Self-Repairing Hardware with Astrocyte-Neuron Networks

Junxiu Liu, Jim Harkin, Liam Maguire, Liam McDaid, John Wade, Malachy McElholm
School of Computing and Intelligent Systems,
University of Ulster, Magee Campus,
Derry, Northern Ireland, UK, BT48 7JL
{j.liu1, jg.harkin, lp.maguire, lj.mcdaid, wade.j, m.mcelholm}@ulster.ac.uk

***Abstract*—A Self-rePAiring spiking Neural NEtwoRk (SPANNER) hardware architecture is presented in this paper. It is based on a software model of an astrocyte-neuron network which previously demonstrated the ability to self-detect faults and self-repair autonomously. Experimental results in this paper show that when faults occur at the synapse, remaining healthy synapses of the same neuron are enhanced by the feedback from the astrocyte, which enables system functionality to be maintained. This is the first time that astrocyte cells merged within spiking neurons demonstrate a self-repairing capability in hardware. This repair capability achieves a much more fine-grained repair capability in hardware compared to the conventional fault tolerance techniques.**

Keywords—astrocyte; SNN; repair; fault tolerance; FPGA

I. INTRODUCTION

Reliability has traditionally been a design challenge in mission critical electronics, however this challenge is now migrating into everyday non-critical systems where engineers must aim to design reliable systems on unreliable fabrics. Traditional approaches to fault-tolerant computing incorporate redundancy/replication models, error correction techniques and radiation hardening. However, such approaches only provide limited levels of reliability as inherent architectural constraints are placed on the number and type of faults that can be tolerated and the level of granularity with which repairs can be implemented. More recent approaches to improving system reliability have looked to nature for inspiration in exploiting "self-x" properties including self-repair and self-organisation [1]. The reasonable success to date of these bio-inspired 'evolutionary' approaches over conventional approaches underpins the belief that to fully realise fault-tolerant and self-repairing capabilities, future systems will need to harness similar mechanisms that are found in nature. Recent publications have highlighted that astrocytes continually exchange information with multiple synapses and consequently play a crucial role in brain re-wiring by regulating synaptic formation/elimination, synaptic morphology and structural plasticity [2]. The authors have developed a computational model which captures this behaviour [3], [4] and have demonstrated how astrocytes cells merged within spiking neurons can perform distributed and fine grained self-repair under the presence of faults. Progress has been made in implementing astrocyte cells in hardware where the aim is to explore Spiking Neural Network (SNN) behaviour [5]. In this paper, we go further and propose a hardware architecture of SPANNER which emulates the self-repairing SNN [3], and demonstrates how significant levels of fine-grained repairs can be achieved in real hardware. This is the first time that astrocytes cells merged within spiking neurons have demonstrated a self-repairing capability in hardware. Section II outlines the authors' current self-repairing SNN software model. Section III introduces the hardware architecture of the self-repairing SNN and section IV reports on results demonstrating the repair capability. Section V provides a summary and conclusion.

II. SPIKING ASTROCYTE-NEURON NETWORKS

This section provides a brief background on how astrocytes interact with spiking neurons. Astrocytes enwrap many synapses which are connected to a neuron, and it is this process by which bi-directional communication is performed. This type of synapse is known as a tripartite synapse. In a tripartite synapse, the arrival of an action potential axon, releases glutamate across the cleft and binds to receptors on the post-synaptic dendrite causing a depolarization of the post-synaptic neuron; then voltage gated channels on the dendrite allow the influx of calcium (Ca^{2+}) into the dendrite causing endocannabinoids to be synthesized and subsequently released from the dendrite. The 2-arachidonyl glycerol (2-AG), a type of endocannabinoid and retrograde messenger, is known to feed back to the pre-synaptic terminal in two ways: (a) *Directly*. 2-AG binds directly to type 1 Cannabinoid Receptors (CB1Rs) on the pre-synaptic terminal. This results in a decrease in transmission probability (PR) and is termed Depolarization-induced Suppression of Excitation (DSE); and (b) *Indirectly*. 2-AG binds to CB1Rs on an astrocyte which enwraps the synapse increasing IP_3 levels within the astrocyte and triggering the intracellular release of Ca^{2+}. This results in the astrocytic release of glutamate which binds to pre-synaptic group I metabotropic Glutamate Receptors (mGluRs). Such signalling results in an increase of synaptic transmission PR termed e-SP. In the authors' works [3], [4], the neurons are modelled using the Leaky Integrate-and-Fire (LIF) model where the firing threshold voltage is 9mv; and the synapses are probabilistic-based. The associated PR of each synapse is affected by the DSE and e-SP (1):

$$PR(t) = (\frac{PR(t_0)}{100} \times DSE(t)) + (\frac{PR(t_0)}{100} \times eSP(t)) \qquad (1)$$

where $PR(t_0)$ is the initial PR. The signal exchange and chemical transmission in the tripartite synapse are briefly described as follows: when a post synaptic neuron fires, 2-AG is released and described in (2) by:

$$\frac{d(AG)}{dt} = \frac{-AG}{\tau_{AG}} + r_{AG}\delta(t - t_{sp}) \quad (2)$$

where AG is the quantity of the released 2-AG; τ_{AG} and r_{AG} is the decay and production rate of 2-AG, respectively; t_{sp} is the time of the post-synaptic spike. When the 2-AG binds to CB1Rs on the astrocyte, IP_3 is generated which is dependent on the amount of released 2-AG and is given by (3).

$$\frac{d(IP_3)}{dt} = \frac{IP_3^* - IP_3}{\tau_{ip_3}} + r_{ip_3}AG \quad (3)$$

where IP_3 is the quantity within the cytoplasm, IP_3^* is the baseline of IP_3 when the cell is in a steady state and receiving no input, τ_{ip_3} and r_{ip_3} is the decay and production rate of IP_3. The Ca^{2+} dynamics within the cell is described by (4).

$$\frac{d(Ca^{2+})}{dt} = J_{chan}(Ca^{2+}, h, IP_3) + J_{leak}(Ca^{2+}) - J_{pump}(Ca^{2+}) \quad (4)$$

where J_{chan} is the IP_3 and Ca^{2+} -dependent Ca^{2+} release, J_{pump} is the amount of Ca^{2+} pumped from the cytoplasm into the Endoplasmic Reticulum (ER), J_{leak} is the Ca^{2+} which leaks out of the ER. A linear correspondence is assumed between the released 2-AG and the DSE, and given by:

$$DSE = AG \times K_{AG} \quad (5)$$

where AG is the amount of released 2-AG, K_{AG} is a scaling factor used to convert the level of 2-AG into the desired negative range. The intracellular astrocytic calcium dynamic are used to regulate the release of glutamate from the astrocyte which drives e-SP. The quantity of released glutamate targeting group I mGluRs is given by:

$$\frac{d(Glu)}{dt} = \frac{-Glu}{\tau_{Glu}} + r_{Glu}\delta(t - t_{Ca}) \quad (6)$$

where Glu is the quantity of glutamate, τ_{Glu} and r_{Glu} are the decay and production rate of glutamate respectively, and t_{Ca} is the time of the Ca^{2+} threshold crossing. The level of e-SP is dependent on the quantity of glutamate, which is modelled by:

$$\tau_{eSP}\frac{d(eSP)}{dt} = -eSP + m_{eSP}Glu(t) \quad (7)$$

where τ_{eSP} is the decay rate of e-SP, m_{eSP} is a weighting constant. The signal pathways of DSE in (5) and e-SP in (7) modulate the PR of the synapse as shown by (1). More details on the self-repair mechanism in the SPANNER can be found in our previous works [3], [4].

III. Self-Repair Architecture of SPANNER

This section presents the SPANNER architecture, focusing on the concept and implementation. Fig. 1 presents the concept of SPANNER which shows the self-repairing mechanism, e.g. astrocyte, A, enwrapping the synapses connected to neurons N1 and N2. Both the direct and indirect feedback signals compete at each synapse to alter PR and enable a stable state to be reached. If a fault occurs at the synapse associated with a specific neuron, both the indirect/direct feedback for that synapse cease. This creates an imbalance in PR at the synapses with the result that the PR of healthy synapses increase to restore the functionality. This is due to indirect retrograde feedback from other neurons. This is the principle of self-repair which allows detection of fault(s) in synapses and the repair of healthy synapses via increasing their PR. This process is regulated by the astrocyte and previous experiments demonstrate the capability [3-4]. Fig. 2 describes the proposed hardware architecture of SPANNER which consists of two neurons and one astrocyte, as depicted in Fig. 1. This small network is used as a demonstrator of the repair principle in hardware and its software model [3] provides a benchmark. Each neuron is associated with several synapses. A core component in SPANNER is the *neuron facility* (e.g. Neuron #1, #2 in Fig. 2). Its inputs are the signals from the synapses and output signals include neuron spike, 2-AG and DSE. After the *neuron facility* receives inputs from the synapses, it calculates the membrane potential, and compares it with a firing threshold (v_{th}). If the former is smaller than the latter, no spike is output; otherwise the neuron fires a spike. After firing, the membrane potential is clamped at 0V for 2ms thereby implementing the refractory period. For the *neuron facility*, the generation of membrane potential, 2-AG and DSE are the core processes, which calculate and update values of the membrane potential, and also the released quantity of 2-AG and DSE in real time, see (2) and (5) in Section II. The released 2-AG binds to the CB1Rs on the astrocyte process which results in the *indirect* increase of synaptic transmission PR. However, the DSE has a direct effect on the pre-synaptic terminal, i.e. decreasing the synaptic transmission PR.

The second key component is the *astrocyte facility*. As shown in Fig. 2, the *astrocyte facility* receives the 2-AG from the post-synaptic neuron, and generates e-SP for the synapses which can increase the synaptic transmission PR. The internal working mechanism of the *astrocyte facility* is described as follows and Equations (3), (4), (6) and (7) are implemented in this facility: after receiving the 2-AG from the post-synaptic neurons, the IP_3 is generated and updated. As outlined in Section II, the release of IP_3 binds to inosotil 1, 4, 5-trisphosphate Receptors (IP_3R_s) on the ER, which allows the release of Ca^{2+} from the ER into the cytoplasm. The channels of J_{chan} , J_{leak} and J_{pump} are used to describe the Ca^{2+} dynamics within the cell. For the signal processing in these channels, the generation of Ca^{2+} is the core function; and the release amount of Ca^{2+} is calculated by (4). A Ca^{2+} register is used to store and update the Ca^{2+} value. The change of Ca^{2+}

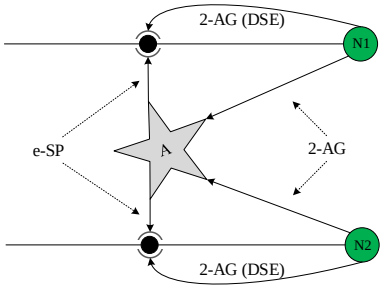

Fig. 1. Self-repair mechanism of SPANNER.

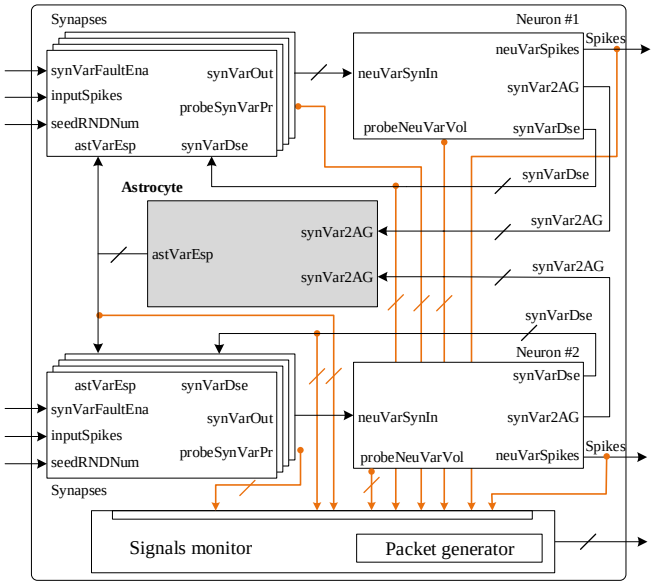

Fig. 2. SPANNER hardware architecture.

leads to the intracellular astrocytic calcium dynamics which regulate the release of glutamate from the astrocyte and drives the e-SP. Inside this facility, a Ca^{2+} spike generator is used to model the calcium dynamics; a glutamate generator and a register are used to calculate and update the quantity of glutamate of (6); an e-SP generator and an e-SP register are used to calculate and update the quantity of e-SP release of (7). Therefore, in the final step the e-SP is calculated as the output signal of the *astrocyte facility*. The output signals of e-SP/DSE in the astrocyte/neuron facilities are connected to each *synapse facility* and have different effects on the pre-synaptic terminals, e.g. increasing and decreasing the synaptic transmission PR, which are discussed in the synapse facility. Note that the *astrocyte facility* in Fig. 2 only connects to two neurons and several synapses, however the astrocyte can enwrap up to $\sim 10^5$ synapses and contact ~6 neurons within the cortex and hippocampus [6]; therefore although the astrocyte model is complex, the *astrocyte facility* does not prohibit the scalability of astrocyte-neuron systems, as only few *astrocyte facilities* are required in a large-scale SNN system.

The *synapse facility* implements a probabilistic-based synapse model and has several inputs, e.g. the pre-synaptic spikes, the DSE/e-SP from neuron/astrocyte facilities, and a seed signal for the random number generator. In addition, fault injection enable signals are used to inject the faults into the synapses, e.g. reduced synapse PR. The *synapse facility* has one output (synVarOut) which is connected to the neuron facility. Each synapse has an initial value of transmission PR (e.g. 0.5 in this approach) for every time step. However, when the fault injection signal is enabled, the PR initial value is set to be very low (e.g. 0.1), which is stored in the PR register. The output value of PR register goes to a comparator which compares the PR value with a random number; if the former is larger, the synapse has a valid output; otherwise the synapse doesn't allow a spike output. In this process, a pseudorandom number generator is employed to generate the random number when the input pre-synaptic spike is presented. The PR value can be set low by the fault injector component, e.g. mimic

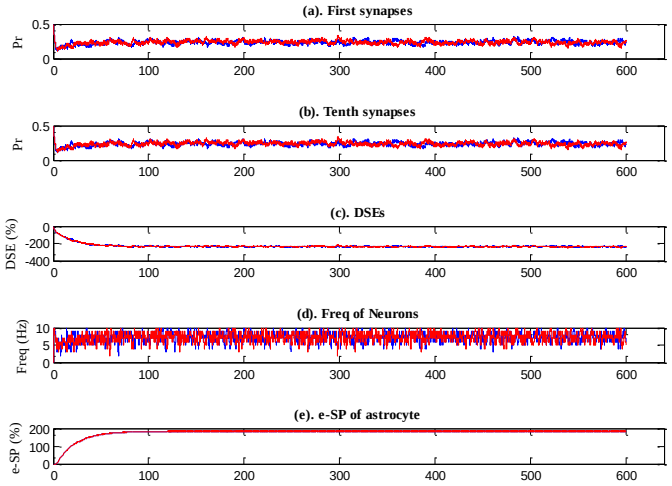

Fig. 3. Neuron #1 (blue) and #2 (red) in the SPANNER under no fault.

hardware fault. In SPANNER, the synaptic transmission PR is regulated by the DSE and e-SP; only the fault injector can override this. In the synapse facility, two PR adjustors (i.e. one is for direct signalling from DSE, e.g. decreasing action, the other is for indirect signalling from e-SP, e.g. increasing action) are employed to adjust the PR value of a synapse. The outputs of these two adjustors are connected to the PR generator and calculate the change of PR (i.e. ΔPR) which is output to the PR register for updating. As discussed in Section II, the indirect signalling pathway is the catalyst for self-repair of damaged or low PR synapses. A range of experiments are presented in the next section which demonstrates the principle of the self-repairing mechanism in hardware.

IV. Experimental Results

This section presents the test bench setup and results of experiments on the proposed SPANNER hardware architecture. The example SPANNER architecture shown in Fig. 2 was implemented on FPGA hardware which includes two neurons and one astrocyte; where each neuron has ten synapses. All synapses have an initial transmission PR of 0.5 and receive an input spike train with average frequencies of 10Hz. A monitoring framework in [7] is employed which can: probe the SPANNER signals, collect the data from the hardware in real-time, and upload the data to a PC for analysis. All FPGA probed signals are coloured brown in Fig. 2. The fault rate is used to denote the percentage of faulty synapses connected to a specific neuron, e.g. fault rate 80% means 8 out of 10 synapses connected to a neuron are faulty. The following experiments were performed in real-time on a Xilinx Virtex-7 XC7VX485T-2FFG1761C device, over a 600 second period. All results have been captured in real-time from the FPGA and the results are compared with the software model of [3].

In the first experiment, we consider when there are no faults, i.e. the fault rate is 0%. In this scenario, neuron #1 and #2 are firing as a result of the presented pre-synaptic stimuli and coupling with the astrocyte via the 2-AG signal. This causes the release of glutamate which acts on mGluRs receptors situated on the pre-synaptic terminals of both neurons. Fig. 3 shows the captured data related to neuron #1 and #2. The synaptic transmission PR values of the first and

tenth synapses of neuron #1 are presented. The PR values of the synapses have similar profiles, e.g. the average value is ~0.25. Neuron #2 related signals are similar to neuron #1. The e-SP is the same for both neurons as it is a global signal. The output frequencies of both neurons were recorded at ~7Hz which match the average firing rate of the model in software [3]. This confirms that the hardware, under non faulty conditions, correctly matches the software model in [3].

In the next experiment, we observe the dynamic behaviours of the SPANNER hardware when 8 synapses connected to neuron #2 are set to be faulty (80% fault rate). The faults were injected temporarily into the 8 synapses of neuron #2 between time 200 and 400 seconds. Fig. 4 provides the results from both the neurons. It can be seen that the PR value of one of the faulty synapses associated with neuron #2 (e.g. first synapse in red in Fig. 4(a)) is set to be very low during time 200-400 which mimics a temporary fault. When the fault disappears, the PRs return to the normal initial values. During the time period of the temporary fault, the healthy synapses (e.g. tenth synapse in red in Fig. 4(b)) are enhanced by the feedback from astrocyte facility to have a larger PR. This is the repair process which aims to maintain, as close as possible, to the target output frequency of 7Hz. The DSE of neuron #2 is also decreased during this time period which is an expected behaviour. During the fault period we can also see that the average firing rate has a slight decrease as SPANNER implements its repair, however, after the fault period (>400s) it returns to its target frequency. Note that neuron #1 (coloured blue in Fig. 4) is not affected as expected and the results are similar to Fig. 3. The results under temporary faults clearly show the dynamic behaviour of the SPANNER hardware, and most importantly, demonstrate the principle of the proposed self-repairing mechanism, i.e. the system can self-adapt and self-adjust to different fault conditions.

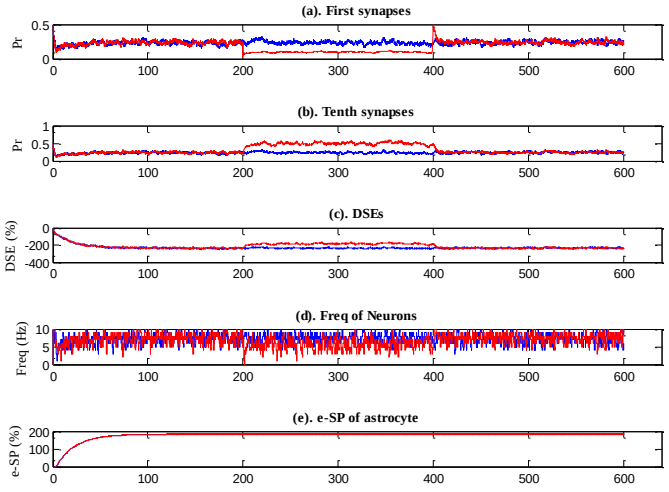

Fig. 4. Neuron #1 (blue) and #2 (red) under temporary 80% fault rate.

To verify the hardware accuracy, Table I outlines the average frequencies of both neurons from software [3] and FPGA implementation under different fault rates. The difference between the software simulation and hardware system is between 0.01Hz to 0.09Hz. As the synapse model is a probabilistic-based model, the results of different trials are slightly different; therefore the minor difference are acceptable. This verifies that the SPANNER hardware achieves the same results as the software simulation. In addition, the computing latency of FPGA device is less than 2,500 clock cycles.

TABLE I. AVERAGE FREQUENCIES OF NEURONS

Fault rate	Platform	Neuron 1	Neuron 2
0%	Simulation	7.19	7.20
	Hardware	7.28	7.27
80%	Simulation	7.38	5.68
	Hardware	7.37	5.75

The occupied resource area of SPANNER is summarized in Table II with the astrocyte, synapse and neuron facilities presented. The astrocyte facility occupies the largest resource since the astrocyte includes the complex equation models. The synapse and neuron facilities occupy less area. It should be noted that this paper focused on emulating the SPANNER self-repair capabilities in hardware and therefore optimisation of the hardware was not the goal of this work. Future work aims to explore this objective.

TABLE II. HARDWARE RESOURCE

	Astrocyte	Synapse & Neuron	DSE generator
Slice LUTs	11,394	9,865	4,353
Slice Registers	11,666	10,383	4,688
Slice	3,552	3,120	1,394
Block RAM Tile	5	0.5	0
DSPs	42	45	14

V. CONCLUSION

A novel SPANNER hardware architecture has been presented which emulates the self-repairing mechanism of the brain. Results demonstrate that the proposed work enables the capabilities of self-detection and repair to be implemented in the electronic hardware with fine-grained repairs achievable compared to conventional fault-tolerant techniques. Future works includes the optimized hardware design of large-scale SPANNER implementations, and the hardware large-scale astrocyte-neuron system design and analysis.

REFERENCES

[1] J. Liu, J. Harkin, M. McElholm, L. McDaid, A. Jimenez-Fernandez, and A. Linares-Barranco, "Case Study: Bio-inspired Self-adaptive Strategy for Spike-based PID Controller," in ISCAS, 2015, pp. 2700–2703.

[2] M. De Pittà, N. Brunel, and A. Volterra, "Astrocytes: Orchestrating Synaptic Plasticity?," Neuroscience, pp. 1–19, 2015.

[3] J. Wade, L. McDaid, J. Harkin, *et. al*, "Self-Repair in a Bidirectionally Coupled Astrocyte-Neuron (AN) System based on Retrograde Signaling," Frontiers in Computational Neuroscience, 6 (76), pp. 1–12, Jan. 2012.

[4] M. Naeem, L. J. McDaid, J. Harkin, *et. al*, "On the Role of Astroglial Syncytia in Self-Repairing Spiking Neural Networks," IEEE Transactions on Neural Networks and Learning Systems, pp. 1–11, 2015.

[5] M. Hayati, M. Nouri, S. Haghiri, and D. Abbott, "A Digital Realization of Astrocyte and Neural Glial Interactions," IEEE Transactions on Biomedical Circuits and Systems, pp. 1–12, 2015.

[6] M. M. Halassa, T. Fellin, H. Takano, J.-H. Dong, and P. G. Haydon, "Synaptic Islands Defined by the Territory of a Single Astrocyte," The Journal of Neuroscience, vol. 27, no. 24, pp. 6473–6477, 2007.

[7] J. Liu, J. Harkin, Y. Li, L. Maguire, and A. Linares-Barranco, "Low Overhead Monitor Mechanism for Fault-tolerant Analysis of NoC," in IEEE 8th International Symposium on Embedded Multicore/Many-core Systems-on-Chip, 2014, pp. 189–196.