Thermal Neutrons: a Possible Threat for Supercomputers and Safety Critical Applications

Daniel Oliveira, Sean Blanchard, Nathan DeBardeleben, Fernando F. dos Santos, Gabriel Piscoya Dávila, Philippe Navaux, Stephen Wender, Carlo Cazzaniga, Christopher Frost, Robert C. Baumann, Paolo Rech

***Abstract*—The high performance, high efficiency, and low cost of Commercial Off-The-Shelf (COTS) devices make them attractive for applications with strict reliability constraints. Today, COTS devices are adopted in HPC and safety-critical applications such as autonomous driving. Unfortunately, the cheap natural Boron widely used in COTS chip manufacturing process makes them highly susceptible to thermal (low energy) neutrons.**

In this paper, we demonstrate that thermal neutrons are a significant threat to COTS device reliability. For our study, we consider an AMD APU, three NVIDIA GPUs, an Intel accelerator, and an FPGA executing a relevant set of algorithms. We consider different scenarios that impact the thermal neutron flux such as weather, concrete walls and floors, and HPC liquid cooling systems. Correlating beam experiments and neutron detector data, we show that thermal neutrons FIT rate could be comparable or even higher than the high energy neutron FIT rate.

I. Introduction

Reliability is one of the most important considerations in the field of High Performance Computing (HPC) [1], [2], [3]. An unreliable system can negatively affect not only the throughput of a computer but also the correctness of operations. Reliability can be increased through redundancies in chip architectures, improved manufacturing processes, transistor layout changes, or other hardening solutions [4]. However this added reliability comes at an added cost in terms of additional engineering, more expensive manufacturing, and added power consumption. This creates a trade off between lower cost and higher reliability such that only specialized safety critical industries, such as aerospace or medical, are willing to pay the additional cost of highly reliable parts. This is in contrast to Commercial Off-The-Shelf (COTS) devices which are generally not built to the highest achievable levels of reliability due to the low margins of the markets that consume these parts. Most consumers of COTS parts are primarily interested in performance and low price. They are typically willing to suffer lower reliability in exchange [4], [5]. The majority of the HPC community builds systems out of COTS parts and there is a constant struggle between the drive for ever increasing compute power and the potential of lower scientific productivity due to lower reliability [6].

In this paper we compare the reliability risk to HPC systems from *high energy* neutrons to that of boron-10 (^{10}B), which makes devices vulnerable to *thermal neutrons* generated from either fast neutrons that have lost energy through multiple interactions [5], [7] or are emitted from naturally occurring radioactive isotopes. ^{10}B has a relatively large capture cross section for thermal neutrons and the resulting excited state of ^{10}B quickly decays into Lithium-7 and a 1.47 MeV alpha particle. It is this high energy alpha particle that is known to contribute to upsets in semiconductors. Eliminating boron all-together or using depleted ^{11}B would make the device immune to thermal neutrons. However, depleted boron is expensive and boron is necessary for the manufacture of modern semiconductors, so many COTS devices contain ^{10}B. Modern data centers contain large masses of materials that can potentially increase the flux of thermal neutrons, in the form of concrete slab floors, cinder block walls, and water cooling units. In order to accurately estimate the effects of thermal neutrons we deployed a neutron detector to measure the natural background rate variation due to materials used in a modern data center. Our initial measurements indicate that these materials can increase the thermal neutron counts, and thus the COTS device's error rate, by as much as 20%.

The details of how ^{10}B is used in modern chips is proprietary and not publicly available. The only way to evaluate boron concentration in a chip, and the associated increased sensitivity to thermal neutrons, is through controlled radiation exposure. We studied the effects of fast and thermal neutrons on modern computing devices executing a representative set of benchmarks. We show that all the considered devices are vulnerable to thermal neutrons. For some devices, the probability for thermal neutrons to generate an error appears to be higher than the probability due to high energy neutrons.

The main contributions of this paper are: (1) an experimental evaluation of the probability for a high energy vs. thermal neutron to generate an error in modern computing devices; (2) an estimation of the thermal neutrons flux modification due to materials heavily present in a supercomputer room, based on homemade thermal neutrons detectors; (3) the evaluation, based on (1) and (2), of the contribution of thermal neutrons to the error rate of computing devices.

II. Background and Related Works

Radiation is a known cause of upsets in computers [8]. The interaction of particles, primarily neutrons for terrestrial machines, with transistors can reverse the value of the bits stored in memory or create current spikes in logic operations. These faults can create an undetected error known as Silent Data Corruption (SDC), or create a Detected Unrecoverable Error (DUE). It is well known that thermal neutrons can affect electronic devices [5], [7]; however, only devices containing ^{10}B are susceptible to thermal neutrons. Approximately 20% of naturally occurring boron is ^{10}B with the rest primarily

being ^{11}B. Depleted boron, where the ^{10}B content is low, is expensive in large quantities and generally not used in COTS parts.

Recently, ^{10}B was found in the manufacturing process of COTS devices [9]. It is worth noting that ^{10}B presence does not depend on the technology node but on the quality of the manufacturing process (smaller transistors will have less Boron, but also less Silicon. The Boron/Silicon percentage is not necessarily reduced). As devices produced for the user market are now employed in HPC and safety-critical applications, we must expect ^{10}B to be present. Some previous work has studied the sensitivity of SRAM and FPGA devices to thermal neutrons [10], [11], [12]. Weulersse et al. [13] compared the error rates of some memories (SRAM, CLB, caches) induced by thermal neutrons, 60MeV protons, and 14MeV neutrons.

A. High Energy and Thermal Neutrons

High energy neutrons, or fast neutrons, are neutrons with energies that range from 1 to over 1,000 MeV are known to disturb the function of electronic devices and are considered a main cause of faults in terrestrial electronic devices [5], [4]. The flux of high energy neutrons in the atmosphere has been thoroughly studied since Hess' discovery [14], [15]. The flux is known to vary across the surface, as a consequence of the earth's magnetic field, and increases exponentially with altitude, reaching a maximum at about 60,000 ft. Under normal solar conditions, the fast neutron flux is almost constant for a given latitude, longitude, and altitude.

Thermal neutrons, or slow neutrons, are low energy neutrons (lower than 0.5 eV), produced by the moderation of high energy neutrons in materials or the emission of neutrons from nuclear decay. The flux of thermal neutrons, in contrast to high energy neutrons, can be difficult to predict as it strongly depends on the environmental conditions as well the presence of other materials (primarily hydrogen containing) in the device's immediate surroundings (like concrete, water, a fuel tank, etc..) in addition to latitude, longitude, and altitude. Various authors have made calculations to evaluate thermal fluxes in realistic cases [7], [16], [17], [18]. As a result, when predicting the error rate caused by thermal neutrons, it is essential to measure rates in realistic settings.

We have built and deployed a neutron detector in order to have a precise understanding of the thermal neutron flux inside a representative data center. We measured the rates of thermal neutrons in the proximity of water, and demonstrate that cooling water, for instance, can increase the thermal neutron flux (and thus error rate) by up to 20%.

B. Supercomputer Cooling

One of the main challenges in designing HPC systems is the dissipation of heat. A modern supercomputer can push more than 750 watts per square foot which can easily overwhelm traditional cooling systems [19]. One notable and growing trend in data centers is the use of liquid cooling [20]. Eight of today's Top10 supercomputers use some form of liquid cooling [21]. These surrounding materials, as demonstrated in Section V, act as a moderator for neutrons energy and, thus, increase the thermal neutron flux.

III. Methodology

In this section, we describe the devices and applications chosen to test the impact of high energy and thermal neutrons in modern computing devices reliability. We also detail the radiation experiments setup used for this work and describe the detector we used to measure the impact of materials in the thermal neutron flux.

A. Devices

We selected six devices for this study using different technologies and vendors to have an in-depth insight of thermal neutrons sensitivity on a breadth of modern devices. It is worth noting that both the fabrication process and the foundry can significantly impact the amount of ^{10}B in the device.

Intel Xeon Phi is designed for HPC systems, built using a $22nm$ **Intel's 3-D Tri-gate technology**.

NVIDIA K20 is a GPU built with the $Kepler$ architecture and fabricated in a $28nm$ **TSMC CMOS technology**.

NVIDIA TitanX is a GPU built with the $Pascal$ architecture and fabricated in a $16nm$ **TSMC FinFET technology**.

NVIDIA TitanV is built with the $Volta$ architecture and fabricated in a $12nm$ **TSMC FinFET technology**.

AMD Accelerated Processing Unit (APU) integrates CPU and GPU in the same chip fabricated in a $28nm$ **SHP Bulk Process at Global Foundries**.

FPGA is the Zynq-7000 designed by Xilinx using a $28nm$ **TSMC technology**.

B. Codes

The set of devices we consider covers a wide range of architectural and computational characteristics. Using the same code for each device would bias the reliability evaluation, in favor of the devices that are more efficient in executing the chosen code. To have a fair evaluation, then, we choose for each class of devices the codes that better fit with its computational characteristics. For Xeon Phi and GPUs we chose four codes representative of **HPC**: MxM, LUD, LavaMD, and HotSpot. We selected three **heterogeneous** codes specially made to fully utilize the APU architecture: SC, CED, and BFS. Finally, on GPUs and FPGA we tested two **neural networks** to represent codes that have a significant impact on self-driven vehicles: YOLO and MNIST.

Matrix Multiplication (MxM) is representative of highly arithmetic compute-bound codes used in HPC and for features extraction in CNNs [21].

LUD is a linear algebra method that calculates solutions for a square system of linear equations.

LavaMD simulates particle interactions using Finite Difference Methods [22].

Hotspot is representative of stencil solvers [22], it estimates the processor temperature using an architectural floor plan and simulated power measurements.

Stream Compaction (SC) is a memory-bound code used in databases and image processing applications. SC is composed

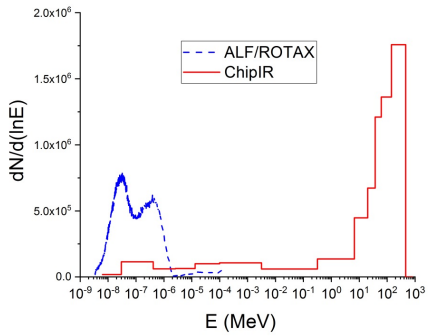

Fig. 1: The neutron spectra of the beamlines used for irradiation in lethargy scale.

of a data manipulation primitive that removes elements from an array.

Canny Edge Detection (CED) extracts information from images and reduce the amount of data to be processed. CPU and GPU concurrently work on different frames.

Breadth First Search (BFS) is a search in graphs algorithms that performs non-uniform memory access widely used in GPS Navigation Systems.

YOLO is a Convolution Neural Network (CNN) used for object classification and detection [23].

MNIST is a CNN used for classifying handwritten digits [24]. We have tested MNIST only on FPGAs as it is a minimal network that would not exercise sufficient resources on GPUs or Xeon Phis.

C. Radiation Experiments Setup

To evaluate the sensitivity of our devices to high energy and thermal neutrons we exposed the devices on two different beamlines at the ISIS spallation neutron source in the UK: ChipIR for high energy neutrons and ROTAX for thermal neutrons. ChipIR [25] is the reference beamline dedicated to the irradiation of microelectronics and it features a high energy neutron spectrum, as similar as possible to the atmospheric one. The flux with neutron energy above 10 MeV is $5.4\times10^6 n/cm^2/s$, while the thermal component ($E < 0.5eV$) is $4 \times 10^5 n/cm^2/s$ [26]. ROTAX [27] is a general purpose beamline with a thermal neutron spectrum generating a flux of $2.72 \times 10^6 n/cm^2/s$. Here the thermalization is achieved by moderation of the neutrons using liquid methane.

The spectra of the two beamlines are compared in Figure 1 on a log-log scale where the fluxes are proportional to the areas under the curves. As Figure 1 suggests, most neutrons in ROTAX are thermals and most neutron in ChipIR are high energy one.

To evaluate the sensitivity to thermal and high energy neutrons we align the devices with the beam, while executing the codes. The device output is compared with a pre-computed fault-free copy and any mismatch is marked as an SDC. If the device stops responding we count a DUE. To eliminate any setup-dependent differences between thermal and high energy neutrons, we irradiate the same physical devices executing the

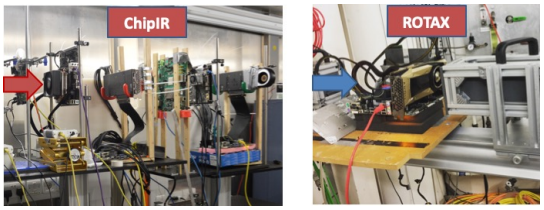

Fig. 2: Experimental setup in ChipIR and ROTAX. The arrow indicates the direction of the neutron beam.

codes with the same input both in ROTAX and in ChipIR. The only difference between the two experiments is that at ChipIR we can align various boards with the beam, as shown in Figure 2. Using a derating factor that takes distance into account we can measure the sensitivity of multiple devices in parallel. In ROTAX, as the irradiate device block most of the incoming neutrons, we must test one device at a time.

D. Thermal Neutrons Detector

We have designed and deployed a thermal neutron detector, called Tin-II, to measure the flux of thermal neutrons in different conditions. Ultimately, Tin-II will be used to measure the flux of thermal neutrons inside the data center housing the Trinity supercomputer at LANL.

Tin-II counted thermal neutron events over the course of several days. To estimate the effect of some of the characteristic materials in modern data centers on the thermal neutron flux, we placed a box containing 2 inches of water close to the detector. The count difference with and without the water, shown in details in Section V, indicates its influence in the thermal neutrons flux.

IV. Cross Section Results

In this section, we compare the cross section measured at ChipIR and ROTAX for the tested devices and codes with the methodology described in Section III-C. We emphasize that we used exactly the same device and setup for both ChipIR and ROTAX experiments. As we show, the cross section to thermal neutrons is far from being negligible, indicating the presence of ^{10}B in the silicon doping. Reported data have been normalized to the lowest cross section for each vendor to prevent the leakage of business-sensitive data while allowing a direct comparison between codes and devices of the same vendor. We also report error bars considering Poisson's 95% confidence interval.

Figure 3 shows the **Xeon Phi** SDC and DUE cross sections for high energy and thermal neutrons. On average the thermal neutrons cross section is much lower $(1/20)$ than the high energy neutrons' one, for both SDC and DUE. This low sensitivity to thermal neutrons is a sign that either little boron is used in the production of Xeon Phi or depleted boron is used.

For SDCs, the high energy neutron cross sections vary significantly depending on the code being executed (more than 2x across codes), which is in accordance with previous work [28], [29]. The SDC cross sections for thermal neutrons,

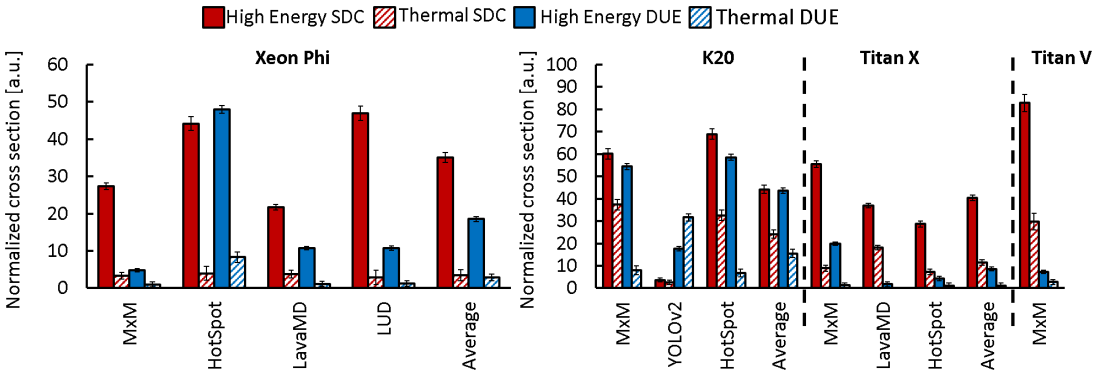

Fig. 3: High energy and thermal neutrons normalized cross sections for Xeon Phi and GPUs.

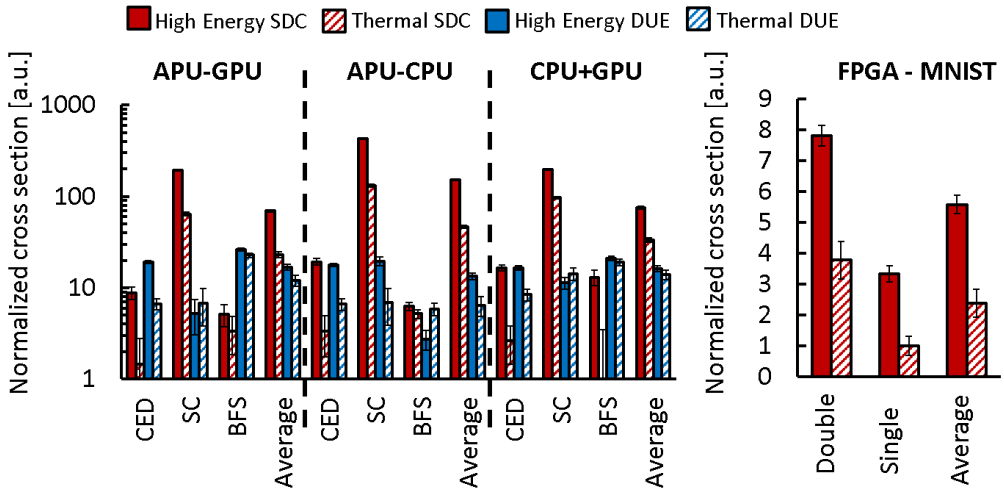

Fig. 4: High energy and thermal neutrons normalized cross sections for AMD APU and FPGA.

however, have a very low variation between codes (less than 20%) which may be an artifact of the low number of SDCs observed. This result suggests there is a negligible sensitivity to thermals in the chip resources that are responsible for the variation between error rates in the high energy SDC results. DUEs, on the other hand, have a similar trend for high energy and thermal neutrons.

Figure 3 shows the sensitivity of **NVIDIA GPUs** to thermal and high energy neutrons. For the K20, on the average, both the SDCs and DUEs thermal cross sections are very high, being 60% and 50% of the high energy neutrons ones. This indicates the presence of a significant amount of ^{10}B in the manufacturing process. The thermal neutrons SDC cross section trend across codes is also similar to the high energy neutrons one, in the sense that the code with the largest thermal neutrons cross section (i.e., HotSpot) is also the code with the largest high energy neutron cross section. This suggests that ^{10}B is present in the computing resources and memory of these devices, and that the fault locations are similar for both kind of neutrons. It is also interesting to notice that YOLOv2 is the only code for which DUEs are more likely than SDCs, for both kind of neutrons. This result follows previous work that shows low SDC sensitivity in CNN based object detection [30].

For Titan X and Titan V, on the average, the thermal neutron cross section is an order of magnitude lower than the high energy one. The impact of thermal neutrons is lower for the newest GPUs than on the mature K20. This may imply that FinFET based GPUs are less susceptible to thermal neutrons than CMOS GPUs (K20 is built using CMOS planar transistors, Titan X and Titan V using FinFET). However, for the MxM tests, Titan V ($12nm$) shows an almost doubled thermal neutron SDC cross section compared to the Titan X ($16nm$). Unfortunately, we were not able to test more codes on the Titan V to confirm if the increased thermal neutron cross section is intrinsic of smaller FinFET technologies.

The **AMD APU** cross sections are shown in Figure 4. As described in Section III-A, the APU embeds a GPU and a CPU. We test the three heterogeneous codes described in Section III-B (CED, SC, and BFS) as executed on the GPU only, on the CPU only, and distributing concurrently 50% of the workload to the CPU and 50% to the GPU (CPU+GPU).

The APU-GPU, APU-CPU, and CPU+GPU SDC cross section for both thermals and high energy neutrons vary of more than an order of magnitude, forcing the use of logarithmic scale for APU data in Figure 4. The reported data shows that, on the average, the thermal neutrons cross section is between $1/4$ and $1/5$ the high energy neutron's, for CPU,

GPU, and CPU+GPU. All APU configurations, on average, are more sensitive to SDCs than DUEs. It is also worth noting that the APU-CPU has, on average, a higher SDC sensitivity than APU-GPU. This is in accordance with previous work that shows a much lower probability for a fault in the AMD GPU to impact the application output than a fault in the CPU [31].

Figure 4 shows **Xilinx FPGA** SDC cross section when executing the MNIST CNN. It is worth noting that neutron-induced errors in the configuration memory of SRAM FPGAs have a *persistent* effect, in the sense that a corruption changes the implemented circuit until a new bitstream is loaded in the device. The observation of an error at the FPGA output indicates that the bitstream has probably been corrupted. We reprogram the FPGA at each observed output error to avoid the collection of a stream of corrupted data, making the observation of DUEs very rare. In fact, as FPGA executes operation without any operating system, interfaces, or control-flow involved, a considerable amount of errors would need to accumulate in the configuration memory to have the circuit functionality compromised. We never observed a DUE in FPGAs during our experimental campaign.

We have tested two different versions of the neural network, one using double and the other using single precision floating-point arithmetic. When comparing the high energy and thermal neutrons cross sections for the two configurations, we can clearly perceive that the Xilinx FPGA is more sensitive to high energy neutrons. However, the thermal neutrons cross section is far from being negligible.

The double precision version takes about twice as many resources to be implemented in the FPGA. As the neutrons cross section is directly related to the circuit's area, the cross section is expected to be higher for the double version of MNIST. Experimental results for both high energy and thermal neutrons confirm this intuition. The thermal neutrons cross section for the double version is particularly higher than the single one, being almost four times larger.

Our results show that different codes executed on the same device can have very different high energy vs thermal neutrons sensitivities. The physical interaction of a thermal neutron and, consequently, the resulting fault model (i.e., the way the physical fault is manifested at circuit level) and the impact on the code execution is highly different from the high energy neutron one. Software fault-injection can emulate predefined fault models and study their effects, but cannot be used to study the fault manifestation nor to define different fault models. One way to investigate the different fault models would be to simulate the physical implementation of a transistor in a given technology and observe the effect of neutron strikes at different energies [32]. However, transistor implementation details are not available for COTS devices, which makes the comparison of the beam experiment cross sections of various codes the only possible way to highlight code-dependent thermal vs high energy neutrons induced error rates.

V. FIT Rate Analysis

The cross sections reported and discussed in Section IV, represent the device's sensitivity to thermal or high energy

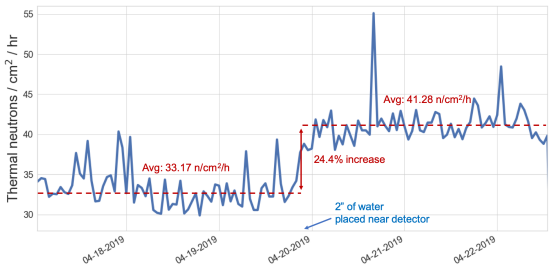

Fig. 5: Tin-II thermal neutron detector measurements with two inches of water placed over detector on 20^{th} April 2019.

neutrons. To have an understanding of the impact of thermal and high energy neutrons in the device error rate, we need to consider also the natural background radiation fluxes of each. FIT rates can then be calculated by multiplying the experimentally measured cross sections by the neutron fluxes. We only show in percentages the contribution of thermal and high energy neutrons to the device's FIT rates to avoid the leakage of business sensitive data. This information allows us to evaluate how much thermal neutrons increases the FIT of each device. This also tells us how much the FIT rate of each device is underestimated if thermal neutrons are not considered.

The flux for high energy (fast) neutrons in the atmosphere can be precisely estimated considering the altitude, longitude, latitude, and solar activity. However, the environment and the materials that surround a device significantly impact neutron flux and energy. For instance, during thunderstorms the rain droplets act as moderators slowing high energy neutrons into lower energy ones. The thermal neutron flux, as measured in [7], can be as much as $2\times$ higher during a rain storm than on a sunny day. Thermal neutron rates may be as much as 20% higher over a large slab of concrete such as in a parking lot or the concrete floor of a machine room. Water cooling systems can also have the side effect of significantly increasing the proportion of thermal neutrons that strike a device.

In order to empirically measure the impact of materials in the thermal neutron flux in a data center, we placed the Tin-II detector (details in Section III-D) in a building similar to the one containing the Trinity supercomputer. We collected data over the course of several days, then placed 2 inches of water in a box over the detector starting on 20^{th} April 2019. Figure 5 shows that when water is placed over the detector the thermal neutron counts abruptly increase of about 24%. This increase shows that the presence of water in the cooling system can significantly increase the rates of thermal neutrons in a system, which in turn will increase the rates in the devices sensitive to those neutrons as seen in section IV.

These same considerations exist when trying to understand the thermal neutron component of faults in autonomous vehicles. The road material, concrete or asphalt, the vehicle is driving on makes a difference, as does the weather, and the type and volume of fuel the vehicle uses. In addition, the number of passengers will change the thermal neutron flux, as humans are primarily composed of water which makes us excellent neutron moderators.

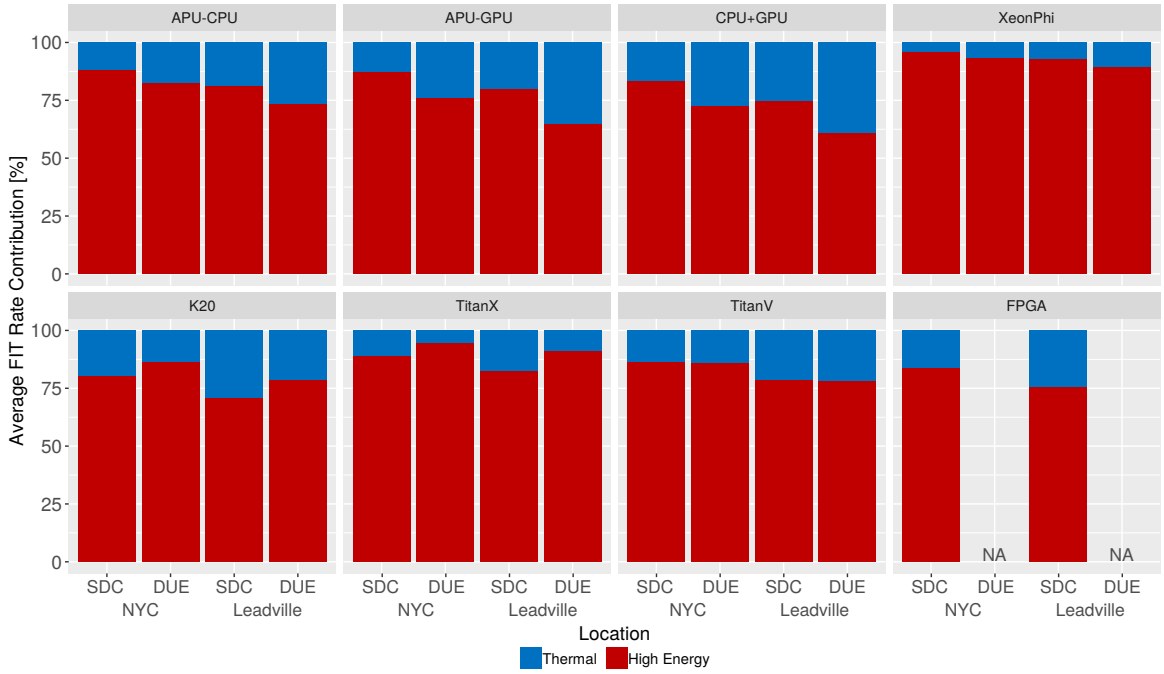

Fig. 6: Percentage of total FIT rate due to high energy and thermal neutrons. All tested parts except Xeon Phi show significant errors due to ^{10}B levels.

A. *High Energy vs Thermal Neutrons FIT*

Figure 6 shows the percentage of the total FIT rates due to high energy and thermal neutrons. These calculations use measured values of neutrons at sea level (NYC) and in Leadville, CO (10,151 ft in altitude). The thermal rates used have been adjusted to compensate for back scattered neutrons from a concrete slab and water cooling as measured by Tin-II detector, an overall increase of 44% in the thermal flux. Note that on a rainy day the thermal flux may be as much as doubled over the rates used in this graph and the corresponding FIT rate on those days will increase in a corresponding way [7].

Xeon Phi processors, as stated in Section IV, have a low sensitivity to thermals, which is a symptom of the use of either depleted boron or a reduction in boron usage. Thus, the thermals FIT rate seen in figure 6 is a relatively small percentage of the overall FIT rate (from 4.2% at NYC SDC up to 10.6% for Leadville DUE). The other tested devices, especially the K20 and CPU+GPU devices, have thermal FIT rates comparable to the FIT rates from high energy neutrons. At Leadville, K20 has 29% of the SDC FIT rate caused by thermal neutrons while APU CPU+GPU has 39% of DUEs caused by thermal neutrons.

B. *Discussion*

Figure 6 shows that if thermal neutrons contribution to the device error rate is not considered both the DUE and SDC FIT rates could be significantly underestimated, posing unconsidered risks to a safety critical application or reducing the HPC server productivity unexpectedly. Of particular interest in Figure 6 is the relatively high percentage of faults that result in Silent Data Corruption (SDC) on several of the tested devices. In general, HPC systems are designed and engineered to maintain SDC rates low and manageable, where corrupted calculations are rare and often noticeable to users. However, anything that increases the SDC rate is always concerning. In safety critical applications, SDCs should be strictly avoided as they could put the system in unexpected states, and they could potentially lead to unpredictable actions.

The elevated DUE rates are also of concern as they result in a system crash and loss of some portion of a calculation's run time. It is worth noting that even with thin layers of shielding, embedded devices in vehicles can suffer from a much higher thermal flux than the one considered in Figure 6 due to moderation and reflection from the surrounding materials [33].

VI. Conclusions

In this paper we have experimentally investigate the differences between high energy and thermal neutron induced error rates in modern HPC devices. By irradiating devices with high energy and thermal neutrons while executing representative applications, we have demonstrated that thermals significantly impact device reliability. We have demonstrated that the impact of high energy and thermal neutrons depends not only on the specifics of the hardware, but also on the executed code. The impinging neutron energy has more or less effect depending on how the code accesses memory and executes instructions.

We have also shown that the FIT rates can vary based on the physical layout of the machine room in which a system resides and variations such as weather conditions external to the building. The reported data attests the importance of thermal neutron reliability evaluation, which can significantly raise the total device error rate.

REFERENCES

[1] R. Lucas, "Top ten exascale research challenges," in *DOE ASCAC Subcommittee Report*, 2014.

[2] A. Cohen, X. Shen, J. Torrellas, J. Tuck, Y. Zhou, S. Adve, I. Akturk, S. Bagchi, R. Balasubramonian, R. Barik, M. Beck, R. Bodik, A. Butt, L. Ceze, H. Chen, Y. Chen, T. Chilimbi, M. Christodorescu, J. Criswell, C. Ding, Y. Ding, S. Dwarkadas, E. Elmroth, P. Gibbons, X. Guo, R. Gupta, G. Heiser, H. Hoffman, J. Huang, H. Hunter, J. Kim, S. King, J. Larus, C. Liu, S. Lu, B. Lucia, S. Maleki, S. Mazumdar, I. Neamtiu, K. Pingali, P. Rech, M. Scott, Y. Solihin, D. Song, J. Szefer, D. Tsafrir, B. Urgaonkar, M. Wolf, Y. Xie, J. Zhao, L. Zhong, and Y. Zhu, "Interdisciplinary research challenges in computer systems for the 2020s," National Science Foundation, USA, Tech. Rep., 2018.

[3] J. Dongarra, H. Meuer, and E. Strohmaier, "ISO26262 Standard," 2015. [Online]. Available: https://www.iso.org/obp/ui/#iso:std:iso:26262:-1:ed-1:v1:en

[4] J. Ziegler and H. Puchner, *SER–history, Trends and Challenges: A Guide for Designing with Memory ICs*. Cypress, 2004.

[5] R. Baumann, "Radiation-induced soft errors in advanced semiconductor technologies," *Device and Materials Reliability, IEEE Transactions on*, vol. 5, no. 3, pp. 305–316, Sept 2005.

[6] M. Snir, R. W. Wisniewski, J. A. Abraham, S. V. Adve, S. Bagchi, P. Balaji, J. Belak, P. Bose, F. Cappello, B. Carlson *et al.*, "Addressing failures in exascale computing," *International Journal of High Performance Computing Applications*, p. 1094342014522573, 2014.

[7] J. Dirk, M. E. Nelson, J. F. Ziegler, A. Thompson, and T. H. Zabel, "Terrestrial thermal neutrons," *IEEE Transactions on Nuclear Science*, vol. 50, no. 6, pp. 2060–2064, 2003.

[8] JEDEC, "Measurement and Reporting of Alpha Particle and Terrestrial Cosmic Ray-Induced Soft Errors in Semiconductor Devices," JEDEC Standard, Tech. Rep. JESD89A, 2006.

[9] S.-J. Wen, S. Pai, R. Wong, M. Romain, and N. Tam, "B10 finding and correlation to thermal neutron soft error rate sensitivity for srams in the sub-micron technology," in *2010 IEEE International Integrated Reliability Workshop Final Report*. IEEE, 2010, pp. 31–33.

[10] S. Lee, I. Kim, S. Ha, C.-s. Yu, J. Noh, S. Pae, and J. Park, "Radiation-induced soft error rate analyses for 14 nm finfet sram devices," in *2015 IEEE International Reliability Physics Symposium*, IEEE. IEEE, 2015, pp. 4B–1.

[11] Y.-P. Fang and A. S. Oates, "Characterization of single bit and multiple cell soft error events in planar and finfet srams," *IEEE Transactions on Device and Materials Reliability*, vol. 16, no. 2, pp. 132–137, 2016.

[12] P. Maillard, M. Hart, J. Barton, P. Jain, and J. Karp, "Neutron, 64 mev proton, thermal neutron and alpha single-event upset characterization of xilinx 20nm ultrascale kintex fpga," in *2015 IEEE Radiation Effects Data Workshop (REDW)*. IEEE, 2015, pp. 1–5.

[13] C. Weulersse, S. Houssany, N. Guibbaud, J. Segura-Ruiz, J. Beaucour, F. Miller, and M. Mazurek, "Contribution of thermal neutrons to soft error rate," *IEEE Transactions on Nuclear Science*, vol. 65, no. 8, pp. 1851–1857, 2018.

[14] V. F. Hess, "Über den ursprung der durchdringenden strahlung," *Z. Phys.*, vol. 14, p. 610, 1913.

[15] J. F. Ziegler, "Terrestrial cosmic rays," *IBM Journal of Research and Development*, vol. 40, no. 1, pp. 19–39, Jan 1996.

[16] A. Hands, P. Morris, K. Ryden, C. Dyer, P. Truscott, A. Chugg, and S. Parker, "Single event effects in power mosfets due to atmospheric and thermal neutrons," *IEEE Transactions on Nuclear Science*, vol. 58, no. 6, pp. 2687–2694, 2011.

[21] J. Dongarra, H. Meuer, and E. Strohmaier, "TOP500 Supercomputer Sites: November 2018," 2018. [Online]. Available: http://www.top500.org

[17] R. Baumann, "Soft error characterization and modeling methodologies at texas instruments," in *Proc. Semiconductor Research Council 4th Topical Conf. Reliability.[CD-Rom] SemaTech CD-ROM*. USA: SemaTech, 2000, pp. 0043–3283.

[18] R. Sheu and S. Jiang, "Cosmic-ray-induced neutron spectra and effective dose rates near air/ground and air/water interfaces in taiwan," *Health physics*, vol. 84, no. 1, pp. 92–99, 2003.

[19] M. K. Patterson and D. Fenwick, "The state of datacenter cooling," *Intel Corporation White Paper. Available at http://download.intel.com/technology/eep/data-center-efficiency/stateof-date-center-cooling. pdf*, 2008.

[20] A. Capozzoli and G. Primiceri, "Cooling systems in data centers: state of art and emerging technologies," *Energy Procedia*, vol. 83, pp. 484–493, 2015.

[22] S. Che, M. Boyer, J. Meng, D. Tarjan, J. W. Sheaffer, S.-H. Lee, and K. Skadron, "Rodinia: A benchmark suite for heterogeneous computing," in *Proceedings of the IEEE International Symposium on Workload Characterization (IISWC)*. Austin, TX, USA: IEEE, 2009, pp. 44–54.

[23] J. Redmon, S. K. Divvala, R. B. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," *CoRR*, vol. abs/1506.02640, 2015. [Online]. Available: http://arxiv.org/abs/1506.02640

[24] L. Deng, "The mnist database of handwritten digit images for machine learning research [best of the web]," *IEEE Signal Processing Magazine*, vol. 29, no. 6, pp. 141–142, 2012.

[25] C. Cazzaniga and C. D. Frost, "Progress of the scientific commissioning of a fast neutron beamline for chip irradiation," in *Journal of Physics: Conference Series*, vol. 1021, IOP Publishing. IOP, 2018, p. 012037.

[26] D. Chiesa, M. Nastasi, C. Cazzaniga, M. Rebai, L. Arcidiacono, E. Previtali, G. Gorini, and C. D. Frost, "Measurement of the neutron flux at spallation sources using multi-foil activation," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 2018.

[27] H. Tietze, W. Schmidt, and R. Geick, "Rotax, a spectrometer for coherent neutron inelastic scattering at isis," *Physica B: Condensed Matter*, vol. 156, pp. 550–553, 1989.

[28] D. Oliveira, L. Pilla, N. DeBardeleben, S. Blanchard, H. Quinn, I. Koren, P. Navaux, and P. Rech, "Experimental and analytical study of xeon phi reliability," in *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, ser. SC '17. New York, NY, USA: ACM, 2017, pp. 28:1–28:12. [Online]. Available: http://doi.acm.org/10.1145/3126908.3126960

[29] V. Fratin, D. Oliveira, C. Lunardi, F. Santos, G. Rodrigues, and P. Rech, "Code-dependent and architecture-dependent reliability behaviors," in *2018 48th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*, IEEE. IEEE, 2018, pp. 13–26.

[30] F. F. d. Santos, P. F. Pimenta, C. Lunardi, L. Draghetti, L. Carro, D. Kaeli, and P. Rech, "Analyzing and increasing the reliability of convolutional neural networks on gpus," *IEEE Transactions on Reliability*, vol. 68, no. 2, pp. 663–677, June 2019.

[31] H. Jeon, M. Wilkening, V. Sridharan, S. Gurumurthi, and G. H. Loh, "Architectural vulnerability modeling and analysis of integrated graphics processors," in *IEEE 10th Workshop on Silicon Errors in Logic - System Effects (SELSE)*. IEEE, 2013.

[32] P. E. Dodd, "Physics-based simulation of single-event effects," *IEEE Transactions on Device and Materials Reliability*, vol. 5, no. 3, pp. 343–357, Sep. 2005.

[33] W. R. Leo, *Techniques for nuclear and particle physics experiments: a how-to approach*. Springer Science & Business Media, 2012.