Institution: University of Edinburgh

Unit of Assessment: 11

Title of case study: Novel architecture for high-performance computing accelerates data intensive software applications and enables release of new server lines

Period when the underpinning research was undertaken: 2015 – 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:
Adrian Jackson	Senior Research Fellow	2002 – present
Mark Parsons	Professor	1994 – present
Michèle Weiland	Senior Research Fellow	2006 – present

Period when the claimed impact occurred: 2016 – 2020

Is this case study continued from a case study submitted in 2014? No

1. Summary of the impact

Research led by the University of Edinburgh (UoE) in collaboration with Intel and Fujitsu has produced a ground-breaking novel architecture for high-performance computing (HPC) used, for example, in meteorological forecasting. UoE researchers worked with an existing Intel product (non-volatile RAM technology Optane DCPMM) to create a system of hardware and software named NEXTGenIO, that can be used for intensive memory storage. Fujitsu commercialised the resulting prototype into their new HPC server lines which have sold to multiple organisations, including Japan's Meteorological Research Institute. German system integration company SVA also used the research to attract approximately [text removed for publication] worth of new contracts. In further impact, the research led to software strategy changes for the European Centre for Medium-Range Weather Forecasts, permanently altering the way the centre's weather forecasts are created.

2. Underpinning research

Reading and writing data, in short referred to as I/O, can present a key performance bottleneck when scaling memory- and data-intensive applications on supercomputers. University of Edinburgh (UoE) researchers Adrian Jackson, Michèle Weiland, and Mark Parsons (Personal Chair in High Performance Computing) managed the development of the NEXTGenIO computing platform that accelerates I/O for high-performance computing (HPC) scientific and data applications. The NEXTGenIO platform was the key output of the UoE-led H2020 project Next Generation I/O for the Exascale (01/10/2015 – 30/09/2019). It exploits a new generation of byte-addressable non-volatile random-access memory (NVRAM). Jointly developed by Intel and Micron, the new NVRAM technology is known as Optane Data Center Persistent Memory (DCPMM).

The raw I/O performance of this NVRAM technology is extremely high – orders of magnitude better than even the fastest solid state disk – and although slower than conventional RAM, it has much greater capacity (on the order of TBs), and can maintain data without power being supplied to it. The UoE researchers developed methods to make DCPMM usable in HPC applications. The research was centered around the design and development of a platform architecture that exploits the new hardware, as well as the system software architectures to

support that hardware, based on the requirements of scientific and data-intensive computing applications [3.1].

DCPMM can be used in both memory and storage modes depending on how it is addressed. To enable users to make use of this functionality without being forced to make fundamental changes to the implementation of their HPC applications, the platform architecture must support secure application- and user-level access to both modes without requiring elevated system privileges. Additionally, the storage is local to individual compute notes, raising the challenge of access in a distributed HPC system.

The architecture developed by the UoE researchers includes software for data movement and management that enables node-local storage to be used by applications across compute nodes, and the development of new scheduling technologies that reduce the need to reconfigure the storage modes on individual nodes – an operation that takes 30 minutes. The research also developed methods to improve the efficiencies of end-to-end user workflows in the presence of NVRAM, through intelligently scheduling jobs to minimize the data movement between nodes, in accordance with workflow dependencies. Taken as a whole, these innovations enable applications to exploit performance benefits of NVRAM without significant changes or in-depth knowledge of the hardware [3.2].

Evaluations of the platform showed that there are significant performance and energy efficiency benefits for memory and I/O intensive applications, demonstrated across workloads including weather forecasting, computational fluid dynamics, and material modelling [3.3, 3.4].

For example, using DCPMM with the NEXTGenIO system software, weather forecasting applications, such as the European Centre for Medium-Range Weather Forecasts' IFS forecasting model, no longer experience any performance slowdown due to I/O operations, as data movement requirements of the forecasting software can be fully absorbed by the hardware – in contrast to using a network attached parallel file system. The demonstrated I/O bandwidth (i.e. the amount of data that can be transferred per second) enabled by the NEXTGenIO system architecture already supports the requirements of future generations of the IFS forecasting application that are only projected to be used in production within the 2025 timescale. Therefore, when utilizing this hardware, I/O is no longer a bottleneck for weather forecasting applications – not only in the current setting but for many years into the future. These features, unique to the NEXTGenIO architecture, have paved the way for the development of ground-breaking new products, which are now altering the possibilities for supercomputing in a range of fields, including weather forecasting and drug development.

The NEXTGenIO platform has now been commercialised by Fujitsu in their latest range of HPC servers.

3. References to the research

The papers below are some of the outputs of the underpinning research, outlining both the breadth of the research undertaken in this collaborative project, and its novelty. The papers describe the NEXTGenIO architecture [3.1] and how it can be exploited [3.2], as well as representing the first publications of performance results [3.3, 3.4] from the new hardware and software architectures in leading Computer Science conferences.

3.1. Jackson, A., Weiland, M., Parsons, M., & Homölle, B. (2019). An Architecture for High Performance Computing and Data Systems using Byte-Addressable Persistent Memory. In M. Weiland , G. Juckeland, S. Alam , & H. Jagode (Eds.), *High Performance Computing: ISC High Performance 2019* (pp. 258-274). (Lecture Notes in Computer Science). Springer. https://doi.org/10.1007/978-3-030-34356-9_21
3.2. Weiland, M., Jackson, A., Johnson, N., & Parsons, M. (2018). Exploiting the Performance Benefits of Storage Class Memory for HPC and HPDA Workflows. *Supercomputing Frontiers and Innovations*, *5*(1), 79-94. https://doi.org/10.14529/jsfi180105 **(6 citations)**
3.3. Miranda, A., Jackson, A., Tocci, T., Panourgias, I., & Nou, R. (2019). NORNS: Extending Slurm to Support Data-Driven Workflows through Asynchronous Data Staging. In *2019 IEEE International Conference on Cluster Computing (CLUSTER)* (pp. 1-12). Institute of Electrical and Electronics Engineers (IEEE). https://doi.org/10.1109/CLUSTER.2019.8891014 **(2 citiations)**
3.4. Weiland, M., Brunst, H., Quintino, T., Johnson, N., Iffrig, O., Smart, S., ... Parsons, M. (2019). An Early Evaluation of Intel's Optane DC Persistent Memory Module and its Impact on High-Performance Scientific Applications. In *SC'19 Proceedings of the International Conference for High Performance Computing, Networking, Storage, and Analysis* [76] ACM. https://doi.org/10.1145/3295500.3356159 **(8 citations)**

Citations based on Google Scholar, 2020-12-09.

Key research grants
European Commission: Next Generation I/O for Exascale (671591, GBP1,137,005)

4. Details of the impact

Through the NEXTGenIO platform, the University of Edinburgh (UoE) team's development of system architectures to support NVRAM enabled innovative products to be created for the supercomputing market, and led to commercial success for a number of companies, including Fujitsu, Intel, and SVA.

Fujitsu credits this work with enabling it to launch two server lines with capabilities unique to its business, ahead of the competition, and to subsequently reap commercial gain. Fujitsu states:

> *[text removed for publication]* [5.1]

In 2019, Fujitsu launched the following server lines, incorporating the NEXTGenIO technology [5.1]:

- PRIMERGY server models: TX2550 M5, RX2530 M5, RX2540 M5, RX4770 M5, CX2560 M5, CX2550 M5 and CX2570 M5
- PRIMEQUEST server models: PQ3800E2 and PQ3800B2

Fujitsu described the servers as "breakthrough", emphasizing their ability to speed up research discoveries and to model complex systems more accurately [5.2]. In their press material, Fujitsu explained that the UoE research enabled the company to build "an entirely new generation of computing platforms with the speed and performance required to power tomorrow's data-intensive and real-time applications" [5.2, penultimate para.]. In particular, the company cites the high performance of the NEXTGenIO platform, alongside its low power consumption, as a significant feature [5.3, para. 7]. The uniqueness of the product

has given Fujitsu a significant market edge, enabling the company to bring new HPC server models to the market ahead of their competitors [5.1].

To date, multiple large-scale PRIMERGY and PRIMEQUEST servers have been sold to specialist institutes in Japan and Australia, where they form the basis of supercomputers used in modelling and simulation [5.4, 5.5]. Their end purpose applications are varied and include medical diagnoses, autonomous driving technology [5.4, para. 4] and vaccine development in relation to the COVID-19 pandemic. [5.5, para. 6] In Australia, Gadi – a Fujitsu-server powered HPC utilising the NEXTGenIO technology – has achieved the country's highest ranking in the Global supercomputer index (25), and is reported to be eight times more powerful than its predecessor [5.5, paras. 1, 9].

In March 2020 Fujitsu announced the sale of their PRIMERGY CX2550 M5 server (which utilises the NEXTGenIO technology) to Japan's Meteorological Research Institute (JMRI), as part of a custom-built HPC architecture the company was commissioned to construct [5.3]. The institute uses supercomputing technology in the prediction of global warming, earthquakes, tsunamis, and volcanoes. JMRI's Senior Director stated that the new architecture meets the institute's needs "to accurately predict storm and heavy rain associated with typhoons as well as local downpours including linear precipitation zones that cause major disasters." [5.3, final para.]

Positive commercial impact has also cascaded down to Fujitsu client SVA, a German system integration company. SVA states [text removed for publication] [5.6, para. 3]. The company reports attracting approximately [text removed for publication] of contracts, by using the DCPMM architecture infrastructure and the NEXTGenIO platform long before their competitors were able to harness such niche knowledge in the HPC field. [text removed for publication] [5.6, paras. 4-5].

Intel reaped similar commercial gains, benefitting from the opportunity to refine an existing product – their DCPMM technology, originally designed for business computing – for an entirely new market in HPC. Significantly, the research project enabled Intel to experiment with the technology before its release, leading (as with Fujitsu) to a notable market advantage. The company states that:

> *[text removed for publication]* [5.7, para. 4]

In addition to its positive commercial impacts, the NEXTGenIO research has benefitted public services and accelerated strategy changes in professional practice.

The European Centre for Medium-Range Weather Forecasts (ECMWF) changed the strategy of its processes to monitor and predict weather, as a result of the enhanced capabilities offered by the NEXTGenIO architecture. ECMWF is the world leader in medium-range weather prediction, and states that NEXTGenIO has had "significant impact" on its software strategy and will "clearly affect its next supercomputer procurement" [5.8, para. 11].

Following the research, "ECMWF has been able to prove that such a platform can deliver staggering performance improvements…in particular for complex workflows…" [5.8, para. 8]. ECMWF calculates that the research project has generated cost savings in infrastructure investment and maintenance, and has enabled it to tighten its operational schedules, targets which "did not seem achievable" before NEXTGenIO [5.8, paras. 9-10]. Software

developments achieved during the project have already changed the running of weather forecasts at ECMWF, and the technology is now being used to facilitate a rolling archive of real-time climate monitoring data [5.8, para 11].

NEXTGenIO has created a chain of impact that starts with its economic benefits to multinational technological companies, and through those companies' clients, demonstrates HPC performances formerly not thought possible.

5. Sources to corroborate the impact

5.1. Letter of corroboration from Fujitsu Head of Category Management, Products Europe
5.2. Fujitsu Public Relations. (2019, August 13). Fujitsu Paves the Way to Exascale Computing With Powerful New NEXTGenIO Hardware Platform. Retrieved March 3, 2020, from https://www.fujitsu.com/emeia/about/resources/news/press-releases/2019/emeai-20190813-fujitsu-paves-the-way-to-exascale-computing.html
5.3. Fujitsu Limited. (2020, March 2). Fujitsu's New Supercomputer System for Japan's Meteorological Research Institute Commences Operations. Retrieved March 4, 2020, from https://www.fujitsu.com/global/about/resources/news/press-releases/2020/0302-02.html
5.4. Trader, T. (2020, February 3). Fujitsu A64FX Supercomputer to Be Deployed at Nagoya University This Summer. Retrieved October 23, 2020, from https://www.hpcwire.com/2020/02/03/fujitsu-arm64fx-supercomputer-to-be-deployed-at-nagoya-university/
5.5. Knowles, C. (2020, June 25). Australian supercomputer Gadi ranks 25 in Global 500. Retrieved October 23, 2020, from https://itbrief.com.au/story/australian-supercomputer-gadi-ranks-25-in-global-500
5.6. Letter of corroboration from SVA Executive Director
5.7. Letter of corroboration from Intel Vice President and General Manager, High Performance Computing
5.8. Letter of corroboration from ECMWF Director-General and Research Department Deputy Director