SEE Evaluation of a Serverless Computing Architecture under 14-MeV Neutrons

David Pacios, Mohammadreza Rezaei, José Luis Vázquez-Poletti, Sara Ignacio-Cerrato and Juan A. Clemente

***Abstract*—This paper presents the effects of 14-MeV neutrons on a NVIDIA Jetson Nano running a serverless computing architecture. SDCs and DUEs were observed in its CPU and GPU when executing a serverless-based FFT function.**

***Index Terms*—Neutron Radiation, Serverless Computing, Detected Unrecoverable Errors, Silent Data Corruption.**

I. Introduction

The increasing use of electronic components in radiation-harsh environments, such as space and high-altitude atmospheric locations, necessitates a comprehensive understanding of how neutron radiation affects these devices. This study specifically focuses on the resilience and vulnerability of serverless computing systems (comprising CPUs and GPUs) to neutron radiation. These systems are becoming increasingly essential in high-performance computing environments where reliability under radiation exposure is critical.

The literature demonstrates how radiation can impact the performance and reliability of semiconductor devices [1], [2]. Furthermore, the robustness of Multi-Processor Systems-on-Chip (MPSoCs) against radiation-induced Single Event Effects (SEEs) is well documented [3], highlighting the need for targeted error mitigation strategies in critical computing environments.

Building on these insights, this research employs an experimental approach by simulating a serverless computing environment on an NVIDIA Jetson Nano subjected to 14-MeV neutron radiation. The objective is to understand how these conditions affect computational accuracy and performance during high-complexity tasks—specifically, Fast Fourier Transform (FFT) calculations, which have been used in previous studies to detect atmospheric anomalies on Mars [4]–[6]. This computation model is chosen due to its intensive nature and critical role in real-time astrophysical data analysis. By examining the impact of neutron radiation on the CPU and GPU embedded in the Jetson Nano, this work contributes to the development of more resilient systems capable of operating effectively in radiation-harsh environments.

Previous studies [7], [8] have shown that radiation affects the GPU in some cases, while in others, only the CPU is impacted. In this work, we aim to evaluate whether both components, in conjunction with OpenFaaS technology, are influenced by radiation.

A prior study was conducted to review previous research on the impact of radiation [8] on GPUs [9]–[12] and CPUs [12], [13]. The study in [8] examines radiation effects on Static Random Access Memory (SRAM)-based processors, measuring the impact of 15.6-MeV protons and 14.6-MeV neutrons at ultralow bias voltage.

Additionally, the study in [9] investigates the impact of radiation on GPUs by deploying algorithms on different GPUs and measuring SDCs. Another experimental study [10] focuses on measuring the effects of radiation on an NVIDIA Jetson Nano. Furthermore, this study [11] explores the effects of ionization on GPU units, specifically measuring Single Event Transients (SETs) and Single Event Latchups (SELs).

Regarding CPUs, we also found research on the effects of ionization [12], where proton testing was performed on an AMD Ryzen 3200G processor to assess degradation in its cores and GPU units.

In conclusion, previous research has evaluated a Jetson Nano Developer Kit exposed to radiation. This study also reported a series of Silent Data Corruption (SDC) errors and Detected Unrecoverable Errors (DUEs) caused by radiation exposure.

On the other hand, previous studies related to serverless computing [4]–[6], [14] demonstrate various use cases.

The first study [4] presents an algorithm deployed in a serverless environment for detecting atmospheric anomalies on Mars. In our case, we will use a similar approach by deploying an algorithm on a worker node to detect auroras under radiation conditions.

Secondly, the study [5] explores the use of an algorithm for processing ionograms via cloud computing. The architecture used in this case is similar to that of the previous study [4] but it is oriented to a different field of research, with a large amount of data and which is a challenge for the serverless architecture.

The next study [6] introduces an initial approach to using serverless computing for processing large volumes of data. This large amount of data is related to the surface of Mars and is also processed by serverless technology.

Finally, the last study [14] examines the deployment of various functions in an edge-computing environment, which focuses on bringing peripheral servers closer together to enhance processing efficiency.

This paper presents a study on the SEE vulnerability of a serverless architecture running on the CPU and GPU embedded in an NVIDIA Jetson Nano under 14-MeV neutron expo-

David Pacios, M. Rezaei, J. L. Vázquez-Poletti, Sara Ignacio-Cerrato and J. A. Clemente are with the Computer Architecture Department, Facultad de Informática, Universidad Complutense de Madrid (UCM), E-28040 Madrid, Spain, e-mail: {dpacios, mrezaei, jlvazquez, signacio, juanancl}@ucm.es.

This work was supported in part by the Spanish MINECO project PID2020-112916GB-I00, "RESHYLIENCE". This work has also received funding from the European Union's 2020 research and innovation programme under grant agreement No 101008126, corresponding to the RADNEXT project.

979-8-3315-5914-4/25/$31.00 ©2025 IEEE

sure. While under radiation, the device executed Fourier transform algorithm designed for detecting atmospheric anomalies on Mars [4]. This algorithm was deployed in a serverless computing architecture designed to mimic AWS Lambda functions [4]–[6], [15] for parallel processing tasks.

II. Experimental Setup

The NVIDIA Jetson Nano used in the experiment was modified to allow direct access to the GPU and CPU by removing part of its hardware. Additionally, the device was adapted to run the operating system from an external USB drive connected via a long cable to prevent irradiation from the neutron source. Several testing environments were set up on this drive.

A daemon prepared the operating system upon detecting an electric current, automatically powering on the device, logging in, and initiating a script. This script first performed a test on the CPU, followed by an analysis on the GPU. The analysis lasted 3 minutes, with readouts taken every 15 seconds to collect data from both the CPU and GPU. This data included information on the Fourier transforms, the number of green pixels, GPU and CPU performance, and the processed images.

The experiments were conducted in February 2024 using the Frascati Neutron Generator (FNG) at the ENEA-Frascati research center, where samples were exposed to a total fluence of 4×10^7 neutrons/cm^2 generated by standard D-T reactions. Funding for access to the ENEA facilities was provided by the RADNEXT project [8].

III. Experiment Application: Ionogram Analysis and Interpretation using MARSIS AIS Mode

A. MARSIS AIS

The Mars Advanced Radar for Subsurface and Ionosphere Sounding (MARSIS) is a radar operated by the European Space Agency (ESA) [4]. The radar operates in two distinct modes: one that focuses on probing the surface and subsurface using selectable frequencies of 1.8, 3, 4, and 5 MHz, and the other, Active Ionospheric Sounding (AIS), which involves transmitting a sweep of 250 frequencies ranging from 0.1 to 5.5 MHz. In AIS mode, the radar investigates the ionosphere from the spacecraft's position to its maximum ionized region, generating ionograms. These ionograms plot the echo reflections of each frequency against the signal's time of flight and frequency. An example of an ionogram is shown in Figure 1.

Ionograms provide extensive details about the ionosphere, including its primary trace formed by reflections at the plasma frequency. This nadir reflection allows for the retrieval of electron density profiles from the spacecraft's altitude to the maximum ionization region (typically around 130 km at the subsolar point). Additionally, AIS can determine local plasma electron density, magnetic field strength, surface reflections for frequencies above the plasma frequency, and occasionally oblique echoes from non-nadir reflections.

Oblique echoes occur when AIS probes areas above or near crustal magnetic fields, where the ionosphere deviates from spherical symmetry, particularly in Mars's southern hemisphere. These regions exhibit locally lifted ionospheric layers due to the influence of crustal magnetic fields. Oblique echoes [4], primarily associated with vertical crustal magnetic fields, are not consistently present in all ionograms. Their study may provide insights into ionospheric nonuniformity and the effects of crustal magnetic fields on plasma distribution.

B. Fast Fourier Transform

To determine the location of the oblique echo in frequency space, the FFT function will be used to isolate it in the ionograms. This process will transform each point in the image into frequency space and apply filtering. Specifically, frequencies that do not correspond to the green ones, but instead match those of the oblique echo, will be filtered out.

C. Ionogram

Ionograms are a representation of the Martian atmosphere in frequency space. The reflected signal is represented using color coding (green and blue). There is a time delay between the signal's transmission and detection. This delay is represented on the vertical axis, while the frequency of the signal is shown on the horizontal axis.

To determine the location of the oblique echo, the FFT function is applied to each image. First, the axis portion is removed from the ionograms. Then, the functions are applied to remove the irrelevant frequencies, and finally, the image with the oblique echo (green zone) is obtained. Figure 1 shows the result of this proccesing.

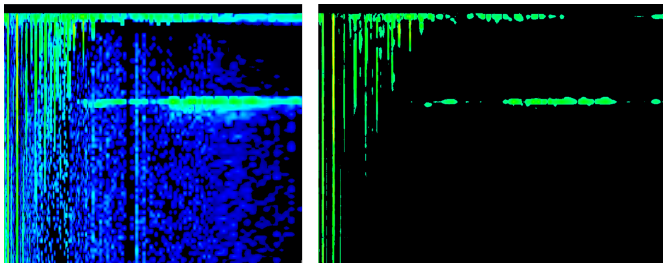

Fig. 1. Ionogram before (left) and after (right) the Fourier transform was applied

IV. Serverless-based system and application deployment

To deploy an architecture mimicking serverless, the FFT functions were deployed on OpenFaaS [14], [16]. This platform is used to streamline multiple functions, allowing for on-demand scaling and managing the entire process. The application is based on a command-line tool, which simplifies programming in multiple languages and frameworks. In this case, the workers are responsible for executing the code in isolation.

For this purpose, a daemon was implemented to handle image processing using the FFT functions described in the previous section. This system was deployed on both the CPU and GPU of the Jetson Nano.

The workers [17] performed the operations for applying the FFT functions. These workers were organized in an isolated

structure to manage computing resources. This structure mimics a serverless [18] distribution.

The image processing was performed as follows: First, the daemon reads the image path, and then the images from that path (20 images, as the one shown in Figure 1) are processed on the target CPU or GPU. The time to perform the Fourier transform on each image to detect the green color was 3 minutes. Running the algorithm generates an output file with the number of green pixels counted during the process, along with the CPU and GPU usage in percentage. The detected green color corresponds to the oblique echo on the Martian surface. Once this process is complete, the detected echo in the ionogram is obtained, as it was shown in Figure 1. The functions implemented by the daemon are executed by each one of the workers. The experiment mimics the use of workers to execute the code as AWS Lambda would, which is designed to scale automatically and handle increases in demand by instantiating multiple parallel executions of the function.

The FFT function was applied to the MARSIS instrument data, with all images processed on both the CPU and GPU. These processes were performed independently, as each time it was determined whether the processing should be run on the CPU or GPU. After the FFT processing, an image like the one shown at the right side of Figure 1 was obtained.

In our local framework, OpenFaaS was used to execute workers, allowing each FFT function to be executed autonomously. This setup enabled parallel execution, optimizing resource utilization across workers and enhancing the processing of the images. The automatic scaling mechanism responds to different workload fluctuations.

When emitting radiation in the instrument, the effects on the input and output images, as well as their processing, will be analyzed. There is a diverse body of literature [8]–[13] on how radiation can affect GPUs and CPUs. Next section will present the results obtained from performing these operations on the CPU and GPU.

V. Results and Discussion

Performance metrics such as CPU and GPU utilization, execution times, and memory usage were recorded to assess the impact of radiation on the system. Additionally, radiation-induced events, such as SDCs DUEs, were also recorded.

A. CPU and RAM Usage before and after radiation

Figure 2 compares performance measurements of the system in a normal environment with the ones under the radiation. The X-axis represents time in minutes, with the test lasting 3 minutes, and the Y-axis shows the usage in percentage. The measurement interval for CPU and GPU usage was 0.25 minutes. Each time the usage is measured, it is marked with an "×" — blue for CPU and red for RAM. CPU usage fluctuates between 10% and 60% in the majority of the time. However, there is a peak at around minute 2.2 where the CPU experiences up to 100% of usage. In contrast, RAM usage remains relatively stable, clustering around 60%, suggesting consistent memory consumption. As a result, DUEs occurred and were observed between 2 and 2.5 minutes. It worth mentioning that after the irradiation the USB drive was not functioning, hence, it was not possible to obtain any metrics neither from the CPU nor the GPU.

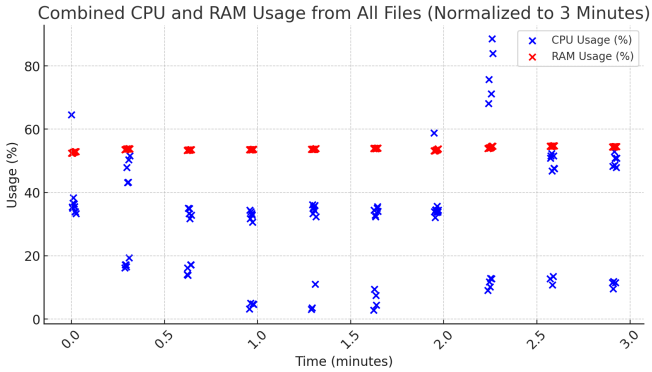

Fig. 2. System performance prior to and following neutron radiation exposure, depicting a decrease in memory usage and processing power

Figure 3 shows a comparison between CPU and RAM usage under neutron radiation, normalized to 3 minutes. RAM usage remains relatively constant at around 40% over time. On the contrary, CPU usage shows greater variability, fluctuating between 10% and 50%, with intervals where it clusters around 30%.

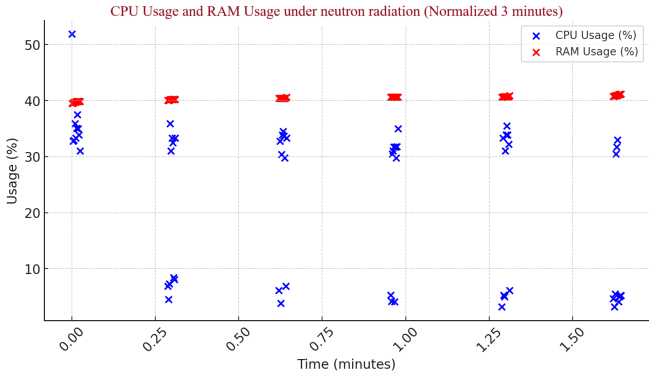

Fig. 3. Output consistency of Fourier transform calculations pre- and post-neutron exposure (this system was subjected to a total fluence of $4 \times 10^7 \text{neutrons/cm}^2$), indicating computational resilience

B. GPU Usage before and after radiation

Figure 4 compares GPU memory usage on normal environment with GPU memory usage under neutron radiation over time. The X-axis represents time in minutes, while the Y-axis indicates GPU memory usage as a percentage. Before irradiation (blue points), GPU memory usage gradually increases over time, reaching nearly 100% around 1.75 minutes. The memory usage under radiation also increases over time, but at a different rate and with a less uniform distribution of points. After the reduction in GPU usage, SDCs were detected.

As mentioned before no data could be extracted after the radiation due to a hard error in the USB drive.

Table I shows the DUEs identified from the data in Figure 2. These errors were detected at 6 points (blue points related to

CPU usage) starting from the second minute. From this point onward, the FFT was not executed, and DUEs were found. A total of 7 points were identified, with CPU usage starting at around 60.5%. On the other hand, SDCs were observed from 1.25 minutes. Four points were found, with GPU usage at about 60%.

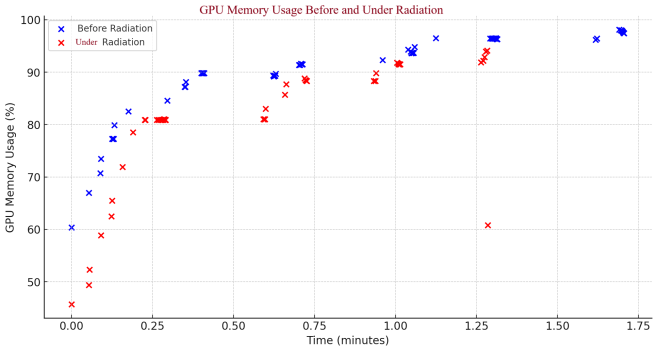

Fig. 4. GPU memory usage before and under neutron radiation (this system was subjected to a total fluence of $4 \times 10^7 \text{neutrons/cm}^2$)

TABLE I
DUEs observed in the experiments, with the percentage of using CPU; and SDCs observed, with the percentage of using GPU

Minute	DUE	CPU Usage (%)	Minute	SDC	GPU Usage (%)
2	1	60	1.25	1	60
2.25	2	71.2	1.25	2	90.5
2.25	3	73	1.25	3	91.2
2.25	4	78	1.25	4	93.9
2.25	5	80.3	- -	- -	- -
2.25	6	85	- -	- -	- -

As shown in Figures 2, 3 and 4, there is a significant difference in the impact of radiation on CPU and GPU usage. In particular, GPU usage increases substantially under radiation exposure, as confirmed by the SDC and DUE statistics in Table I.

C. Discussion

The experiment highlights the complex interplay between hardware vulnerabilities and software resilience in a serverless computing architecture exposed to high levels of neutron radiation. We have tested the NVIDIA Jetson Nano under 14-MeV neutrons by removing its hardware. It was operated with a script emulating the functionality of an AWS-Lambda function. The use of AWS Lambda-like functions for managing parallel processing tasks effectively demonstrates the feasibility of deploying serverless computing technologies in radiation-rich environments, provided that adequate safeguards are in place to mitigate the adverse effects on hardware.

VI. Conclusions

This contribution presents metrics to understand how radiation influences the CPU and GPU. A serverless architecture was deployed through the workers' actions, and their performance was evaluated when exposed to radiation. High resilience of both the CPU and GPU was observed throughout all the tests. These results were seen in the form of SDCs and DUEs that occurred throughout the tests.

References

[1] K. Bartlett *et al.*, "Proton Irradiation Damage and Annealing Effects in ON Semiconductor J-Series Silicon Photomultipliers," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 969, p. 163957, 2020.

[2] S. Fiore *et al.*, *Radiation and Magnetic Field Effects on New Semiconductor Power Devices for HL-LHC Experiments*, pp. 664–668.

[3] D. Agiakatsikas *et al.*, "Single Event Effects Assessment of UltraScale+ MPSoC Systems under Atmospheric Radiation," *IEEE Transactions on Reliability*, vol. 73, no. 1, pp. 771–783, 2024.

[4] D. Pacios *et al.*, "Serverless Architecture for Data Processing and Detecting Anomalies with the Mars Express MARSIS Instrument," *The Astronomical Journal*, vol. 166, no. 19, 2023.

[5] J. L. Vázquez-Poletti *et al.*, "Serverless On-Demand Marsis Ionogram Processing on a Public Cloud Computing Infrastructure," in *The Ninth Moscow Solar System Symposium 9M-S3*, pp. 233–235, 2018.

[6] J. L. Vázquez-Poletti *et al.*, "Advances in Cloud Computing for Mars Data Processing," in *The Eleventh Moscow Solar System Symposium 11M-S3*, pp. 87–88, 2020.

[7] D. Morgan *et al.*, "The processing of electron density profiles from the Mars Express MARSIS topside sounder," *Radio Science*, vol. 48, no. 3, pp. 197–207, 2013.

[8] R. García-Alía *et al.*, "Heavy Ion Energy Deposition and SEE Intercomparison Within the RADNEXT Irradiation Facility Network," *IEEE Trans. Nucl. Sci.*, vol. 70, no. 8, pp. 1596–1605, 2023.

[9] D. A. de Oliviera *et al.*, "Evaluation and mitigation of radiation-induced soft errors in graphics processing units," *IEEE Transactions on Computers*, vol. 65, no. 3, pp. 791–804, 2015.

[10] W. S. Slater *et al.*, "Total ionizing dose radiation testing of NVIDIA Jetson nano GPUs," in *2020 IEEE High Performance Extreme Computing Conference (HPEC)*, pp. 639–641, IEEE, 2020.

[11] P. Rech *et al.*, "Neutron radiation test of graphic processing units," in *2012 IEEE 18th International On-Line Testing Symposium (IOLTS)*, pp. 55–60, IEEE, 2012.

[12] J. L. Taggart *et al.*, "Total ionizing dose and proton single event effects in AMD Ryzen processor fabricated in a 12-nm bulk FinFET process," in *2023 IEEE Radiation Effects Data Workshop (REDW)(in conjunction with 2023 NSREC)*, pp. 6–11, IEEE, 2023.

[13] J. M. Badia *et al.*, "Analysing the radiation reliability, performance and energy consumption of low-power SoC through heterogeneous parallelism," *Sustainable Computing: Informatics and Systems*, vol. 44, p. 101049, 2024.

[14] G. Russo *et al.*, "Serverless functions in the cloud-edge continuum: Challenges and opportunities," in *2023 31st Euromicro International Conference on Parallel, Distributed and Network-Based Processing (PDP)*, pp. 321–328, IEEE, 2023.

[15] M. Villamizar *et al.*, "Infrastructure cost comparison of running web applications in the cloud using AWS lambda and monolithic and microservice architectures," in *2016 16th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid)*, pp. 179–182, IEEE, 2016.

[16] Yang *et al.*, "Flame: A Centralized Cache Controller for Serverless Computing," in *Proceedings of the 28th ACM International Conference on Architectural Support for Programming Languages and Operating Systems*, vol. 4, pp. 153–168, 2023.

[17] D. Balla *et al.*, "Adaptive scaling of Kubernetes pods," in *NOMS 2020-2020 IEEE/IFIP Network Operations and Management Symposium*, pp. 786–790, IEEE, 2020.

[18] M. Malawski *et al.*, "Serverless execution of scientific workflows: Experiments with Hyperflow, AWS Lambda and Google cloud functions," *Future Generation Computer Systems*, vol. 110, pp. 502–514, 2020.