Institution: Swansea University

Unit of Assessment: 9

Title of case study: BSMBench: an open source tool for robust measurement of the performance of supercomputing systems

Period when the underpinning research was undertaken: 2008 – 2020

Details of staff conducting the underpinning research from the submitting unit:

Name(s):	Role(s) (e.g. job title):	Period(s) employed by submitting HEI:
Biagio Lucini	Professor	October 2005 – present
Edward Bennett	Lecturer, Senior Research Software Engineer	October 2014 – present
Antonio Rago	Postdoctoral researcher	2007-2009
Agostino Patella	Postdoctoral researcher	2008-2010

Period when the claimed impact occurred: January 2017 – December 2020

Is this case study continued from a case study submitted in 2014? N

1. Summary of the impact

High-performance computing (HPC) is the highest-end segment of the computer market in terms of computational power. Currently HPC hardware platforms (supercomputers) can outperform conventional desktop computers by 8-10 orders of magnitude and are used in cutting-edge applications in industry and research. BSMBench is an innovative performance-measurement software suite developed at Swansea University that has been widely adopted by leading supercomputing vendors. BSMBench robustly measures the performance of supercomputing systems and is the preferred benchmarking tool of the HPC-AI Advisory Council, the Dell/Intel Competence Centre for Cloud and HPC, Intel and NVIDIA Networking for the independent and vendor-neutral assessment of fast network interconnect (fabric) performance. In 2018, BSMBench was used by Intel to demonstrate to their customers the effectiveness of their solution to the highly concerning Spectre and Meltdown security vulnerabilities, which could have affected the sector to an estimated value of USD2,400,000,000.

2. Underpinning research

The parallel computer benchmarking tool BSMBench originated from Lucini's research in strongly interacting dynamics beyond the Standard Model (BSM) of particle physics. The Standard Model (SM) is the theory of how elementary particles interact through strong and electroweak forces. An essential component of the SM is the Higgs field, which generates the mass of the carriers of the weak interactions. The SM is believed to be incomplete and the main problem of the SM is to explain the stability of the Higgs mass. The fundamental research of Lucini focuses on the explanation of the stability of the Higgs mass through two paradigms known respectively as "near-conformal infrared dynamics" and "Higgs compositeness". According to these theories, the Higgs boson is not an elementary particle, as it is assumed in the SM, but a bound state generated by a novel and still undiscovered elementary strong force. In order to verify this hypothesis, a set of specific gauge theories obeying appropriate conditions must be studied. Since the novel force must be a strong interaction, quantitative progress in this field can be made only by using numerical simulations on large supercomputers performed within the framework of lattice gauge theory. The related investigations involve the study of the quantum dynamics of Yang-Mills theory with non-Abelian group G and fermionic matter in multiple representations of G. As G, the matter representation and the number of matter fields vary, the relevant models span the set of possible confining theories that are free at large energies. Current phenomenological bounds offer only loose guidance on the physically relevant models. Therefore, both G and the matter content need to be identified through the results of the numerical calculations. Having to study numerically wide classes of models presents non-trivial technical challenges in the development of the related research code. Consequently, the resulting programme of investigation is based on an interdisciplinary synergy involving particle physics (for informing the model building), mathematics (devising methods for fast numerical computations in the relevant classes of models) and high-performance computing (for the implementation aspects).

Starting from 2008, Lucini and his team at Swansea (working originally in collaboration with Prof. Luigi Del Debbio's group at Edinburgh University) developed a flexible and efficient software infrastructure called *HiRep*, a parallel Monte Carlo simulation framework for non-Abelian gauge theories discretised on a 4D spacetime lattice. In these types of simulations, the spacetime grid is spread across various nodes of the supercomputer, and the performance of the code is affected both by the computations that can be performed locally and the data that need to be communicated (generally, the boundaries of the local grid). Since *HiRep* was built to target a large class of models, its ability to simulate different gauge theories offers an opportunity to vary the relative relevance of local computation and communication [**R1, R2, R4, G1-G5**].

The inner communication structure of *HiRep* is based on an original implementation of *latency masking* for communication across the nodes of supercomputers. This method allows to overlay computation and communication, ultimately resulting in high computational efficiency. The devised computational strategy can be successfully adopted outside the specific research field in which it was developed to provide a performance measurement tool of the capabilities of supercomputers [**R3**], leading to the development of BSMBench in 2011. BSMBench is freely available from https://gitlab.com/edbennett/BSMBench. BSMBench extracts representative numerically and computationally intensive deterministic routines from *HiRep* and packages them in a software tool that exposes the number of operations performed in pre-set scenarios based on selected physical models, implementing validation of the result as a cross-check of correct execution. Building on the infrastructure provided by *HiRep*, BSMBench allows users to extract robust quantitative data for the performance of the software on a given hardware system and for different scenarios of relative importance of communications and local computations. Comparing the results across platforms, one can get a reliable measure of hardware efficiency.

3. References to the research

All papers have been peer reviewed and supported by external funding from STFC and from the Royal Society. [R1], [R2] and [R4] are published in Q1 journals (JCR 2019). [R3] is the result of a collaboration with a leading HPC industry, which provides the affiliation for one of the authors.

[R1]. Del Debbio, L., Lucini, B., Patella, A., Pica, C., Rago, A. (2010) The infrared dynamics of minimal walking technicolor. Physical Review D, 82, 014510. https://doi.org/10.1103/PhysRevD.82.014510

[R2]. Del Debbio, L., Patella, A., Pica, C. (2010) Higher representations on the lattice: Numerical simulations. SU(2) with adjoint fermions. Physical Review D, 81, 094503. https://doi.org/10.1103/PhysRevD.81.094503

[R3]. Bennett, E., Lucini, B., Del Debbio. L., Jordan, K., Patella, A., Pica, C., Rago, A. (2016) BSMBench: a flexible and scalable HPC benchmark from beyond the standard model physics. 2016 International Conference on High Performance Computing & Simulation (HPCS), Innsbruck, 2016, pp. 834-839. https://doi.org/10.1109/HPCSim.2016.7568421

[R4]. Bennett, E., Hong, D.K., Lee, J-W., Lin, C-J., Lucini, B., Piai, M., Vadacchino, D. (2018) Sp(4) gauge theory on the lattice: towards SU(4)/Sp(4) composite Higgs (and beyond). Journal of High Energy Physics 185. https://doi.org/10.1007/JHEP03(2018)185

The research programme has been supported by the following research grants:

[G1]. Lucini, B. (PI) (10.2015-09.2010) Royal Society University Research Fellowship. The Royal Society of London, [UF051628], GBP300,698.

[G2]. Lucini, B. (PI) (10.2010-09.2013) Royal Society University Research Fellowship. The Royal Society of London, [UF090003], GBP325,343.

[G3]. Lucini, B. (PI) (04.2008-03.2010) STFC Special Programme Grant, Lattice Gauge Theories Beyond QCD. STFC, [PP/E007228/1], GBP190,195.
[G4]. Hands, S., (PI) Lucini, B (Co-I) et al. (11.2019-10.2012) Computer Resources and Software Support for the UKQCD Physics Programme. STFC, [ST/H008829/1],GBP1,203,993.
[G5]. Lucini, B. (PI) (2016-2017) Wolfson Research Merit Award. Monte Carlo Investigations of Sp(2N) Composite Higgs Models. The Royal Society of London, [WM170010], GBP75,000.

4. Details of the impact

In the HPC environment, fast network interconnects (referred to as fabrics) are a key factor in determining the overall efficiency of a supercomputer under the highest workloads, which provide the most relevant usage scenario of BSMBench. In recent years, the two leading vendors of fabrics have been Intel and NVIDIA Networking (formerly Mellanox Technologies which was acquired by NVIDIA in 2020). Fabrics are complex and performance-critical hardware components. Assessing their robustness and performance needs a dedicated software tool that is flexible enough to emulate different workload scenarios. BSMBench has been the preferred tool of the HPC-AI Advisory Council (known as the HPC Advisory Council before 2018), the Dell/Intel Competence Centre for Cloud and HPC, Intel and NVIDIA Networking for assessing fabric performance in an independent, vendor-neutral way.

Since 2017, BSMBench has been used as a benchmarker by the HPC-AI Advisory Council, the leading and most influential independent advisory organisation in HPC, with over 400 member companies, including AMD, Intel, IBM, NVIDIA Networking and Microsoft. Its central mission is the promotion of best practice in the field. They have used BSMBench to perform an unbiased comparison of the performance of Mellanox InfiniBand (now NVIDIA Networking) adapters with respect to Intel Omni-Path [**C1**], IBM POWER8 processors versus x86 processor technology [**C2**] and of the use of a novel networking technology called adaptive routing versus more traditional methods [**C3**]. BSMBench was chosen for these performance measurements because it "includes the ability to tune the ratio of communication over computations" in contexts that "make up a significant fraction of supercomputing cycles worldwide". The tests were done "to provide best practice" and demonstrate "the scalability of the compute environment" [**C1-C3**]. In all cases, BSMBench has provided clear results that have steered market adoption of the best available product or feature. As an example of the information provided by BSMBench, Fig. 1 [from **C3**] shows the clear advantage in terms of performance when adaptive routing is used.

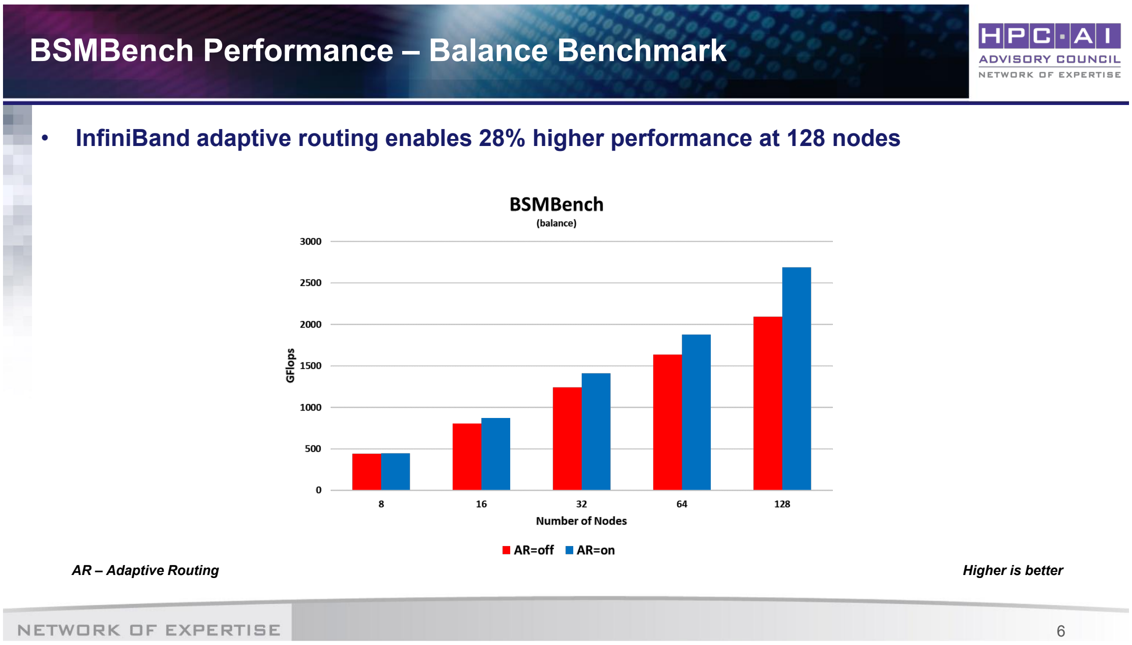

Fig. 1: Tests performed by the HPC-AI Advisory Council with BSMBench to contrast the performance of the fabric obtained with the use of adaptive routing with the case in which this recently developed feature is turned off, for variable size of the supercomputing partition [**C3**].

BSMBench is also widely used in three international supercomputing centres, namely the Dell/Intel Competence Centre for Cloud and HPC (Pisa, Italy), Supercomputing Wales (distributed across Cardiff and Swansea) and the Texas Advanced Computing Center (Austin, Texas), as a monitoring and performance tool. The Dell/Intel Competence Centre for Cloud and HPC is a joint venture between the University of Pisa and Dell EMC fully supported by Intel. One of the main objectives of the Centre is to design and test the next generation CPU and HPC technologies and solutions both in respect to performance and scalability using a suite of benchmarking platforms, of which BSMBench has been a key component since 2017. The CTO of the Centre says:

"BSMBench is very flexible by design, since it offers the possibility to test compute dominated, communication dominated, and balanced workloads. Since it is an independent benchmark, BSMBench offers a genuine measurement of the capabilities of CPUs and interconnects, since the performance results it provides are not informed by hidden features of commercial compilers that could limit the validity of the information one can extract. Scalable, adaptable to several environments and easy to use, BSMBench provides robust quantitative metrics and at the same time an immediate feeling of the capabilities of the HPC solution it is running on. Since adoption in February 2017, it has become our preferred suite for the early evaluation of new platforms" [**C4**].

Among major HPC suppliers, BSMBench has gained wide adoption by Intel and NVIDIA Networking, who use it to develop, test and promote to consumers their fabric and high-performance computing hardware offers. Through improved fault diagnostics, better testing processes and mitigation of losses, these two companies have been the main beneficiaries of the economic impact resulting from BSMBench in the REF period.

From 2018-2020, BSMBench was used at Intel as a performance metric to demonstrate how Intel® Omni-Path Architecture (Intel® OPA) helps maximize cluster performance across applications [**C5**] and to market the corresponding OPA fabric product line. BSMBench was deployed by Intel to compare latency, bandwidth, and message rate between it and a competitor [see **C5**, "application performance" section, OPA performance vs. InfiniBand* Enhanced Data Rate (EDR)]. This analysis enabled Intel to conclude that OPA offers to the customer better total cost of ownership, allowing for more compute or storage hardware to be procured within a fixed HPC budget. In September 2020, the Omni-Path Architecture was licensed by Intel to the newly formed company Cornelius Networks in a deal including a USD20,000,000 investment.

The ability to rely on robust metrics has been particularly important following the discovery of the Spectre and Meltdown security vulnerabilities in 2018. These vulnerabilities, which are caused by hardware design flaws and affect a wide class of modern processors commonly in use, including Intel processors, would allow attackers to execute malicious code on target systems, leaving the latter highly exposed to cyberattacks. These vulnerabilities can be bypassed by current processors only through the use of software fixes that degrade the overall performance of the system. Since the immediate workaround for those vulnerabilities resulted in a 10% reduction in the processing power, the associated cost resulting from the performance loss is estimated to be around USD2,400,000,000. This figure is calculated using market analysis performed by specialised companies such as Hyperion, which has found that the total HPC market, largely dominated by Intel, in 2017, i.e. immediately before the discovery of the vulnerabilities, had a volume of revenue of USD24,300,000,000. Using BSMBench as a measurement tool, Intel was able to demonstrate that even in the aftermath of Spectre and Meltdown the performance of their OPA product had increased [**C6** see first 2 figures for BSMBench application], preventing loss of reputation that could have resulted in decreased sales.

The High Performance Computing Account Manager at Intel UK notes:

[text removed for publication] [**C7**].

BSMBench is also a central part of the development processes at NVIDIA Networking. The SVP of Marketing at NVIDIA notes it "fills a hole in the market", thanks to its advanced design, and:

"Since 2017, we have used BSMBench on all our key systems to benchmark and develop our InfiniBand technology and to demonstrate to our customers the superiority of our HPC products in the field. We chose to use BSMBench because of the robust metrics it provides around the relative importance of communication and computation. In this respect, BSMBench is a unique

High-Performance Computing benchmarking tool that provides essential information on the performance of a supercomputing system. Therefore, when we present our systems to potential clients, we reference the BSMBench metrics to help prove the worth of our devices/systems/interconnects" [SVP of Marketing at NVIDIA, **C8**].

5. Sources to corroborate the impact

[C1]. HPC Advisory Council Best Practice "BSMBench Performance Benchmark and Profiling", January 2017 http://www.hpcadvisorycouncil.com/pdf/BSMBench-Performance-Benchmarking-and-Profiling.pdf (public domain)

[C2]. HPC Advisory Council Best Practice "BSMBench Performance Benchmark and Profiling", February 2017 http://www.hpcadvisorycouncil.com/pdf/BSMBench-Performance-Benchmarking-and-Profiling-IBM-Power8-Based-Cluster-2017-Feb.pdf (public domain)

[C3]. HPC-AI Advisory Council Best Practice "BSMBench Performance Benchmark and Profiling", August 2020, http://www.hpcadvisorycouncil.com/pdf/200804%20BSMBench_Performance_and_Profiling.pdf (public domain)

[C4]. Chief Technology Officer, The Dell | Intel Competence Centre Cloud and High Performance Computing (provided letter)

[C5]. Intel OPA official product page "Intel OPA tested for HPC" https://www.intel.com/content/www/us/en/high-performance-computing-fabrics/omni-path-architecture-performance-overview.html (public domain, retrieved 10/03/2019)

[C6]. Post by Intel High Performance Fabric Marketing Director in IT Peer Network. "Applications on Intel® Omni-Path Architecture Run 4.6 Percent Faster in 2018 than 2017—with Spectre and Meltdown Mitigations", July 2018 (https://itpeernetwork.intel.com/omni-path-architecture-faster/#gs.r379rk (public domain, retrieved 10/03/2019).

[C7]. High Performance Computing Account Manager, Intel UK (provided letter)

[C8]. SVP Marketing, NVIDIA Networking (provided letter).