrastructure, Equipment and Facilities scheme (2014), The Queensland Cyber Infrastructure Foundation (QCIF), The University of Queensland, University of Technology Sydney, Griffith University, Monash University and Queensland University of Technology. This paper contains input from the chief investigators of grant LE140100061, namely Xiaofang Zhou, Debra Bernhardt, Chengqi Zhang, Xingquan Zhu, Dacheng Tao, Ling Chen, Michael Drinkwater, Rodger Tomlinson, Ross Coppel, Yuantong Gu, Kevin Burrage, Lyn Griffiths, Ian Turner, Brendan Mackey, Aijun Du, Kerrie Mengersen and David Edwards. We also acknowledge Dragan Dimitrovici and Werner Scholz from Xenon Systems who built and supplied the system.

REFERENCES

[1] Bueno-Orovio, A., Kay, D., Grau, V., Rodriguez, B. and Burrage, K. "Fractional diffusion models of cardiac electrical propagation: role of structural heterogeneity in dispersion of repolarization", Journal of the Royal Society Interface, 2014 Aug 6;11(97):20140352.

[2] Burston, J., Nose, T., and Tomlinson, R.B. (2013),"Real-time numerical simulation of storm surge inundation using high-performance computing for disaster management, Queensland, In Piantadosi, J., Anderssen, R.S. and Boland J. (eds) MODSIM2013, 20th International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, December 2013, ISBN: 978-0-9872143-3-1.

[3] Burston, J., Ware, D. and Tomlinson, R.B. (2015), The real-time needs of emergency managers for tropical cyclone storm tide forecasting: results of a participatory stakeholder engagement process, Natural Hazards, DOI 10.1007/s11069-015-1794-7

[4] Carapella, V., Bordas, R., Pathmanathan, P., Lohezic, M., Schneider, J. E., Kohl, P., Burrage, K. and Grau, V. "Quantitative study of the effect of tissue microstructure on contraction in a computational model of rat left ventricle", PLOS ONE, April 02, 2014DOI: 10.1371/journal.pone.0092792.

[5] Foster, I. and C. Kesselman, The Grid: Blueprint for a New Computing Infrastructure. 1999, USA: Morgan Kaufmann

[6] Gemmell, P., Burrage, K., Rodrigue, B. and Quinn, T. A. "Population of computational rabbit-specific ventricular action potential models for investigating sources of variability in cellular repolarization", PLoS ONE, 28/2/2014, DOI: 10.1371/journal.pone.0090112.

[7] Han, K., Bernardi, S., Wang, L. and Searles, D. "Water Diffusion in zeolite membranes: Molecular dynamics studies on effects of water loading and thermostat" Journal of Membrane Science, doi:10.1016/j.memsci.2015.08.033 accepted (August 2015).

[8] Hankel, M., Ye, D. Wang, L. and Searles, D.J., "Lithium and Sodium Storage on Graphitic Carbon Nitride" Journal of Physical Chemistry C, 2015. DOI: 10.1021/acs.jpcc.5b07572.

[9] Henriques, B., Thomas, P., Oliver, S. and Roseboom, I. "Monte Carlo Markov Chain Parameter Estimation in Semi-Analytic Models of Galaxy Formation", DOI: 10.1111/j.1365-2966.2009.14730.x

[10] Hey, T. and A. Trefethen, The Data Deluge: An e-Science Perspective, in Grid Computing: Making the Global Infrastructure a Reality, F. Berman, G. Fox, and T. Hey, Editors. 2003, Wiley: England. p. 809-824.

[11] http://cosmologist.info/cosmomc/

[12] http://iic.seas.harvard.edu/research/data-intensive-science

[13] http://www.molecularevolution.org/software/genomics/velvet

[14] http://www.scalemp.com/architecture

[15] https://www.griffith.edu.au/environment-planning-architecture/australian-rivers-institute

[16] https://www.griffith.edu.au/research/research-excellence/griffith-climate-change-response-programLiao, T., Sun, Z., Sun, C. and Searles, D. “Electronic coupling and catalytic effect on H2 evolution of MoS2/graphene nano catalyst” Scientific Reports, 4, 6256/1-7 (2014).

[17] McFredries, P. “The Coming Data Deluge”, IEEE Spectrum, Feb 2011.

[18] Microsoft Inc, “Science@Microsoft - The Fourth Paradigm in Practice”, Edited by David Heckerman, Stephen Emmott, Yan Xu and. Kenji Takeda, http://www.microsoft.com/en-us/researchconnections/science/stories/

[19] Microsoft Inc, “The Fourth Paradigm: Data-Intensive Scientific Discovery”, Edited by Tony Hey, Stewart Tansley, and Kristin Tolle, http://research.microsoft.com/en-us/collaboration/fourthparadigm/

[20] Norman, M. and Snavely, A “Accelerating data-intensive science with *Gordon* and *Dash*”, *Proceedings of the 2010 TeraGrid Conference* (TG '10). ACM, New York, NY, USA, , Article 14 , 7 pages. DOI=10.1145/1838574.1838588 http://doi.acm.org/10.1145/1838574.1838588

[21] Parkinson, D. et al, "The WiggleZ Dark Energy Survey: Final data release and cosmological results” DOI: 10.1103/PhysRevD.86.103518

[22] Shoshani, A and Rotem, D., Scientific Data Management: Challenges, Technology and Deployment, Chapman & Hall, 2009

[23] Sinkovits, R., Cicotti, P., Strande, S., Tatineni, M., Rodriguez, P., Wolter, N. and Balac, N. “Data intensive analysis on the gordon high performance data and compute system”, Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '11). ACM, New York, NY, USA, 747-748. DOI=10.1145/2020408.2020526 http://doi.acm.org/10.1145/2020408.2020526

[24] Strande, S., Cicotti, P., Sinkovits, R., Young, W., Wagner, R., Tatineni, M., Hocks, E, Snavely, A and Norman, M. “Gordon: design, performance, and experiences deploying and supporting a data intensive supercomputer”, Proceedings of the 1st Conference of the Extreme Science and Engineering Discovery Environment: Bridging from the eXtreme to the campus and beyond (XSEDE '12). ACM, New York, NY, USA, , Article 3 , 8 pages. DOI=10.1145/2335755.2335789 http://doi.acm.org/10.1145/2335755.2335789

[25] Sun, Q, Wang, M., Li, Z., Du, A and Searles, D.“A computational study of carbon dioxide adsorption on solid boron”, Physical Chemistry Chemical Physics, 16, 12695-12702 (2014).

[26] Sun, Q., Li, Z. Searles, D., Chen, Y., Lu, G. and Du, A. “Charge-controlled switchable CO2 capture on boron nitride nano materials” Journal of the American Chemical Society, 135, 8246-8253 (2013).

[27] Tao, D., Li, X., Wu, X. and Maybank, S. “General Tensor Discriminant Analysis and Gabor Features for Gait Recognition”. IEEE Trans. Pattern Anal. Mach. Intell. 29(10): 1700-1715 (2007)

[28] Tao, D., Li, X., Wu, X., Hu, W. and Maybank, S. “Supervised tensor learning”, Knowl. Inf. Syst. 13(1): 1-42 (2007)

[29] Tao, D., Li, X., Wu, X., Maybank, S. “General Tensor Discriminant Analysis and Gabor Features for Gait Recognition”. IEEE Trans. Pattern Anal. Mach. Intell. 29(10): 1700-1715 (2007)

[30] Wagner, R., Tatineni, M., Hocks, E., Yoshimoto, K., Sakai, S., Norman, M., Bockelman, B., Sfiligoi, I., Tadel, M., Letts, J. Würthwein, F. and Bauerdick, L. “Using Gordon to accelerate LHC science”, Proceedings of the Conference on Extreme Science and Engineering Discovery Environment: Gateway to Discovery (XSEDE '13). ACM, New York, NY, USA, , Article 59 , 4 pages. DOI=10.1145/2484762.2484834 http://doi.acm.org/10.1145/2484762.2484834

[31] Walmsley, J., Mirams G. R, Pitt-Francis, J. M, Rodriguez, B. and Burrage K. “Application of stochastic phenomenological modelling to cell-to-cell and beat-to-beat electrophysiological variability in cardiac tissue”, Journal of Theoretical Biology, Vol. 365, 21 January 2015, 325–336.

[32] Wang, H., Zheng, K., Zhou, X. and Sadiq, S. “SharkDB: An In-memory Column-oriented Trajectory Storage”, In Proceedings of The 23rd ACM International Conference on Information and Knowledge Management (CIKM 2014), Shanghai, China, Nov 2014

[33] Wang, H., Zheng, K., Zhou, X. and Sadiq, S. “SharkDB:An In-Memory Storage System for Massive Trajectory Data”, In Proceedings of ACM Conference on Management of Data (SIGMOD 2015), demo. Melbourne, Australia, June 2015.

[34] Wu, S., Li, W., Smarr, L., Nelson, K., Yooseph, S. and Torralba, M. “Large memory high performance computing enables comparison across human gut microbiome of patients with autoimmune diseases and healthy subjects”, Proceedings of the Conference on Extreme Science and Engineering Discovery Environment: Gateway to Discovery (XSEDE '13). ACM, New York, NY, USA, , Article 25 , 6 pages. DOI=10.1145/2484762.2484828 http://doi.acm.org/10.1145/2484762.24848

[35] Yu, H., Du, A., Song, Y. and Searles, D. “Graphyne and graphdiyne: Versatile catalysts for dehydrogenation of light metal complex hydrides” The Journal of Physical Chemistry C, 117, 21643-21650 (2013).