When One Fault Breaks the SNN: the Critical Role of the Spike Routing NoC

Giuseppe Chessa [(1)], Michele Favalli [(2)] and Davide Bertozzi [(1)]

[(1)] Dept. of Computer Science, University of Manchester (UK), [(2)] Engineering Department, University of Ferrara (Italy)

As Neuromorphic Computing pushes toward mission-critical edge deployments, the narrative of "intrinsic fault tolerance" must be rigorously challenged [1]. The robustness of SNNs is often showcased under controlled algorithmic perturbations, reflecting only a limited projection of a broader and more complex hardware fault space. Bridging the gap between algorithmic analysis and end-to-end reliability of real-world deployments requires explicitly considering the underlying neuromorphic hardware. Crucially, existing research has focused on faults in synaptic memories and neural cores [2], while largely sidestepping the digital communication fabric – a network-on-chip (NoC) - that orchestrates spike routing across parallel neuromorphic hardware platforms. This is not merely an oversight, but a fundamental gap in current understanding: in fact, given the large fan-in and fan-out of typical SNNs, even localized faults in the spike routing fabric can potentially propagate through the system, amplifying errors and ultimately compromising functionality at scale. Ignoring this layer is no longer defensible. This gap arises because efficient spike-routing NoCs aligned with the event-driven neuromorphic computing principles are inherently asynchronous [3]. As a result, they are not readily available to the broader research community and remain difficult to design in practice, due to the specialized expertise required and the reliance on non-standard tool flows [4].

This study is a first step toward bridging the gap. It introduces the first cross-layer reliability analysis framework that systematically links aging-induced permanent faults in the spike-routing fabric to their ultimate impact on SNN performance. Rather than treating robustness as an isolated algorithmic property, it is explicitly anchored to the integrity of modern asynchronous NoCs, which trade intrinsic robustness for significant gains in area and energy efficiency. This design space, though essential for scalable neuromorphic systems, exposes a critical and largely unexplored vulnerability.

The study focuses on the long-term degradation associated with the aging-induced onset of permanent faults during operation, which can silently degrade system behavior. To capture this phenomenon, we introduce the first cross-layer reliability analysis framework that bridges the dynamics of faulty NoC microarchitectural states in hardware with model-level learning performance. By tightly integrating a gate-level model of an asynchronous NoC, faithfully backannotating post-synthesis delays, with a mainstream SNN simulation framework (snnTorch), we enable, for the first time, end-to-end fault propagation analysis from spike routing to inference accuracy.

When considering recurrent spiking neural networks for a binary navigation task and for a keyword spotting task as case studies, collected statistics on individual sample inferences indicate that on average $\sim$ 10% of the spikes are impacted by single NoC faults, with peaks beyond 80%, and with task accuracy drops that become significant (up to -10% and -36%, respectively). Crucially, these magnitudes are comparable to those typically reported in prior work under much more abstract and disruptive fault assumptions, including neuron-level failures and synapse-level corruptions. In stark contrast, our results demonstrate that similar system-level impact can emerge from a single permanent stuck-at fault affecting a logic value in the whole spike-routing infrastructure. This observation fundamentally reshapes the perceived fault sensitivity hierarchy in neuromorphic systems: it suggests that the communication fabric can act as an equally critical, and in some cases more insidious, source of functional degradation through silent data corruption.

[1] C. D. Schuman *et al*., *International Joint Conference on Neural Networks (IJCNN)*, pp. 1-10, 2020.
[2] H. -G. Stratigopoulos et al., *IEEE International Symposium on Defect and Fault Tolerance in VLSI and Nanotechnology Systems (DFT)*, pp. 1-8, 2023.
[3] Z. Su *et al*., *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 14, no. 3, pp. 409-424, 2024.
[4] S. Ataei *et al*., *IEEE Design & Test*, vol. 38, no. 2, pp. 27-37, 2021.