#25 Thermal Neutrons: a New Threat for Supercomputers and Safety Critical Applications

Email notification
Select to receive email on updates to reviews and comments.

PC conflicts
None

▶ **Other conflicts**

Submitted

Submission 6 Aug 2019 6:14:32pm PDT · 33890c04

▶ **Abstract**

The high performance, high efficiency, and low cost of Commercial Off-The-Shelf (COTS) devices make them attractive for applications with strict

[more]

▶ **Authors (blind)**

D. Oliveira, S. Blanchard, N. DeBardeleben, F. Santos, G. Dávila, P. Navaux, A. Favalli, O. Schappert, S. Wender, C. Cazzaniga, C. Frost, P. Rech [details]

▶ **Topics**

	OveMer	RevExp	WriQua
Review #25A	2	2	2
Review #25B	4	4	4
Review #25C	3	4	4
Review #25D	2	2	4
Review #25E	2	2	3
Review #25F	2	3	3

You are an **author** of this submission.

Edit submission Add response

Reviews in plain text

Review #25A

Overall merit

2. Weak reject

Reviewer expertise

2. I have passing familiarity with this area

Paper summary

This paper performed empirical studies to quantify error rates caused by high-energy versus low-energy (thermal) neutron. The error rate characterization covers a set of applications running on a wide range of computer systems that are manufactured with different CMOS technologies. The paper shows that error rates can be significantly affected by low energy/thermal neutron flux, coming from weather conditions, materials of the datacenter buildings, or materials of cooling solutions.

Strengths

- The SDC/DUE error rate characterization studies and results are interesting and can help inspire new error-mitigating solutions.

Weaknesses

- It is difficult to interpret the error rate results with many dimensions of the experimental methodology varying simultaneously.
- It is unclear why the experiments were evaluated without state-of-the-art ECC solutions in place.
- The applications used in the evaluation are not representative of HPC workloads, especially for real-system studies.

Writing quality

2. Needs improvement

Comments for author

This paper studies an interesting phenomenon, that is how error rates react to neutrons at different energy levels (high energy vs. thermal/low energy neutrons) for HPCs. The study shows that materials used in commercially off the shelf (COTS) devices, cooling solutions, or datacenter buildings can increase thermal neutron counts, thereby increasing COTS error rate by as much as 20%. The error rate trend and characterization results presented in the paper can help inspire new error mitigation techniques.

There are several aspects of the work that can be strengthened:

First, the applications used in the evaluations are not representative of HPC workloads. This is especially problematic when the characterization studies were performed on real HPC systems.

Furthermore, the reviewer also found it challenging to interpret the error rate results with many dimensions of the experimental methodology varying simultaneously. For example, different applications exhibit different error tolerant nature. In addition, different system components (memory vs. compute) also exhibit different error characteristics. System architecture and microarchitecture also play a role in error masking/tolerance. Therefore, it'd make the work much stronger if the paper can design the experiments to isolate the effect of different aspects of a system on error rates under different circumstances.

Next, it is unclear why the experiments were evaluated without state-of-the-art ECC solutions in place. Reliability is critical for HPCs; thus, ECCs are deployed in server-class systems. How does the error characterization results change with ECC-enabled? Are state-of-the-art error mitigation techniques sufficient for thermal-neutron threats?

Questions for authors' response

1. How does the error characterization results change with ECC-enabled? Are state-of-the-art error mitigation techniques sufficient for thermal-neutron threats?

2. Can you please explain the choice of applications and how the experimental setup is representative to HPCs/safely critical applications?

Review #25B

Overall merit

4. Accept

Reviewer expertise

4. I know a lot about this area

Paper summary

An experimental study of thermal neutrons' impact on commercial chips with Boron-10 isotope was conducted with actual hardware systems running programs. Results show that due to the re-introduction of B-10 in silicon process, thermal neurons' impact should not be ignored.

Strengths

Well written paper with data collected on actual systems running with software. This study is important for supercomputers and safety critical applications.

Weaknesses

Some data are missing. For example, Titan V (Fig 5) has only one program.

Writing quality

4. Well-written

Comments for author

1. As the paper seems to suggest (from DRAM) that technology scaling should help ease the impact of radiation as the amount of B-dopant (diffusion) is reduce. It is not clear why Titan V is worst than Titan X since Titan V is manufactured with a more advanced technology.
2. The explanation given as why thermal neutrons' impact on SDC is independent of codes is not totally convincing.
3. It would be nice to collect data for different generation of FPGAs. It would be nice to collect data on SRAM failure rate for FPGAs.

Questions for authors' response

1. We know as technology scales, the amount of B-10 is reduced. Can you offer an explanation on why GPU Titan V is worst than Titan X?
2. Do different DRAM vendors produce different directional flips or the directional is only due to DDR/DDR4 difference? Can you estimate the technology nodes used for DDR4 and DDR3? For derivation the chips are 4Gb for DDR3 and 8Gb for DDR4.

Review #25C

Overall merit

3. Weak accept

Reviewer expertise

4. I know a lot about this area

Paper summary

The paper presents the results of an experimental evaluation of real systems under thermal neutrons and discovers that thermal neutrons can impact the reliability of the systems considerably. The paper compares real systems running a variety of workloads under exposed to both high-energy and thermal neutrons. The authors build a thermal neutron detector and discover that thermal neutrons increase in number in datacenter due to material structure of data centre as well as the water based cooling infrastructure. The results show that the effect of thermal neutrons on various real HW is not negligible although still much smaller than that of high-energy neutrons.

Strengths

The paper is an experimental analysis paper that tests and evaluates thermal neutron radiation effect on various real-systems.

Weaknesses

It is clear that thermal neutrons contribute to the overall FIT rate of the system. However, it is not

clear whether the system designers and/or reliability engineers should treat thermal neutrons differently when developing fault/failure mitigation strategies.

Writing quality

4. Well-written

Comments for author

In general, I like this paper. It is a rigorous study evaluating the reliability of real systems under thermal neutrons in particular HPC systems where low-energy neutrons can be relatively higher.

I have the following 3 questions in the question box and a minor comment as follows:

MINOR COMMENT

In Section 2.1, the reference [11] is cited as a recent publication but the paper was published in 2010, which does not make it very recent.

Questions for authors' response

1) I'd expect that YOLOv2 would have less DUEs than SDCs because after all CNNs are just pattern recognition applications with a lot parallel data computations with very little control code. Pattern recognition applications should show more mismatched outputs at the end of a run if there is bit upset in some flip-flop due to a thermal neutron rather than getting stuck or crash. I have found counter-intuitive that this application on GPU gets more DUEs than SDCs. Can you please elaborate on the possible reasons why this applications observes more DUEs?

2) The abstract of the paper says at the end that thermal neutrons FIT rate could be comparable or even higher than the high energy neutron FIT rate. However, Figure 11 contradicts this in that none of the figures show higher thermal neutron FIT rate than the high-energy neutron FIT rate. Have I missed something?

3) How will this change the fault/failure detection and correction mechanisms already used in safety-critical HW? Should the system designers and architects treat thermal neutrons differently when they develop mitigation schemes? For example, for autonomous vehicles, how would lockstep CPU/GPU architectures used today should change? Similarly, should the ECC in SRAMs and DRAMs change or they are sufficient to detect and correct errors caused by thermal neutrons? Could you provide some discussion about this topic?

Review #25D

Overall merit **Reviewer expertise**

2. Weak reject

2. I have passing familiarity with this area

Paper summary

The paper evaluates the sensitivity of several CPU and GPU devices to thermal low-energy neutrons. While several papers have evaluated high-energy neutrons impact on reliability, no prior papers have focused on the low-energy neutrons that silicon doped with boron-10 is particularly susceptible to. The paper also evaluates several interesting environmental conditions that influence the rate of low-energy neutrons.

Strengths

The paper is very well-written and provides a lot of interesting field analysis.

Weaknesses

The paper evaluates CPUs and GPUs that are a few years old and it is unclear whether chips implemented in more recent technology are susceptible to the same issues.

Writing quality

4. Well-written

Comments for author

The paper is outside my area of expertise, but I found it generally interesting. It was particularly interesting to hear that the use of water cooling may increase the proportion of thermal neutron strikes. However, I did not rank it high because it is unclear to me whether this is a problem in the most recent technology generations. Specifically, it would be interesting to see the paper analyze CPUs or GPUs produced in TSMC's 7 nm process or Intel's 14 nm process. As the paper points out, Intel's 22 nm process did not seem to have much of problem and the authors suspect that is due to the use of a different dopant. I suspect other foundries have made a similar shift and my understanding is the impact of low-energy neutrons drops by multiple orders of magnitude if one avoids boron-10.

Please do not use references as nouns. References are parenthetical and should never be directly referred to. For example, page 6 uses references #42-44 and #39 as nouns. By the way, that list of DDR memories experimented using high energy neutrons should also include Sridharan et al. ASPLOS 2015.

Questions for authors' response

Do you think this is a problem in the most recent technology generations?

Review #25E

Overall merit

2. Weak reject

Reviewer expertise

2. I have passing familiarity with this area

Paper summary

This paper evaluates the impact thermal neutrons on commercial off-the-shlef devices due presence of Boron-10 in these devices. It characterizes the FIT rate of various commercial accelerators due to thermal neutrons on HPC applications.

Strengths

This paper sheds lights on potential vulnerability due to Boron-10 that may not be as well appreciated.

Weaknesses

It is not clear if HPCA is the right conference for this paper. The material covered in the paper may not be easy to appreciate by the chip/system/software architects in the audience. ISC and SC, which are top HPC conferences, may better fit for this paper.

The paper does not offer a solution for the problem that may be within the scope of system architecture. As a mitigation, the discussion section talks about needing to "consider the realistic conditions in which the device will operate ..". This is something that the IT or procurement department can act on, not a computer architect. Such audience are more likely to be at ISC and SC than at HPCA.

Writing quality

3. Adequate

Comments for author

- How was Tin-II calibrated to make sure it was detecting and reporting the thermal neutron events correctly?
- Is Boron-10 is only why thermal neutrons can cause errors? Are there reasons where thermal neutrons can cause errors even in the absence of Boron-10? Would be good to have that zero Boron-10 baseline, otherwise attributing all the errors due to Boron-10, may be misleading.

Review #25F

2. Weak reject

3. I know the material, but am not an expert

Paper summary

Low cost and high performance of commercially off-the-shelf (COTS) components made them popular in building HPC and safety-critical systems. However, many applications running on these systems need a high degree of reliability guarantee. Unfortunately, Boron-ten atoms (10B) which are present as impurity in the manufacturing process can introduce errors when they are hit by alpha particles and emit low-energy neutrons.

The authors extensively study how high-energy and low-energy neutrons affect the commonly used COTS HPC components. The authors use a thermal neutron detector to radiate different system components at different environments. The evaluation characterizes the impact of neutrons based on operating environment and processor design/utilization.

Strengths

+ Interesting research direction as the impact of radiation has not been extensively studied for modern-day HPC components.

+ Demonstrates the effect of thermal neutrons at different environments (e.g., near water cooling system or concrete slabs), operating conditions, and geographical locations of an HPC facility.

Weaknesses

- Does not publish real experimental numbers because of vender confidentiality

- Shows per component failure but exposes the whole system to radiation

- Not clear how the metric "cross-section" should be interpreted

Writing quality

3. Adequate

Comments for author

The authors did a fantastic job in motivating the problem of low-energy neutrons and the impact of Boron-10 impurity. The reviewer feels that it can be an exiting paper and can make a huge impact if the authors on focus on more detailed analysis. The paper falls short in different avenues in its current state.

- The metric used is not clear The paper uses the metric "cross-section" as a way to show the impact of error per neutron without considering the difference in the hardware components. It is not clear what is the definition of this metric and how it should be interpreted compared to more popular metrics, such as FIT. It is also not clear how to judge the difference in this metric when the evaluated systems are completely different (CPU, GPU, FPGA, memory).

- The results are not meaningful without absolute numbers A characterization paper is particularly interesting for two reasons: (i) it provides an interesting data points to make some conclusions and (ii) it provides characterization numbers that can be used by the future works. This work is not being able to meet any of these categories. The metric used makes it hard to make any conclusion, and not providing the actual numbers make it impossible for others to use the characterization in the future.
- Lack of extensive analysis on memory errors or FPGA/CPU/GPU errors The authors are making an apples to oranges comparison here. One DDR3 module vs one DDR4 module does not give us much information unless we know the vendor, generation, capacity, true-anti cell combination, workloads tolerance to error, etc. A useful study needs to characterize the impact of neutrons for different modules from the same vendor and same generation and then can show a trend across different maturating process or node technology.
- Methodology regarding per component vs full system beaming is not clear It is not clear what is the impact of aligning the neutron stream to a silicon component. The paper mentions that the full module/card was exposed to neutron stream, which means that not only the specific component tested but also all the circuit components were exposed to the neutron stream. It is not clear how this work makes conclusion about one particular component.

Questions for authors' response

- What will be the impact of high and low energy neutrons in regular environment?
- Does the beam hit the full module/card?
- How should the reader interpret the cross-section metric against bit-flips, FIT, or raw failure rate?

Response

Author [Daniel Oliveira] 22 Oct 2019 **1007 words**

Response submitted.

We thank all reviewers for their valuable comments.

Reviewer-A:

The review raised an interesting point about ECC. Actually, among the tested devices only Xeon-Phi includes ECC (which was enabled). As shown, even with ECC enabled both thermals and high energy

neutrons produce errors. For the other devices, the error rate with ECC enabled can be derived by looking at the fault model (for DDRs all transient faults would be correctable) or based on previously published results (ECC reduces of about one-order-of-magnitude the error rate [1]). We will include a discussion in the paper.

We agree with the reviewer that we test different codes on different devices. Unfortunately, testing the same code in all devices would be unfair (heterogeneous codes do not map well in GPUs, and parallel codes are not efficient on CPUs), so we decided to provide an operative evaluation based on the actual expected utilization of each device using established benchmark suites.

Reviewer-B:

We actually cannot state that B-10 is reduced as technology scales. Since commercial devices are used in HPC or safety-critical applications, we expect B-10 to be present (reducing costs and improve competitiveness). Xeon-Phi (the only non-consumer device), in fact, has the lowest thermal neutron sensitivity. The difference between Titan-X and Titan-V is then in process quality [4,5].

The different direction flips for DDRs are probably due to the use of complementary logic by one of the vendors. The technology nodes were between 22 and 19 nm. We will add this discussion in the paper.

Reviewer-C:

We thank the reviewer for the interesting comments. DUEs are more frequent than SDC in YOLOv2 as it launches several kernels and makes several host-device synchronizations, stressing PCI-express and GPU synchronization logic. This behavior was also shown in [6].

The reviewer is right, and the thermal FIT rate is not higher than the high-energy one. We will rewrite the abstract to reflect Fig. 11 results (highest thermal contribution is 40% for CPU+GPU).

Very interesting point. Any reliability mechanism should be tuned considering thermal-neutrons contribution. A mechanism that guarantees 10FIT considering only high-energy neutrons might not be sufficient for automotive applications or a checkpoint-restart frequency based only on high-energy neutrons might not be optimal. Moreover, since thermal and high-energy neutrons have different impacts on memory and logic, the effectiveness of the adopted mechanism is expected to be different. We will add a discussion on the final paper.

Reviewer-D:

We thank the reviewer for sharing her/his concern. Actually, Boron-10 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 Boron-10 to be present. In fact, purified Boron is expensive and would increase the device price (unjustified for user applications). Xeon-Phi is the only high-end device we tested, and it

shows the lowest thermal-neutrons sensitivity (but still suffers from thermal-neutron errors). We will add a discussion to the motivation of our paper and correct grammar errors.

Reviewer-E:

The detector was calibrated both using the procedure presented in [8] and comparing the counts with expected fluxes presented in [9].

As confirmed by various publications [4,5,7], Boron-10 is the only significant cause of thermal neutron errors in computing devices.

We believe that both software and device architects should be aware of the thermal-neutrons problem. This is the only way to guarantee a safe use of commercial devices in HPC and safety-critical applications.

We will include a discussion on these topics in the paper.

Reviewer-F:

The cross-section measures the probability of a neutron (high-energy or thermal) to induce an error and is an intrinsic characteristic of the device executing a code while FIT rates depend also on the flux of neutrons. The reported cross-sections show the sensitivities to thermal and high-energy neutrons of the tested devices, indicating a high or low presence of Boron-10. Then, we also include FIT rates (Fig.11) considering various scenarios. As the fluxes of high-energy and thermal neutrons are different, we believe that presenting only FIT rates would prevent a deep analysis of the sensitivity of the device.

While it would be interesting to have absolute FIT rates, we still believe that normalized data allows the evaluation of the Boron-10 problem. Absolute FIT rates for high-energy neutrons can be found in literature ([2] for Xeon-Phi, [11] for FPGA, and [10] for DDR). We will include this in the final paper as well as DDR details (without mentioning the vendor).

We selected a neutron beam of 2-inch diameter to irradiate only the computing chip and not the surrounding circuits. We will make this clear in the final paper.

[1] Evaluation and mitigation of radiation-induced soft errors in graphics processing units. IEEE Transactions on Computers, 2015, 65(3), 791-804.

[2] Experimental and analytical study of xeon phi reliability. Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (2017).

[3] Code-dependent and architecture-dependent reliability behaviors. 48th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN 2018).

[4] Thermal neutron-induced soft errors in advanced memory and logic devices, IEEE Transactions on Device and Materials Reliability, vol. 14, no. 1.

[5] Thermal neutron soft error rate for srams in the 90nm–45nm technology range," 2010 IEEE International Reliability Physics Symposium.

[6] Analyzing and Increasing the Reliability of Convolutional Neural Networks on GPUs," IEEE Transactions on Reliability, vol. 68, no. 2, 2019.

[7] Neutron-induced 10b fission as a major source of soft errors in high density srams, Microelectronics Reliability, vol. 41, no. 2, 2001.

[8] Thermal neutron flux characterization at aircraft altitudes with the TinMan detector, IEEE Aerospace Conference, Big Sky, MT, 2017.

[9] Terrestrial thermal neutrons, IEEE Transactions on Nuclear Science, vol. 50, no. 6, Dec. 2003.

[10] Memory errors in modern systems: The good, the bad, and the ugly. ACM SIGPLAN Notices, v. 50, n. 4, p. 297-310, 2015.

[11] Neutron Radiation Beam Results for the Xilinx UltraScale+ MPSoC, IEEE Radiation Effects Data Workshop (REDW), Waikoloa Village, HI, 2018.