NEUROMORPHIC COMPUTING FOR VERY LARGE TEST AND EVALUATION DATA ANALYSIS

MAY 2014

FINAL TECHNICAL REPORT

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

STINFO COPY

AIR FORCE RESEARCH LABORATORY
INFORMATION DIRECTORATE

■ AIR FORCE MATERIEL COMMAND ■ UNITED STATES AIR FORCE ■ ROME, NY 13441

NOTICE AND SIGNATURE PAGE

AFRL-RI-RS-TR-2014-112   HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

FOR THE DIRECTOR:

/ S /
JOSEPH A. CAROLI, Chief
High Performance Systems Branch

/ S /
MARK H. LINDERMAN
Technical Advisor, Computing &
Communications Division
Information Directorate

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)* MAY 2014	2. REPORT TYPE FINAL TECHNICAL REPORT	3. DATES COVERED *(From - To)* OCT 2010 – SEP 2013

4. TITLE AND SUBTITLE NEUROMORPHIC COMPUTING FOR VERY LARGE TEST AND EVALUATION DATA ANALYSIS	5a. CONTRACT NUMBER IN-HOUSE
	5b. GRANT NUMBER N/A
	5c. PROGRAM ELEMENT NUMBER 61102F
6. AUTHOR(S) Bryant Wysocki, Nathan McDonald, Clare Thiem, and Thomas Renz (AFRL) James Bohl (Rome Research Corporation)	5d. PROJECT NUMBER T1NM
	5e. TASK NUMBER IN
	5f. WORK UNIT NUMBER HO

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Air Force Research Laboratory/RITB 525 Brooks Road Rome, NY 13441-4505	8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Research Laboratory/RITB 525 Brooks Road Rome NY 13441-4505	10. SPONSOR/MONITOR'S ACRONYM(S) AFRL/RI
	11. SPONSOR/MONITOR'S REPORT NUMBER AFRL-RI-RS-TR-2014-112

12. DISTRIBUTION AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited. PA# 88ABW-2014-2062
Date Cleared: 5 MAY 2014

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The research conducted under this AFOSR effort can be broken down into two parallel tracks. The first track persued the development of self-reconfigurable neuromorphic computing architectures utilizing memristive nanotechnology. The second track involved the analysis and utilization of newly available hardware-based artificial neural network chips. These two aspects of the program are complementary. The neuromorphic architectures research focused on long term disruptive technologies with high risk but revolutionary potential. The hardware-based neural network research provided the means to apply many of the advantages of biologically inspired processing using advanced chipsets that are available today. Overall, hardware-based neural processing research allows us to study the fundamental system and architectural issues relevant for employing neuromorphic computing concepts towards the development of autonomous systems for perception and data analysis. The utilization of computationally intelligent processors for tasks within the United States Air Force (USAF) Developmental Test and Evaluation (T&E) community was emphasized, with a focus on test data analysis and process control efficiencies. This final report is organized to provide the reader with useful overview of the work.

15. SUBJECT TERMS

Neuromorphic Computing, Nanotechnology, Memristive Device, Artificial Neural Networks, Perception, Zero Instruction Set Computing

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON **CLARE D. THIEM**
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U	UU	54	19b. TELEPHONE NUMBER *(Include area code)* **N/A**

TABLE OF CONTENTS

LIST OF FIGURES

1.0 SUMMARY

The research conducted in this project can be broken down into two parallel tracks, development of self-reconfigurable neuromorphic computing architectures utilizing memristive nanotechnology, and analysis and utilization of newly available hardware-based artificial neural network chips. These two aspects of the program were complementary. The neuromorphic architectures research focused on long term disruptive technologies with high risk but revolutionary potential. The hardware-based neural network research provided the means to apply many of the advantages of biologically inspired processing using advanced chipsets that are available today.

Overall, hardware-based neural processing research allowed us to study the fundamental system and architectural issues relevant for employing neuromorphic computing concepts towards the development of autonomous systems for perception and data analysis. The utilization of computationally intelligent processors for tasks within the United States Air Force Developmental Test and Evaluation (T&E) community was emphasized, with a focus on test data analysis and process control efficiencies. This program resulted in the successful implementation of reconfigurable neuromorphic architecture building blocks. In addition, this work demonstrated the benefits of neuromorphic architectures for increased performance over conventional computing in certain applications, especially in Size, Weight and Power (SWaP) constrained environments.

2.0 INTRODUCTION

The research conducted in this project can be broken down into two parallel tracks. The first track pursued the development of self-reconfigurable neuromorphic computing architectures utilizing memristive nanotechnology with the goal to build the scalable neuromorphic computing concept to achieve intelligent functions. The second track involved the analysis and utilization of newly available hardware-based artificial neural network chips with the goal to develop a relationship with the Air Force test and evaluation community to identify applications on which to focus neuromorphic computing technology. These two aspects of the program were complementary. The neuromorphic architectures research focused on long term disruptive technologies with high risk but revolutionary potential. The hardware-based neural network research provided the means to apply many of the advantages of biologically inspired processing using advanced chipsets that are available today. Overall, hardware-based neural processing research allowed us to study the fundamental system and architectural issues relevant for employing neuromorphic computing concepts towards the development of autonomous systems for perception and data analysis. The utilization of computationally intelligent processors for tasks within the United States Air Force Developmental Test and Evaluation (T&E) community was emphasized, with a focus on test data analysis and process control efficiencies.

Reconfigurable XOR Logic Gate

The amazing computing power of the brain is in part due to its highly parallelized interconnectivity amongst neurons through synaptic connections. The synaptic connection plays an important role in brain activity as these connections can be strengthened or weakened as the brain learns and stores knowledge within the system. Neuron behavior has been characterized as an adding system that provides an output based on the sum of all inputs (through synaptic connections) and its connectivity to other neurons. In addition, the amount of information or knowledge a neuromorphic computer can retain depends on the number of neurons and synaptic connections within the system. For example, the brain of an ant is said to contain approximately 300,000 neurons. Given the large number of neurons and synapses required to design systems capable of mimicking practical brain functionality connections (approximately 1,000 synaptic connections per neuron), it is important for the electronic devices performing the mimicry to be small. This facilitates fabrication within a reasonable physical area, similar to how computer microprocessors are fabricated today. In Sections 3.2 and 4.2, a physical description is presented by which three transistor synaptic circuit functions mimic brain synaptic behavior. The implementation of synaptic circuitry within an adding node and single transistor adding neuron (a simple neuromorphic computing architecture) is also demonstrated. As an example to show the utility of a reconfigurable logic gate, a neuromorphic architecture as a reconfigurable XOR logic gate is demonstrated.

Computer simulations based on the physical principles described above demonstrate the feasibility of the neuromorphic architectures. Variations of neuromorphic computing architectures can be constructed employing transistors, memristors and inverter circuits coupled to floating node neurons or thresholding neurons to perform the desired computations.

Section 3 in this report describes the methods, assumptions and procedures used in this project for memristor device and circuit architecture design, to demonstrate neuromorphic processing. Section 4 provides results of the work in memristive device yield and functional testing and memristor based XOR circuit design. Characterizations of reconfigurable circuits and neural networks are given. Section 5 covers the conclusions about the project and potential future research topics. The sponsor, Air Force Office of Scientific Research (AFOSR), wanted a heavy emphasis on socializing the results of this project to achieve transition. As a result, more than typical emphasis is given to interactions and transition throughout this report. The Appendix describes the technology transitions achieved as a result of this project. Patents, journal papers and talks are listed.

3.0 METHODS, ASSUMPTIONS AND PROCEDURES

This section describes the methods, assumptions and procedures used in this project. Memristive device and circuit architecture design and test procedures are detailed. Work evaluating a commercial neuromorphic processor and its application to real world problems is described.

3.1 Memristive Device Design and Fabrication

The nonvolatility and reconfigurability of memristive devices make them useful candidates for electronic synaptic connections and as memory for the storage of neural network training signatures. Physical and compact models were developed in this project for MATLAB, HSPICE, and Verilog A environments and were used in circuit simulations allowing the development of prototype hardware of in-house designed reconfigurable neuromorphic architectures. Three material configurations of memristor device architectures were examined within this effort with subsequent developments in hardware utilization.

Chalcogenide Based Memristor Devices

Extremely productive collaborations were established at the onset of this program with Professor Kristy A. Campbell and her team at Boise State University (BSU). The groups worked closely to utilize and optimize BSU's chalcogenide based memristive devices for neuromorphic applications. It was observed that the few existing published memristive models did not accurately represent the electrical characteristics of Air Force Research Laboratory (AFRL) memristor device hardware. Therefore, a simple compact model was developed that accurately represents the electrical behavior of chalcogenide based memristors. Such models are required for large scale circuit design and simulations. The model enabled in-house architectural designs and patents.

The chalcogenide based memristor devices were fabricated on 200 mm p-type Si wafers. Isolated tungsten bottom electrodes were patterned on the wafers and a planarized nitride layer was used for device isolation. Vias were etched through the nitride layer to provide contact to the bottom electrode and to define the device active region. The memristor device structure consists of the layers (from bottom electrode contact side to top electrode contact): 300 Å Ge_2Se_3 / 500 Å Ag_2Se / 100 Å Ge_2Se_3 / 500 Å Ag / 100 Å Ge_2Se_3, as shown, alongside a typical IV curve, in Figure 1. Part (a) shows a cross section of the chalcogenide based memristor fabrication design. Part (b) shows a typical measured memristor device Lissajous I-V curve with characteristic behavior at 100Hz frequency. The top and bottom electrode bond pad contacts were 80 μm x 80 μm. A detailed examination of these devices and the compact model is provided in [2].
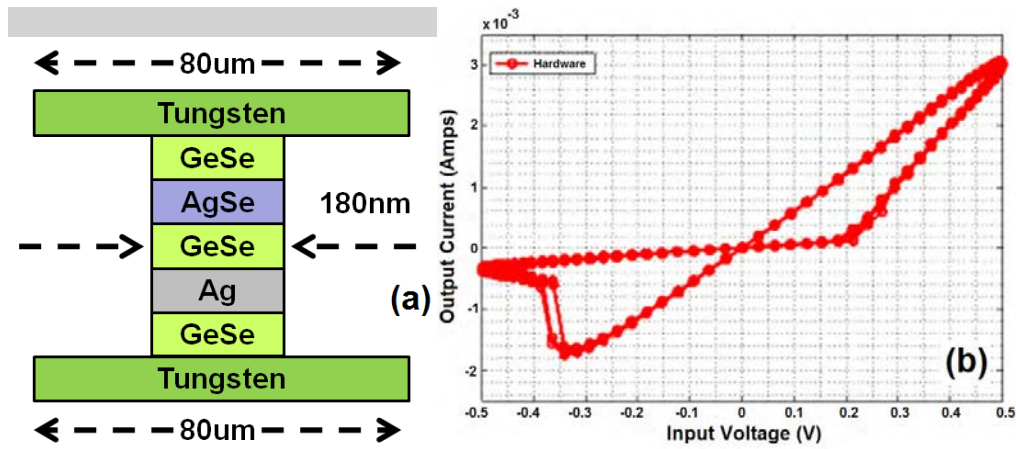
Figure 1. Chalcogenide Based Memristor

In-house Al/Cu$_x$O/Cu Memristive Devices

Memristive devices were developed and fabricated in an effort to better understand memristor behavior and to create memristors with engineered characteristics. This research covered the in-house design, fabrication, characterization and modeling of Al/Cu$_x$O/Cu memristive devices. The devices were created by depositing Al top electrodes atop a Cu$_x$O film grown on a Cu wafer using plasma oxidation. These devices were utilized in simple reconfigurable circuits designed to test physical implementations of neural architectures. Figure 2 shows the device layout and typical testing connections
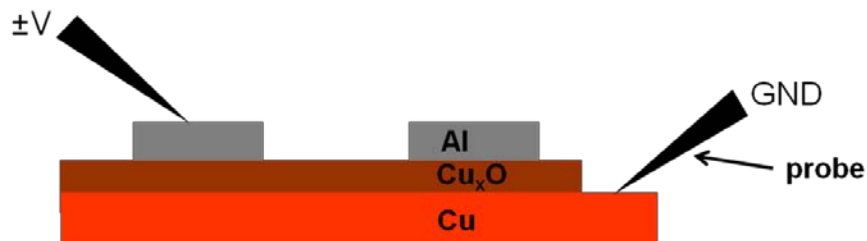
Figure 2. Device Design with Measurement Probe and Electrical Bias Arrangement

The completed memristive devices were electrically characterized using a Cascade M-150 probe station and an Agilent B1500A Semiconductor Device Analyzer. The exact testing parameters varied with the device generation; however, the method of testing remained

consistent. Devices were operated with the voltage bias applied to the Al top electrode and an exposed Cu bottom electrode which was electrically grounded. For both SET and RESET operations, either "single sweep" or "dual sweep" measurements were taken. In single sweep-mode, the applied voltage (V_{app}) sweep progressed from 0 V to $\pm V_{app}$ in ± 10 mV steps. In dual sweep mode, the applied voltage sweep progressed from 0 V to $\pm V_{app}$ then back to 0 V in ± 10 mV steps. The time between each measurement was not specifiable and was measured to be between 0.04-0.05 s. Lastly, during the SET operation only, a current compliance between |0.5 – 10| mA was enforced to prevent the device from shorting during the transition from a high resistance state to a low resistance state.

Physical and empirical models of the devices were created for MATLAB, HSPICE, and Verilog A environments. While the physical model proved of limited practical consequence, the robust empirical model allowed for rapid prototyping of Complimentary Metal-Oxide Semiconductor (CMOS) memristor circuitry. Details on the fabrication, testing, and modeling of these devices are available in [3, 4].

Hafnium Oxide Memristive Crossbars

Fourteen chips containing memristive Resistive Random Access Memory (ReRAM) devices in several different size crossbar architectures were acquired for use in advancing in-house research on neuromorphic computing. Funding supplied through this project was used to purchase these state-of-the-art devices by leveraging concurrent research within AFRL. The ReRAM device technology in crossbar architecture was developed under both the Air Force Research Laboratory Memristive Crossbar Nanoelectronics Architecture Program as well as the current Hybrid CMOS/Memristive Nanoelectronics for AES (Advanced Encryption Standard) Encryption Program with the Albany College of Nanoscale Science & Engineering (CNSE). A schematic of the chip layout is shown in Figure 3. The die is 32mm x 25mm and contains a range of memristor ReRAM configurations from single devices to 128 x 128 crossbar networks.
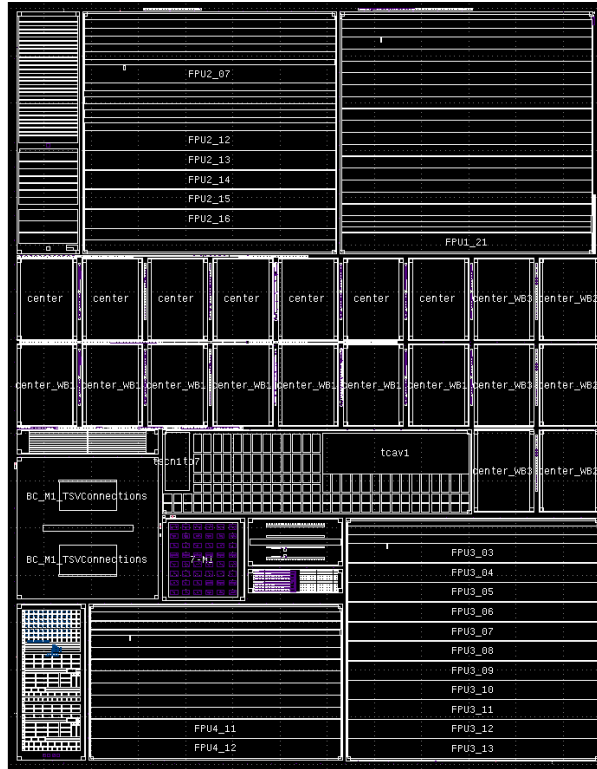
Figure 3. Schematic of the Complete Raptor Die

3.2 Physical Self-Reconfigurable Neuromorphic Computing Architectures

This section briefly describes work conducted in self reconfigurable architectures and application to development of a memristor based XOR logic gate.

Reconfigurable Synaptic System

The synaptic system was defined as a building block employing a variable resistor (either a CMOS transistor operating within the weak inversion and Ohmic region modes or a memristor) and inverter circuit elements [5]. Figure 4 depicts the complete synaptic system where (a) illustrates the circuit implementation of the synapse and (b) illustrates the simplified circuit element form that will be employed when designing the neuromorphic network. In (a), pfet is a p-channel field effect transistor and nfet is an n-channel field effect transistor. The figure shows that the output of the synaptic system is a function of the Vm potential that will either strengthen, weaken, or completely cut-off the connection between the synapse input (SI) and the CMOS inverting circuit operating within its transition bias point range.
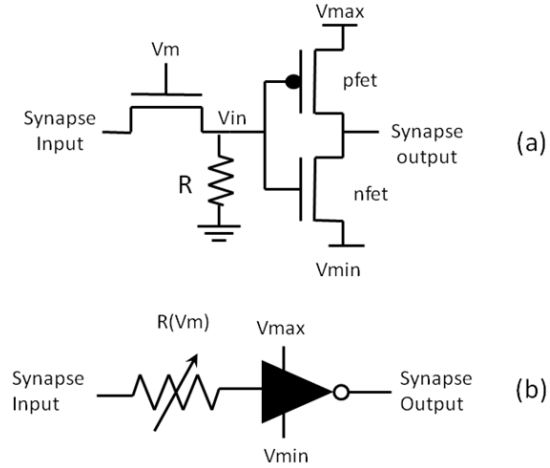
Figure 4. Synaptic System: (a) Circuit Representation, (b) Simplified Circuit Representation

The synapse output, SO was modeled employing a linear approximation to the electronic characteristic behavior of the CMOS inverting circuit by:

$$SO(SI, Vm) = f[I(Vm) * R] \tag{1}$$

The function f represents the transfer function of the CMOS inverter and I(Vm) is current across the transistor channel or memristor device. R is a resistor used to reset and properly bias the synapse. For example, for synaptic inputs between 0 and 2 V, as shown in Figure 5, the inverter transfer functions can be made to range from 3 to ~0 V approximately as a function of the biasing operating voltage Vin as shown in Figures 4 and 5. The resulting multiple synapse circuit is described in Section 4.2.
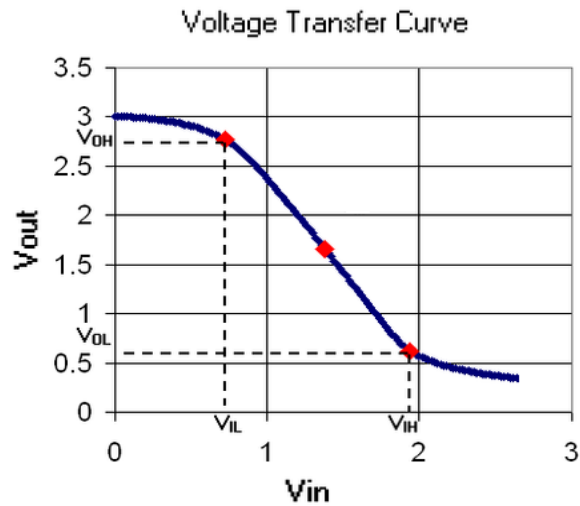
Figure 5. CMOS Inverter Voltage Transfer Curve

Reconfigurable Logic Circuit

Work was conducted with a reconfigurable logic gate circuit that uses the high and low resistance state of memristors to select the logic function. This reconfigurable logic gate circuit could be programmed to perform any logic function. Given an input state, the circuit could be programmed to output a zero or a one for that particular input set. After any input state is trained, the circuit remembered the output for that input state. The circuit worked by storing the desired output for each possible input state using memristors as a memory storage device to store the value of the output for each input state. The total number of memristors needed was equal to the total number of input states. This is equal to 2^n where n is the number of inputs. For example, a two input circuit required 4 memristors, a three input circuit required 8 memristors, and a 4 input circuit required 16 memristors.

The full two input reconfigurable logic gate circuit is shown in Figure 6. In this circuit four resistor-memristor voltage dividers were used to store the output state for each of the four possible input states. Transmission gates, T1 to T4, selected which voltage divider was connected to the comparator. The enabled inputs on the transmission gates were connected to the outputs of a 2 to 4 line decoder which selected which transmission gate to turn on, based on the two inputs. The truth table in Table 1 shows the outputs of the 2 to 4 line decoder, labeled as W, X, Y and Z for inputs A and B.
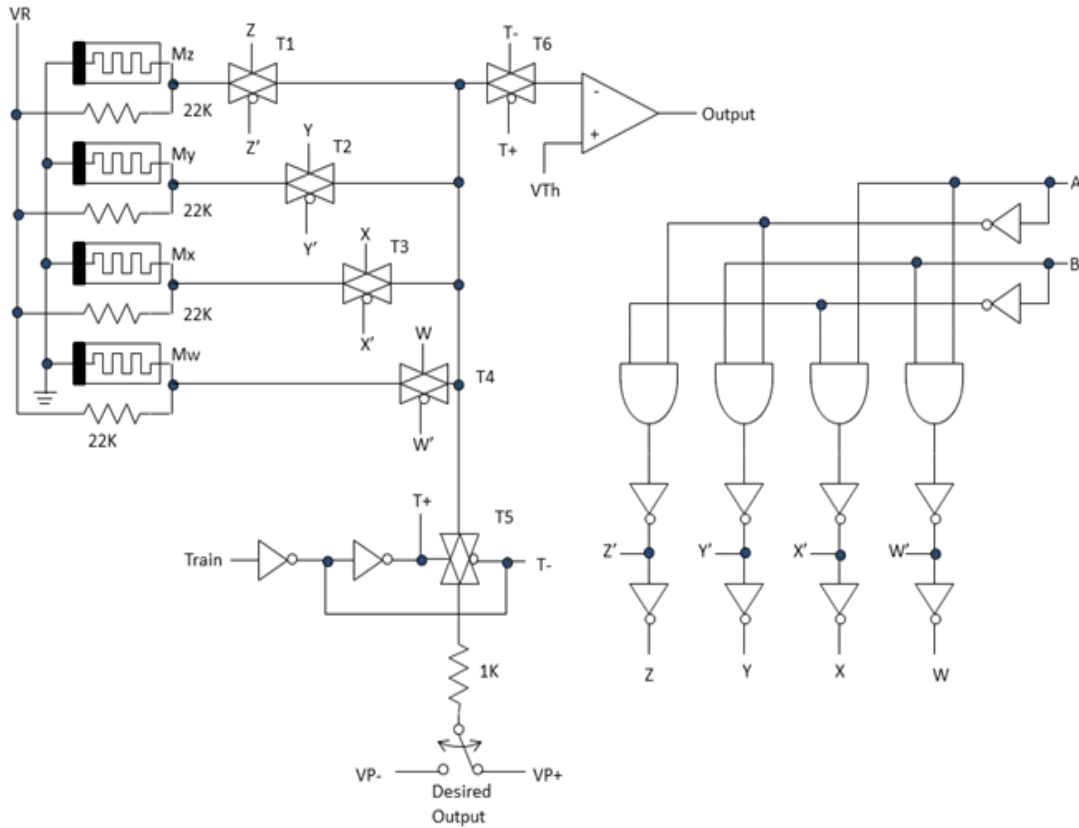
Figure 6. Two Input Reconfigurable Logic Gate

Table 1. Two to Four line Decoder Truth Table

A	B		W	X	Y	Z
0	0		0	0	0	1
0	1		0	0	1	0
1	0		0	1	0	0
1	1		1	0	0	0

3.3 Hardware-Based Artificial Neural Network Chips

Fully parallel, silicon-based neural network chips (CM1K) developed by General Vision / CogniMem Technologies and based on IBM's earlier series of zero instruction set computers were used for pattern recognition applications in large data sources. Hardware-based Artificial Neural Networks (ANN) are ideally suited for mobile or portable platforms with strictly limited SWaP resources such as those used in T&E data acquisitions platforms.

Change Detection in Video Data

In this work, chips containing 1024 parallel neurons were successfully used for change detection and pattern recognition in streaming video data. A graphical user's interface was written by in-house researchers to access and train the chip with the intent of utilizing the least amount of processor resources. The overall system design is given in Figure 7, a block diagram depicting the system configuration where the V1KU is prototyping board designed to analyze vector signatures within a charge-coupled device generated video stream.
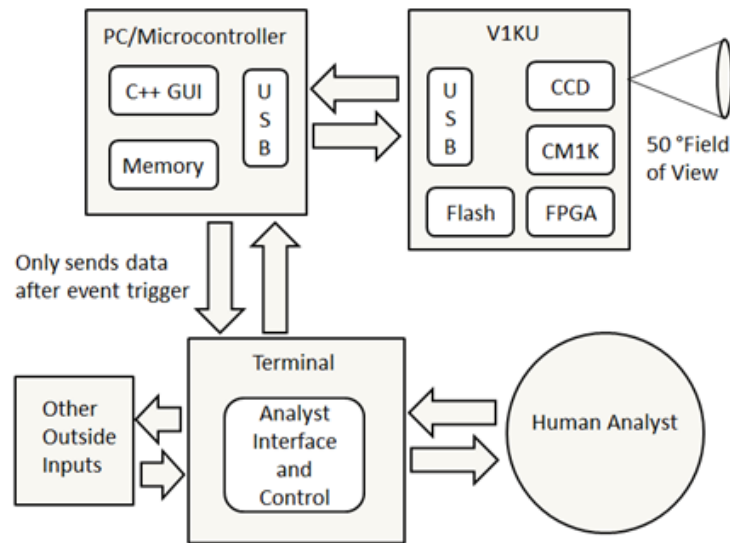
Figure 7. Video Analysis Functional Block Diagram

Change detection and simple object recognition were demonstrated with reduced neuron numbers utilizing only a few, or in some cases one, neuron per category. This simplified approach was used to validate the utility of few neuron networks for use in applications that

necessitate severe SWaP restrictions. The limited resource requirements and massively parallel nature of hardware-based ANNs make them superior to many software approaches in such resource limited systems.

3.4 Test Center Technology Transition

The last half of this effort's performance period saw an increase effort to identify an application area for our technology to address the needs of the T&E community. At the AFOSR T&E Program Review in Niceville, FL, April 2012, initial contact was made with representatives from various T&E locations. It was also learned that the AFOSR's Program Manager's intent for our effort to try to focus the needs of the T&E work at Arnold Engineering Development Complex, AEDC. Two issues that needed to be overcome were: presenting neuromorphic computing in a manner that the T&E community could easily see how the technology will be a benefit to them and getting one's foot in the door at AEDC.

Artificial Neural Network Management of Wind Tunnel Performance Data

Environmental test and evaluation facilities such as climatic laboratories and wind tunnels require significant control over operational factors in order to achieve and maintain critical testing parameters. The use of neurologically inspired systems for process control has the potential to improve environmental stability, operational efficiencies, and testing results, while reducing run times and operating costs. The work outlined in this section explored the use of neural network chips to evaluate and process wind tunnel performance data and built upon concurrent efforts funded by AFOSR, this effort and the Innovative Technology Applications Company / Notre Dame effort.

Hardware based neural networks offer significant benefits over traditional neural software algorithms. First, the parallel nature of the hardware provides increased throughput potential, where all the neurons simultaneously monitor the input data for previously trained signatures. The AFRL in-house system used in this project was built on the CogniMem CM1K platform. It operated with over 40,000 parallel neural nodes where each recognition cycle required 10 microseconds to complete, regardless of number of committed or trained neurons. The AFRL team worked closely with CogniMem engineers to solve power distribution and data latency issues to increase the node count to over 120,000 parallel neurons. Secondly, hardware-based neural networks have greatly reduced SWaP requirements when compared to traditional serial processing methods and software algorithms. For example, a desktop computer running a neural simulation utilizing 1,024 nodes (N=1,024) for pattern recognition can typically consume on the order of 100 W of power while the CM1K architecture consumes around 0.5 W per 1,024 nodes. The CM1K has additionally been shown to be 93 times faster than certain digital signal processors when N=1,024, where the speed-up linearly improves such that performance is enhanced 186 times at N=2,048, 279 times at N=3,072, and so on [6]. Finally, since these hardware-based neural implementations are built on Zero Instruction Set Computer architectures, they require no internal software to operate, effectively reducing the overhead and complexity required to operate a trained system. Drawbacks to the current implementation of the CM1K include its ridged structure and limited memory for signature storage. Reconfiguration and optimization can be more easily achieved in current software approaches but the AFRL team is

investigating adaptive and reconfigurable electronic versions in a complimentary aspect of this AFOSR sponsored program.

The utility of experimental neural chips for the management and analysis of wind tunnel performance data was examined by adapting the modeling and control work of Prof. Mark Rennie and Dr. Alan Cain to hardware based neural systems. A mathematical model that Rennie and Cain developed to predict wind tunnel performance from both calculated and measured physical parameters was used to generate a representative data set that in turn was used to train a neural network to aid in process prediction and control. The details of the modeling process is outlined in [7], to include theoretical assumptions and descriptions of the feature set used in the neural network. Their work demonstrated that neural networks can be used to efficiently manage wind tunnel performance data to produce useful analysis and predictions for operators and test engineers. The network was shown to provide relevant estimates on the state of useful parameters not directly measured by sensors within the wind tunnel unit. AFRL examined ways that hardware implementation of a neural network could enhance performance and provide control upgrades to legacy systems through the use of hardware-based low power neural networks that can be deployed at remote sites or in isolated sections of environmental test facilities where extensive renovations are not practical.

Enhanced SWaP Performance through Neural Pruning and Optimization

Artificial neural networks have demonstrated success in solving difficult nonlinear classification problems but often times are designed as simulations on high performance processors. Most existing neuromorphic concepts are developed without considering the system complexity and power consumption. Human brains have plasticity so that when complex tasks are learned, information is not simply added to memories; the physical structures of the brain is actually altered. Part of this process involves neural pruning which trims neural circuits in order to optimize wiring and reinforce frequently used pathways. Inspired by the biological neuroplasticity of the brain to react to changes in behavior, the environment and neural processes, this portion of the AFRL program investigated a neuromorphic architecture with autonomous optimization ability, capable of changing synaptic configurations in response to the changing application needs and environments. Specifically, a neuronal pruning and multi-objective, heuristic self-optimization approach was investigated to dynamically determine the optimal synaptic structure. This involved evaluating the significance of individual neurons, desired performance, available resource and power goals, rather than operating with statically overprovisioned resources.

Data Processing Methods

The realization of large scale hardware-based neural networks not only offers new capabilities in computational intelligence, but also requires new methods of data manipulation. Such large scale neural networks, for example, offer new possibilities for the speed-up of big data analysis. AFRL designed a data architecture that utilizes identically trained multiple neuronal banks for concurrent processing of high speed data with a linear increase of processing speed as 1/N, where N is the number of neuronal banks employed. The design is based on the single instruction multiple data (SIMD) paradigm as depicted in Figure 8, where multiple parallel

neuron banks perform the same operations but are staggered in time, resulting in increased data throughput.
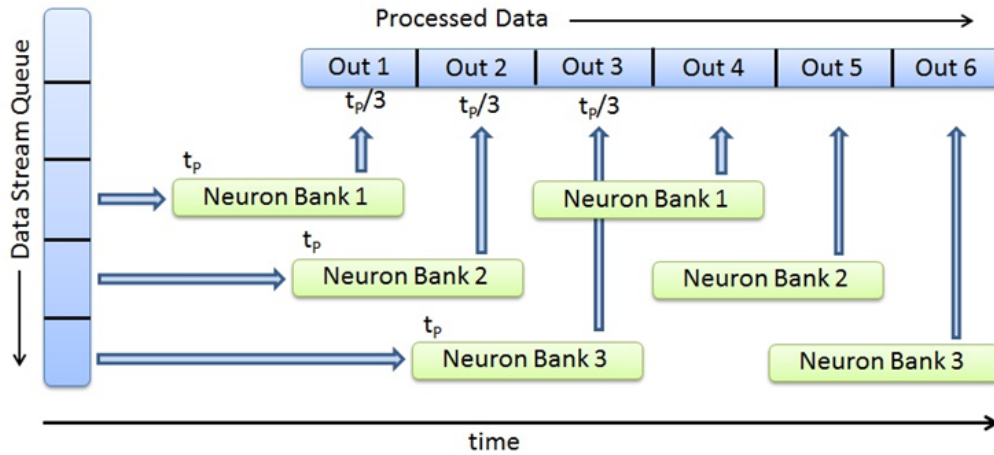
Figure 8. The SIMD Process

This approach takes advantage of the newly available high number of neurons and the massively parallel nature of neuromorphic circuitry. An outline of the varying hierarchies to be applied is given below:

Single Instruction Multiple Data (SIMD)

- Multiple processors performing the same operation but staggered in time

- Data throughput parallelism

- Banks of identically trained neurons

- Increased data rates

- Utilize surplus neurons for increased speed

Multiple Instruction Single Data (MISD)

- Many operations on same data at the same time

- Functional parallelism

- Each neuron in a bank has its own unique target signature

- Increased functionality and speed

Multiple Instruction Multiple Data (MIMD)

- Many sensors processed simultaneously

- Perceptual parallelism

- Distributed memory

- Biologically inspired

Serial Processing

- Allows hierarchical structures

- Redundancy

- Decision trees

- Self-learning

4.0 RESULTS AND DISCUSSION

This section describes the results of the project. Performance and yield testing results are given for memristive devices and application of the memristor XOR circuit architecture to self-reconfigurable neuromorphic processing is shown. Architectures for complete neuromorphic processors are described along with some transition work towards potential applications.

4.1 Memristive Device Design and Fabrication

The memristive device level work centered on three materials sets, a Ag-Ge-Se chalcogenide device, created by BSU a $Al/Cu_xO/Cu$ device created by AFRL and HfO_x devices created by CNSE. This section describes the results of AFRL modeling of the chalcogenide samples and testing of $Al/Cu_xO/Cu$ and HfO_x samples.

Chalcogenide Based Memristive Devices

Sample output from the in-house empirical memristor compact model, showing the simulated behavior of the characteristic Lissajous I-V curve, is given in Figure 9. This model formed the basis for much of the early memristive circuit simulations and designs. Memristor hardware is shown by red circles, the linear model by a green line and the compact model by a blue line, the later clearly showing greatest fidelity to actual device behavior. The compact model follows actual behavior very closely.

(a)   (b)
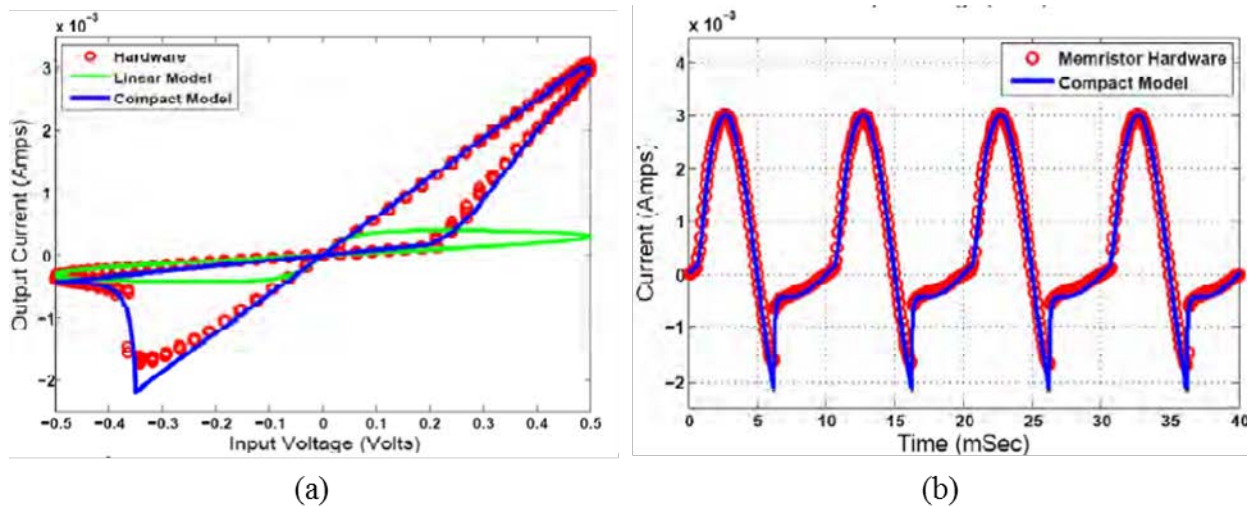
Figure 9. Model Simulations versus Measured Data

In-house Al/Cu$_x$O/Cu Memristive Devices

During fabrication of the Al/Cu$_x$O/Cu based devices, power settings of the plasma oxidation system were shown to control the grown oxide thickness and implanted oxygen concentration, which subsequently affected memristive device behaviors. These memristive devices demonstrated complete nonpolar behavior and could be switched either in a vertical (Al/Cu$_x$O/Cu) or lateral (Al/Cu$_x$O/Cu/Cu$_x$O/Al) manner. The switching mechanism of these devices was shown to be filamentary in nature. Figure 10 gives experimentally measured I-V plots of the nonpolar switching behavior. I-V curves of all four possible switching styles were observed in memristive devices: a) bipolar(+), b) bipolar(-), c) unipolar(+) and d) unipolar(-), where the blue curve is the SET operation and the red curve is the RESET operation
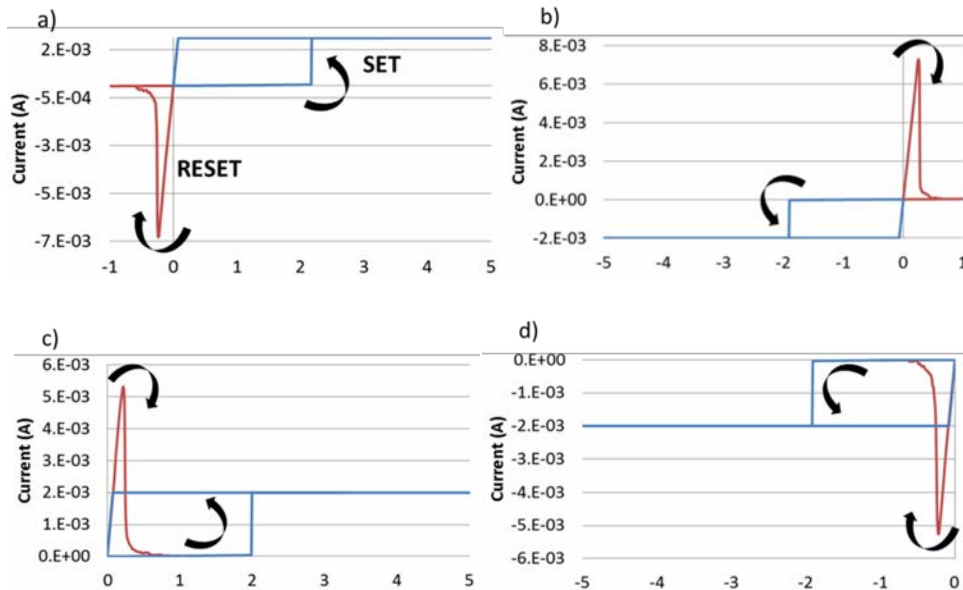
Figure 10. I-V Curves of all Four Possible Switching Styles

These devices demonstrated consistent, complete nonpolar behavior with operating voltages typically within |3| V and required currents under 20 mA. These operating parameters are acceptable for use in CMOS circuitry.

Hafnium Oxide Memristive Crossbars

The active regions of the devices are based on based on hafnium oxide as depicted in Figure 11. The Nano-dimensional Crossbar Architecture is shown, developed with CNSE. In (a) is a schematic of a 2x2 junction, showing the tellurium and copper pathways. In (b) is an

image of interconnecting layers. In (c) is a device stack and in (d) is an image of the cross sectional layers. The smooth analog transitions between resistive states demonstrated by initial testing of these devises shows promise for artificial electronic synapse applications where continuous and predictable evolution allows for higher density synaptic memory. A typical set and reset curve depicting the characteristic hysteresis behavior is shown in Figure 12. In (a) is a Bipolar ReRAM 1Transistor1Resistor configuration and its operational forming charactoristics. A typical SET/RESET curve is given in (b) showing the smooth analog behavior of these devices and their charactoristic hysteresis properties.
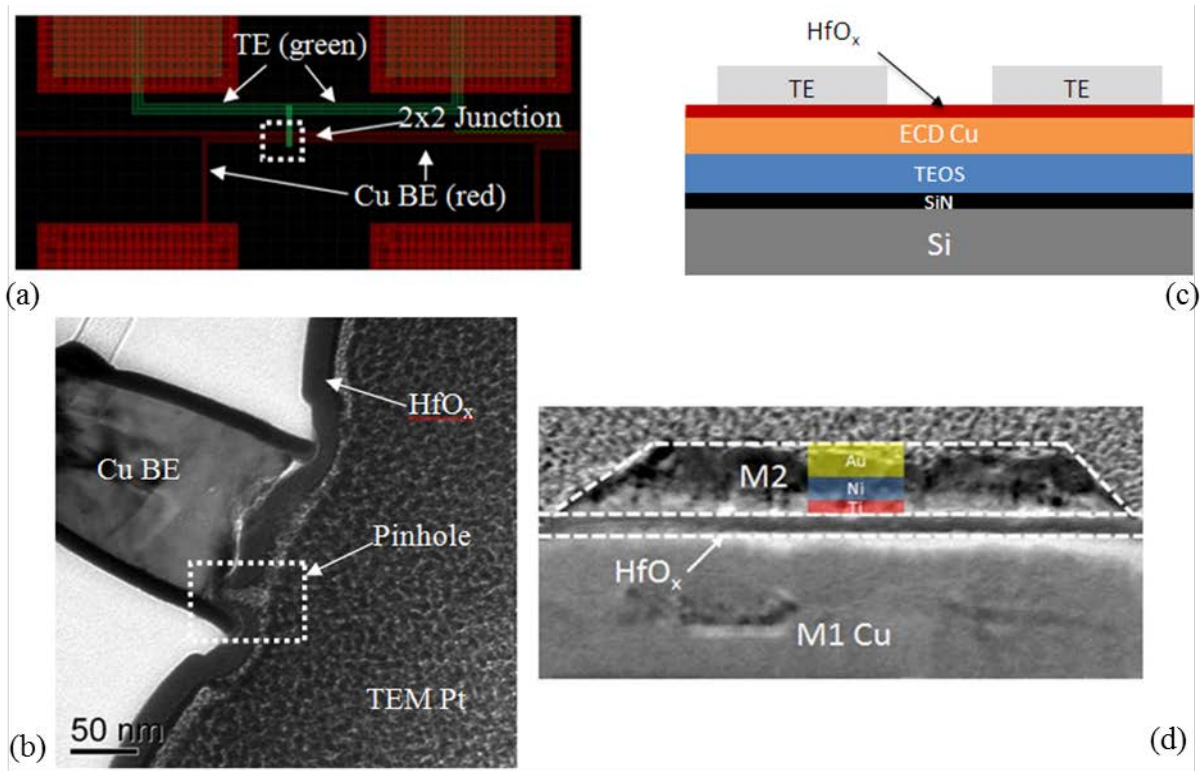
Figure 11. Nano-dimensional Crossbar Architecture
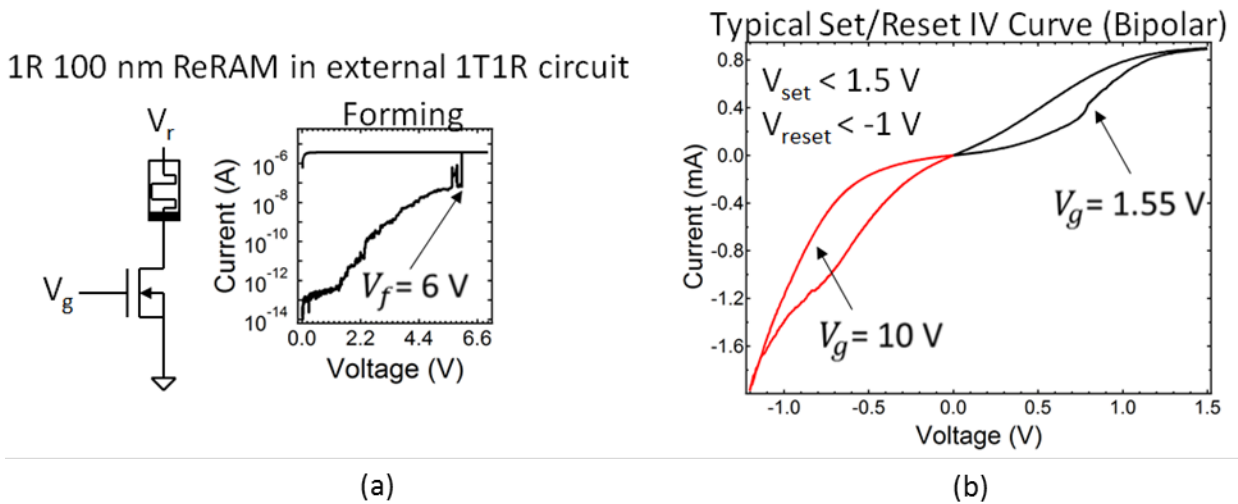
Figure 12. ReRAM Example (Courtesy Albany CNSE)

4.2 Physical Self-Reconfigurable Neuromorphic Computing Architectures

This section briefly describes the results of work conducted under this effort in self reconfigurable architectures and application to development of an XOR logic gate.

Reconfigurable XOR Logic Gate

The implementation of the neuron functionality will be performed with an adding node and/or a single transistor as displayed in Figures 13 and 14. In Figure 13, multiple synaptic outputs converge at the floating adding node and their combined response is fed to the next neuron synaptic layer of the neuromorphic computing architecture. In Figure 14, multiple synaptic outputs converge at the adding neuron where their relative contribution (the adding of all post-synaptic output potentials) will cause the single metal–oxide–semiconductor field-effect transistor, MOSFET neuron to fire as long at the overall synaptic contribution is above its threshold potential. The neuron output Vo will be fed to the next synaptic layer of the neuromorphic computing architecture. The adding node is the physical connection where all post-synaptic outputs will converge.
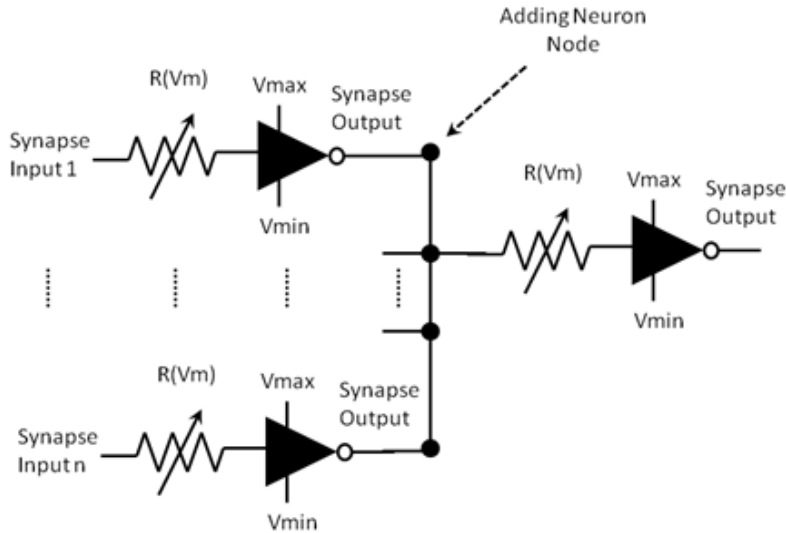
Figure 13. Multiple Synaptic Outputs Converge at the Floating Adding Node
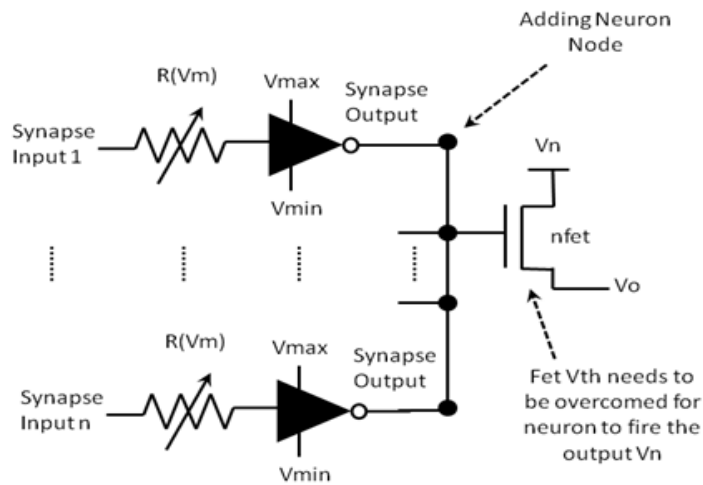
Figure 14. Multiple Synaptic Outputs Converge at the Adding Neuron

As the synaptic outputs converge, they will increase or decrease the potential at the floating neuron adding node. Thus, the resulting added potential will become the input to the following synaptic layer. In addition, if the neuron adding node were connected to the gate of, for example, an NFET transistor and if the total combined potential at the adding node is greater than the threshold voltage, Vth, of the MOSFET transistor, then the output of the neuron will be

Approved for Public Release; Distribution Unlimited.

Vo=Vn. Otherwise, the neuron will not output a high potential (the neuron will not fire). Figure 15 displays the computing architecture required to implement the XOR function with the in-house developed neuromorphic computing architecture [5] classically described in [8]. The circuit in Figure 15 is also an example of a neuromorphic computing concept that demonstrates how a computing architecture can be implemented with thresholding neurons.
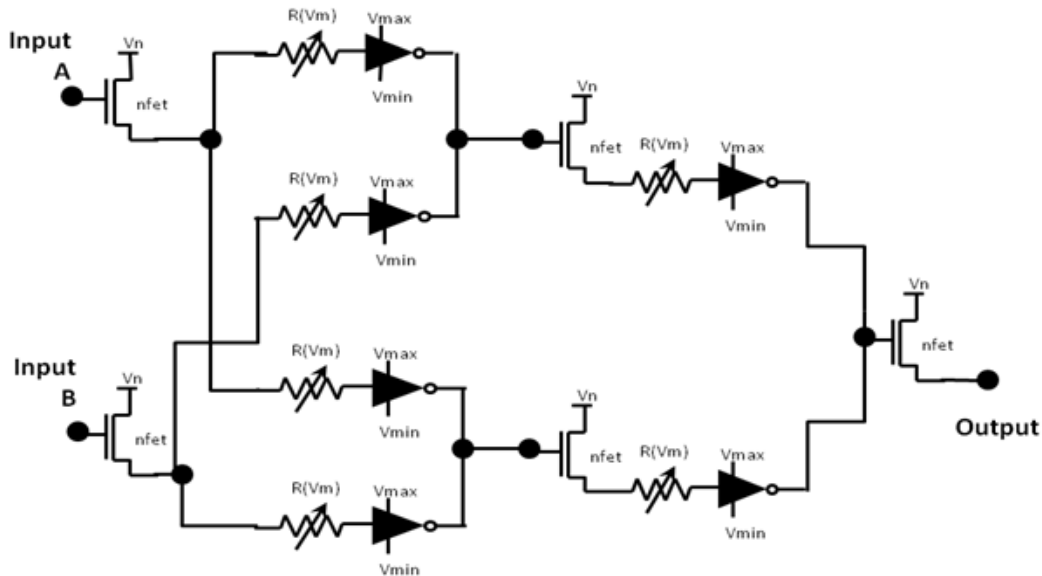
Figure 15. Neuromorphic Computing Implementation of the XOR Function

Demonstration of Neuromorphic Functionality

The AFRL in-house Computational Intelligence research group demonstrated the physical hardware operation of a self-reconfigurable neuromorphic computing architecture in the laboratory. The results illustrated the operation of one input neuron, one memristor-based synapse and one output neuron as shown in the circuit schematic diagram in Figure 16. The figure displays a transistor-based input neuron that is composed of four transistors and two biasing resistors. The synapse was formed by a memristor, two biasing resistors and a CMOS pass-gate to control the appropriate current flow. The autonomous training circuit employed in which a transistor based comparator circuit to take a desired output and send an electronic signal to the memristor device that caused it to change impedance state to switch the output of the circuit to match the desired output electronic signal.
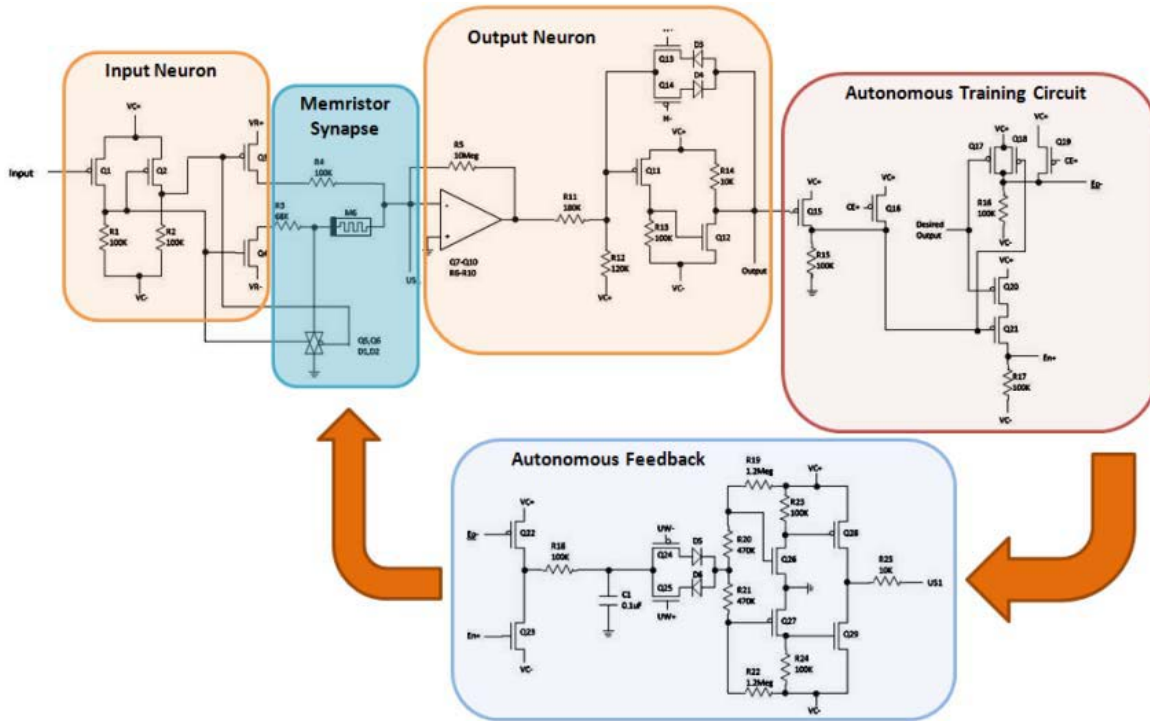
Figure 16. Neuromorphic Self-Reconfigurable Electronic Circuit

The electrical characteristic results of the in-house designed and tested neuromorphic computing architecture are shown in Figure 17 where a 'Desired (V)' electrical input was shown to the circuit, then the circuit was 'Enabled (V)' to learn and reconfigure itself to match the desired electrical signal. Once training was enabled, the circuit calculated and sent an autonomous feedback pulse to the memristor-based synapse causing it to change resistance. Then the circuit output was able to match the desired input. Another very important result demonstrated was that when the power was completely switched off to the entire circuit, the neuromorphic circuit persisted in the state it was in before the loss of power; the circuit "remembered". This important result gave a good first indication that potentially instant on computers could be a real possibility.
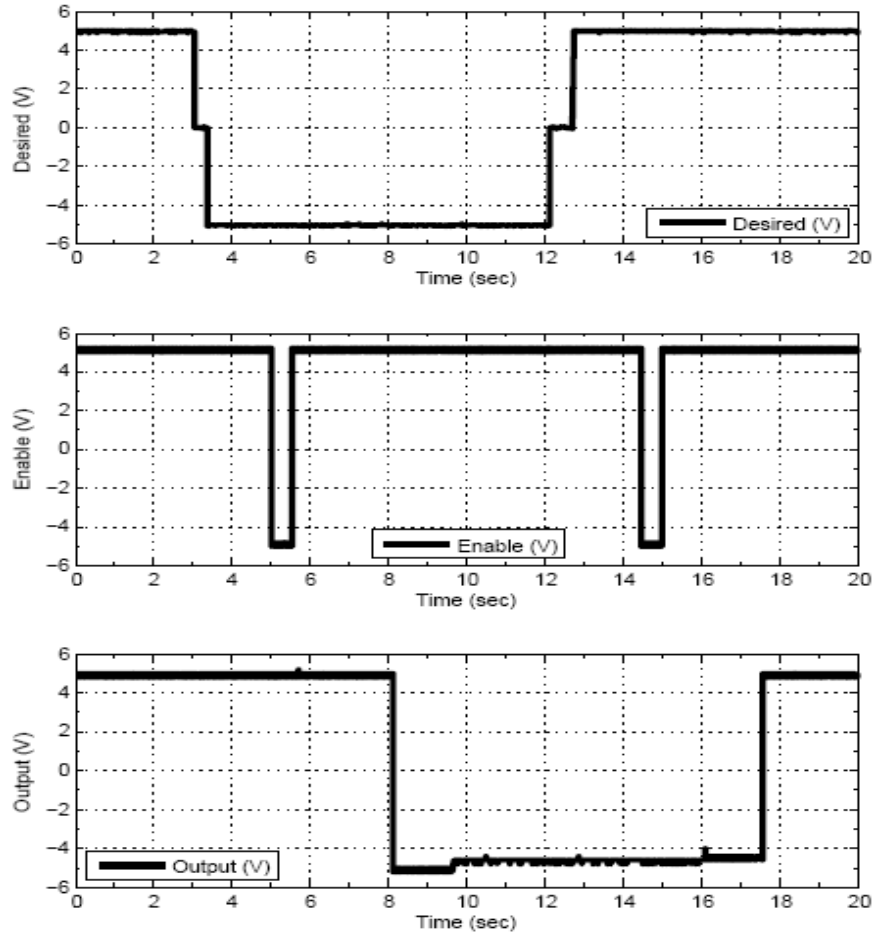
Figure 17. Electrical Results for a Neuromorphic Self-Reconfigurable Electronic Circuit

Memristor-Based Neural Network of 100 Synapses Fabricated & Characterized

The in-house research effort achieved a critical milestone in the design, fabrication and characterization of a 100 synaptic memristor-based neural network. A memristor-based neural network leveraging BSU's ion-conductor chalcogenide-based memristor devices was fabricated by Professor Kris Campbell at the research and development laboratory facilities at BSU. They were characterized by AFRL in collaboration with BSU. The fabrication of the neural chips, led. The characterization results are displayed in Figure 18 where (a) shows the discrete memristance values achieved and (b) shows the contour plots of the memristance maps. The results demonstrated that each synapse or memristor element in the neural network can be programmed individually into various memristive states.
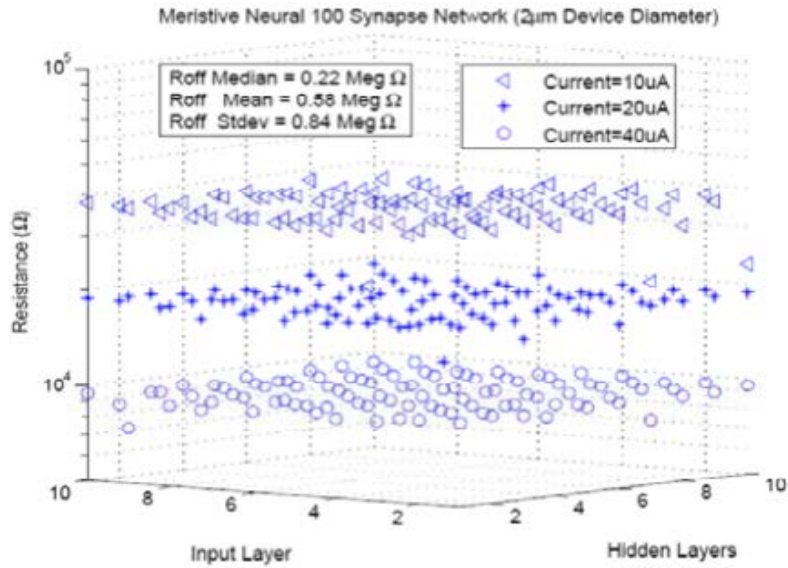
(a)
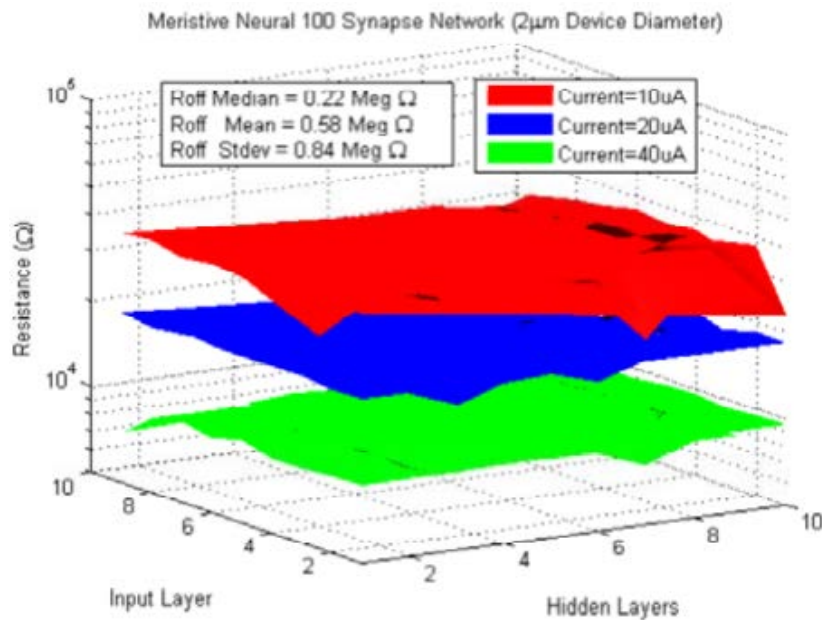
(b)

Figure 18. Characterization of a 100 Synapse Memristor-Based Neural Network

Three distinct memristance states were programmed for each of the 100 memristors in the neural network. Thus, three distinct planes were observed from the characterization results data.

Approved for Public Release; Distribution Unlimited.

However, in practice any memristance state can be programmed from approximately a high of $2x10^6$ to a low $4x10^3$ Ohms. In addition, the narrow distribution exhibited at the various current characterization points indicated that the fabrication process is reproducible with minimal geometric variation within the 2 micrometer fabrication process. The final packaged chips are shown in Figure 19.
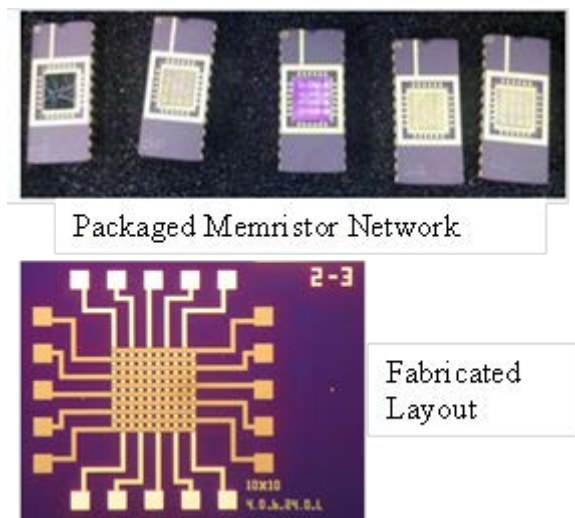
Figure 19. Fabrication of a 100 Synapse Neural Network

Reconfigurable Logic Circuit

A reconfigurable logic circuit was designed with two inputs. Figure 20 shows the two input circuit set up on a breadboard. Figure 21 shows the results of training the reconfigurable logic gate circuit with one memristor input where Vp+ was set to 10 V, Vp- to -10 V, Rc1 to 1 kohms and Rc2 to 1 kohms. At the time the experimental work with the breadboard was performed, there were issues getting the full two input circuit working due to problems with finding enough working memristors. The status of all of the memristors on hand was manually verified.
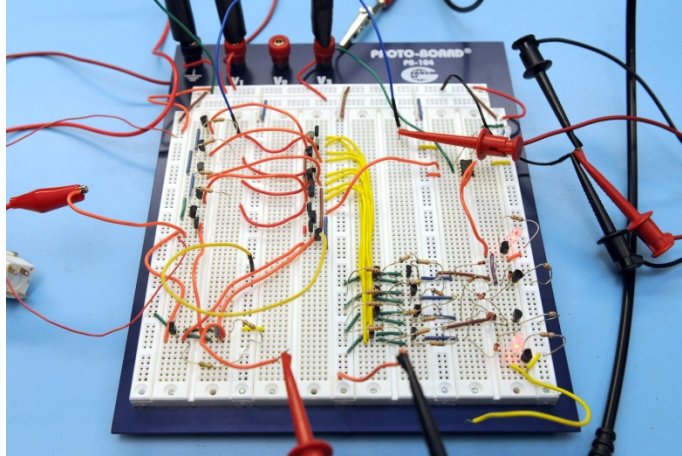
Figure 20. Two Input Circuit on a Breadboard
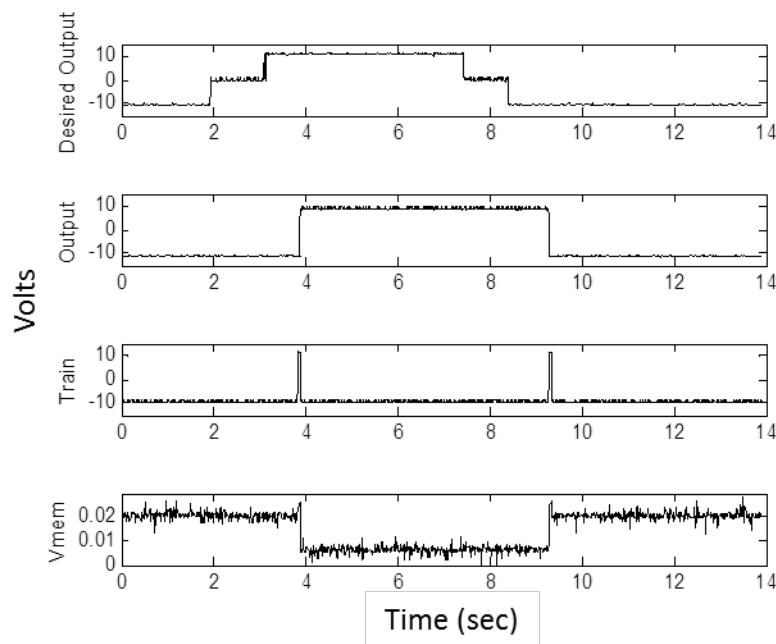
Figure 21. Results of Circuit Operation with One Memristor

A three input reconfigurable logic gate was made in the same way as the two input gate. For a three input gate, eight memristors were needed for the eight possible input states and a three to eight line decoder was used to select the memristor.

A more robust circuit prototype was fabricated for additional experimentation. The reconfigurable logic gate circuit, shown in Figure 22, is a logic gate with two inputs and one output that can be configured to implement any two input, one output logic function. The circuit used four memristors to store the output for each of the four possible input conditions.
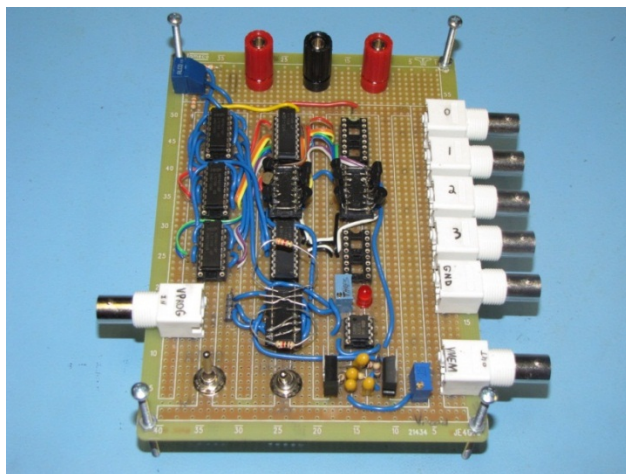
Figure 22. Reconfigurable Logic Circuit

The circuit read the resistance state of each memristor using a voltage divider circuit composed of the memristor to be read and a fixed resistor. The voltage drop across the memristor depended on the resistance of the memristor. A high resistance resulted in a high voltage and low resistance resulted in low voltage. To generate a binary value from the voltage, the voltage was compared with a threshold. If the voltage was greater than the threshold the output was a binary 1. A voltage less than the threshold produced a binary 0.

The circuit was operated in two programming modes: pulse mode and arbitrary waveform mode. Pulse mode was used for applying pulses to the memristors for programming. Arbitrary waveform mode allowed an arbitrary programming waveform to be applied to the memristor from an external waveform generator. While arbitrary waveform mode could be used to apply pulses to the memristor, regular mode allowed for setting and resetting of the memristor to be selected simply by setting a switch. In arbitrary waveform mode, the pulse generator voltages had to be changed each time the memristor was set or reset. Also, in arbitrary waveform mode, the output impedance of the function generator appeared in series with the memristor to be programmed. Pulse mode operation did not have this problem since the pulse voltages were supplied from DC power supplies that had negligible output impedance.

A simplified diagram of the circuit in the pulse mode configuration is shown in Figure 23. When the read/write switch, S1, was closed, the circuit was in the read mode and a voltage divider circuit was formed from Rread and the memristor selected by closing A, B, C or D. In the

actual circuit, A, B, C, and D were controlled by the two binary inputs. A, B, C and D were closed when the inputs were {0,0}, {0,1}, {1,0}, {1,1}, respectively. The voltage developed across the selected memristor was converted to a binary value using op-amp Op2 configured as a comparator. The comparator compares the memristor voltage with the read threshold voltage and outputs a high or low binary value if the memristor is greater or less than the threshold voltage respectively. Op-amp Op1 was configured as a non-inverting buffer amplifier that measured the voltage across the selected memristor without loading the circuit significantly. The output of Op1 could be measured using an oscilloscope or voltmeter.
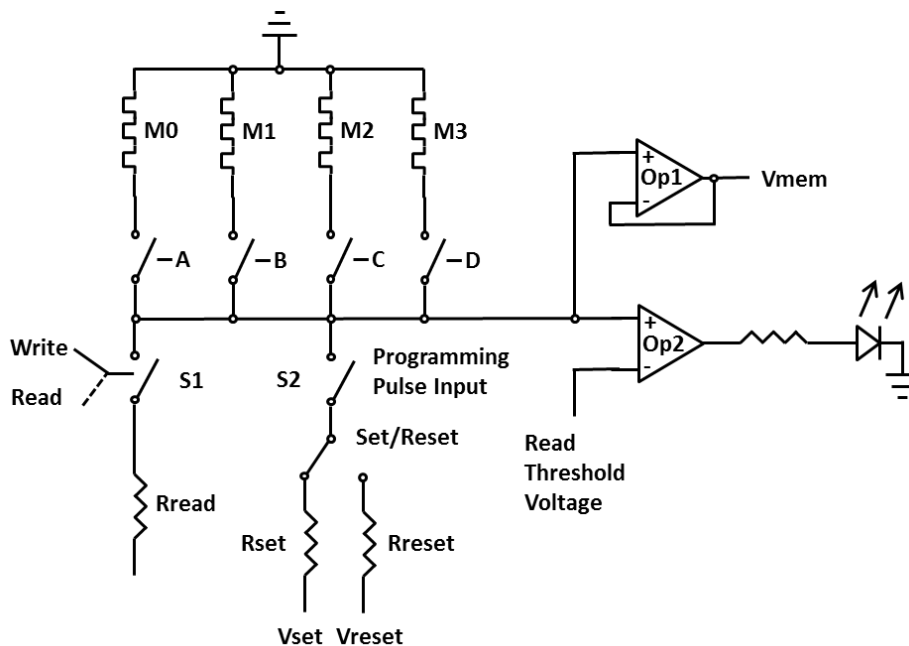
Figure 23. Pulse Mode Simplified Circuit

When S1 was opened, the circuit was in the write mode and the read voltage was removed. Initially, S2 was opened and no voltage was applied across the memristor. The Set/Reset switch was first set such that when the programming pulse was applied, the proper voltage and current limit were selected to switch the memristor to the desired state. A pulse from an external pulse generator was then applied to the programming pulse input. This pulse caused S2 to close and apply the programming voltage across the memristor for the duration of the pulse, in theory causing the memristor to switch to the desired state. Figure 24 shows the pulse mode setup for the test equipment on the lab bench.
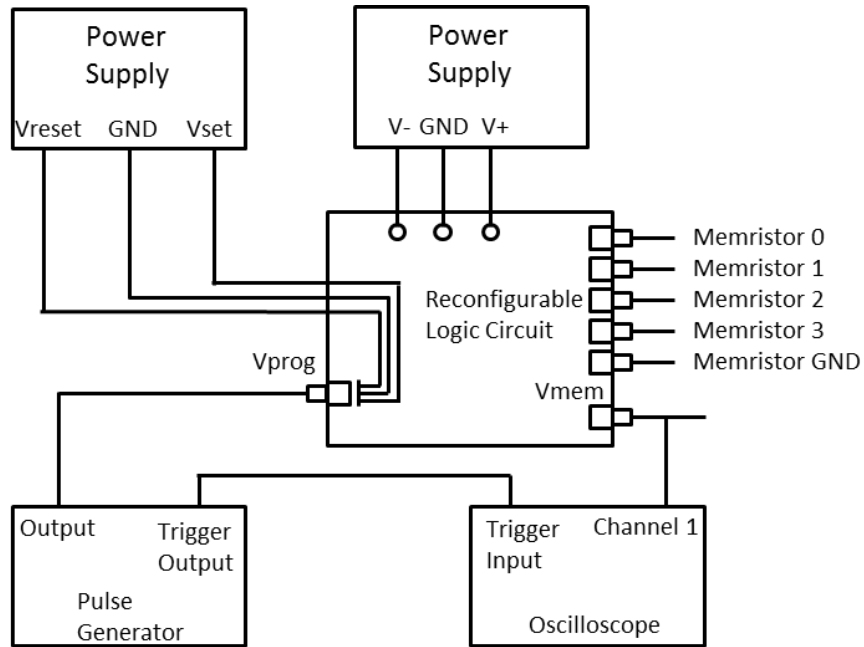
Figure 24. Pulse Mode Setup on Lab Bench

A simplified circuit diagram in the arbitrary waveform mode configuration is shown in Figure 25. In this mode, S1 and S2 were controlled simultaneously. When S1 was closed, S2 was open and the circuit was in read mode. The read operation was the same in arbitrary waveform mode as in pulse mode. When S1 was opened, S2 was closed and the circuit was in write mode. A programming waveform supplied from an external waveform generator was applied across the selected memristor with a series resistance selected by the Set/Reset switch. Figure 26 shows the arbitrary waveform mode setup for the test equipment on the lab bench.
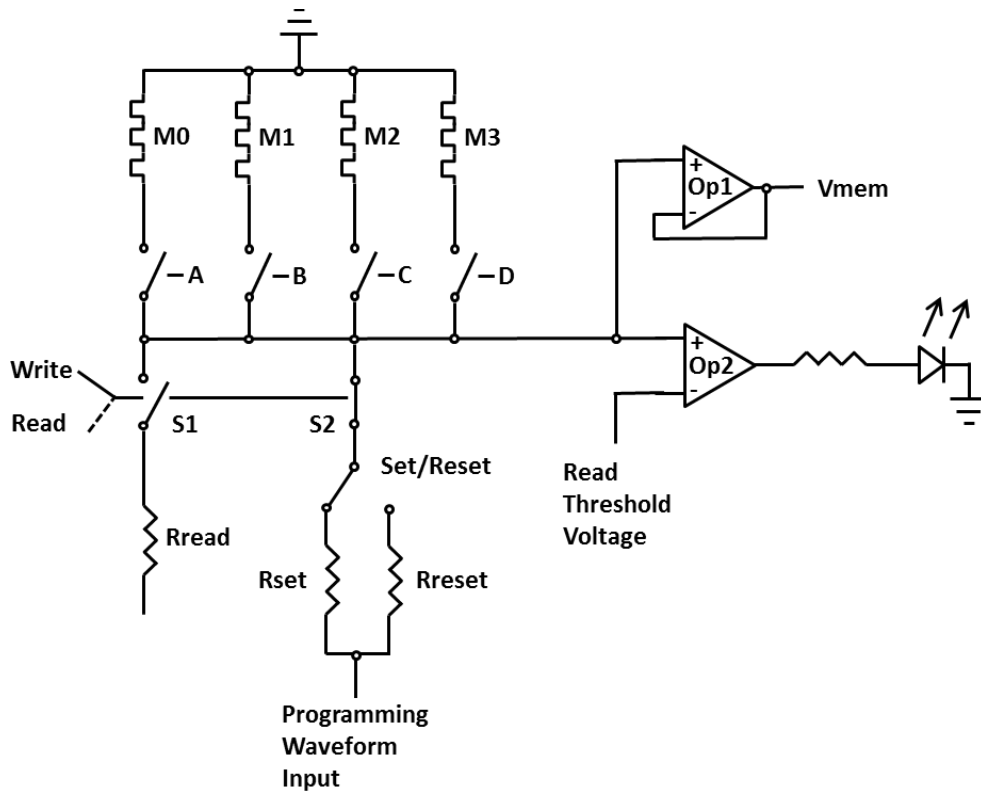
Figure 25. Arbitrary Programming Waveform Mode Simplified Circuit
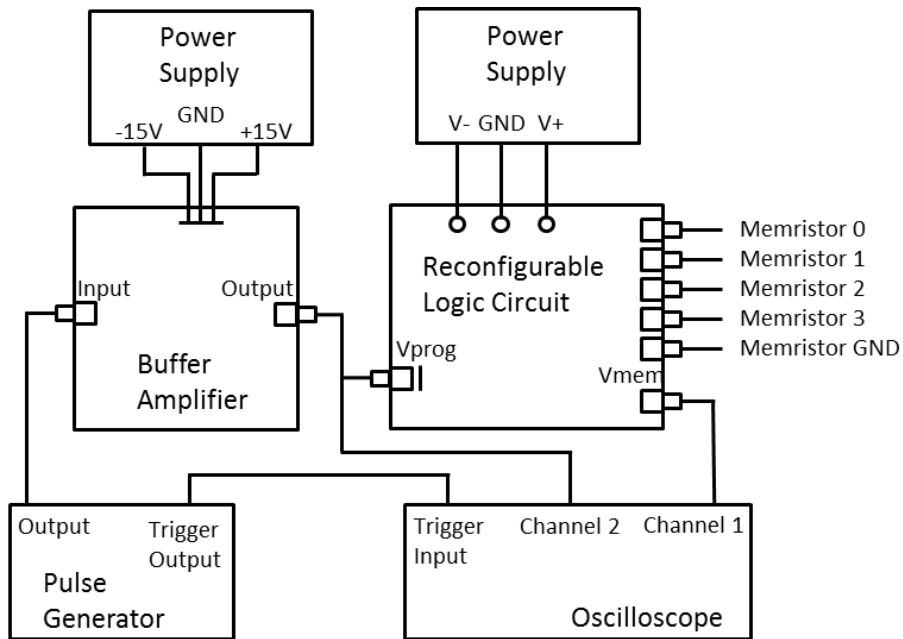
Figure 26. Arbitrary Waveform Setup on Lab Bench

4.3 Hardware-Based Artificial Neural Network Chips

This section describes the results of exploration of the use of neuromorphic processing for DoD related problems.

Change Detection in Video Data

The CogniMem hardware is expected to work with a variety of data types. There was some work that explored the issues of utilizing different types of sensors including their control and data flow. In the case of small unmanned systems, microcontrollers are certainly capable of handling simple onboard processing tasks, however, as the complexity of autonomy increased so did the need for more computing horsepower and on-board memory. In one instance, the use of neural networks for interpreting the optical flow of information was examined. An attempt was made to use a feed forward neural network with a back propagation training algorithm to help determine when a vehicle should stop in front of an obstacle. A straight forward algorithm was found to work better than a more complex neural network. While this was not an intensive study of the use of neural networks for obstacle avoidance, the results do illustrate that the application of neural networks may not be the best solution for all perception problems.

Solid State Neural Networks for Network Intrusion Detection and Cyber Security

The expertise gained from aspects of research funded by this project directly contributed to recent collaborations with the Army Research Laboratory and the National Security Agency in areas concerning the use of solid state neuromorphic architectures for cyber security applications. The joint effort resulted in the formation of an interagency working group, a conference focusing on the application of neuromorphic technologies to network science (The Network Science and Reconfigurable Systems for Cybersecurity Conference, 28-29 Aug 2012) and a technology interchange meeting hosted at the Air Force Research Laboratory, AFRL Information Directorate.

The working group leveraged the capabilities of many DoD, academic and industrial agencies, resulting in stronger and more resilient in-house efforts. The skills gained by such technical interactions were directly applied to the AFOSR T&E effort with possible application to the testing and evaluation of United States Air Force cyber platforms.

To demonstrate the use of neuromorphic processing for network monitoring, a prototype of an in-house developed data monitoring circuit was built. The system, shown in Figure 27, was constructed from a CogniMem PM1K (bottom right), an Arduino microcontroller (top right), and data interface circuit (top left). The Arduino Mega2560 board was used in this prototype, but in conversations with CogniMem Technologies technical staff there was concern about the board being inadequate for the test. Since this work was conducted, the Arduino Duo board with an ARM processor was made available which holds more promise. The use of the Arduino board is now being examined in a follow-on in-house effort.
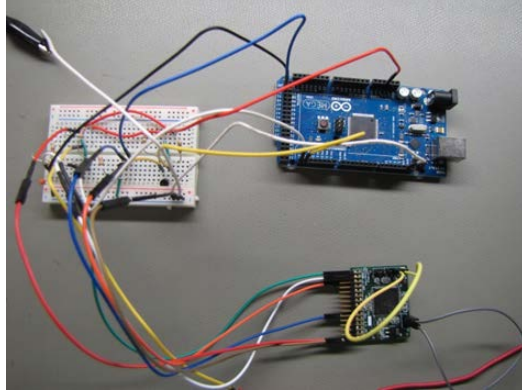
Figure 27. Prototype of a 1024 Neuron Data Monitoring System

4.4 Test Center Technology Transition

This project helped leverage AFOSR's Small Business Technology Transfer Program (STTR) technology investment. The particular effort involved Dr. Alan B. Cain from Innovative Technology Applications Company and Prof. R. Mark Rennie from the University of Notre Dame. The effort involved the use of neural networks to organize and manipulate wind-tunnel performance data into formats useful for tunnel optimization and control. The STTR effort was leveraged as a means to show the T&E community how neuromorphic computing technology can be used to bring more efficiency to their work and provide intelligent control to their systems.

Discussions on applications of neuromorphic processing were conducted with Capt. Alexander Henning, AEDC, including during the 2012 International Test and Evaluation Association Symposium. One thing he reiterated was the need to explain neuromorphic computing in a manner that those with limited technical knowledge of the technology area will understand its potential.

There was also discussion of possibly redirecting the AFRL effort to address the needs of the McKinley Climatic Laboratory at Eglin AFB, FL, if working with AEDC did not pan out. A teleconference was held on 27 September 2012, between the PM, Air Force Research Laboratory Information directorate, AFRL/RI, the Innovative T Technology Applications Co. / Notre Dame team and various individuals from Eglin, AFB. This was an initial discussion to explore the possibility of applying the STTR control system work toward the needs of one of its facilities. From the teleconference, it seemed like it would be easier to obtain real data. Discussions continued and a proposal was put together by AEDC for the Environmental Security Technology Certification Program for their Advanced Building Energy Management and Control Topic Area for FY 2014. The proposal did not make it through the selection process.

Management of Wind Tunnel Performance Data Using Hardware-Based Artificial Neural Network Chips

The techniques demonstrated in early testing were adapted for use in control systems managing complex automated processes. The chosen process concentrated on extracting useful information from sensor data characterizing wind tunnel performance for increased process control and optimization. The team worked closely with CogniMem Technologies' engineers to obtain a 102,400 neuron system (at the time of this effort, the largest system of its kind in the world), theoretically capable of finding a single signature among 100,000+ possible categories within 10μs while using only 150 mW per 1,000 patterns. Figure 28 is an image of the 102,400 neuron system at the AFRL Information Directorate. Each board consisted of 4 individual 1024 neuron chips orchestrated by a Field Programmable Gate Array and a communications bus. Utilization of only 40,960 neurons was successfully demonstrated due to communications latency and power distribution issues. The 40K neurons, however, proved quite adequate for demonstrating the efficiencies of this architecture. Figure 29 shows the current 10 board, CogniBlox configuration being used in AFRL research activity. The 10 board CogniBlox stack was connected to a laptop
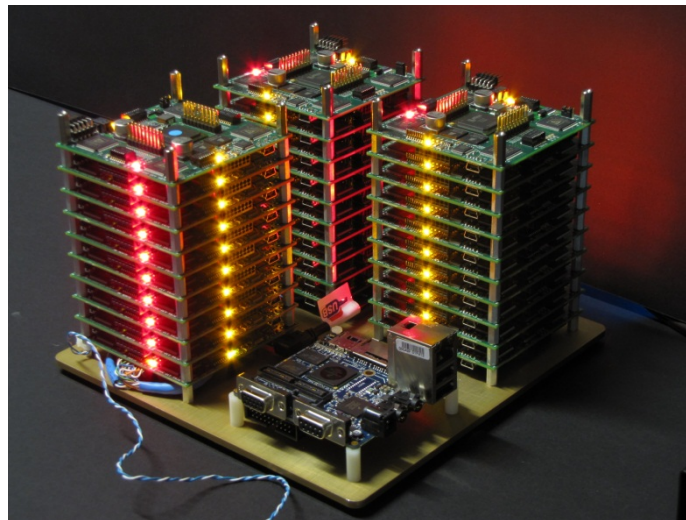
Figure 28. Three Stacks of Modules Making Up a 102,400 Neuron System
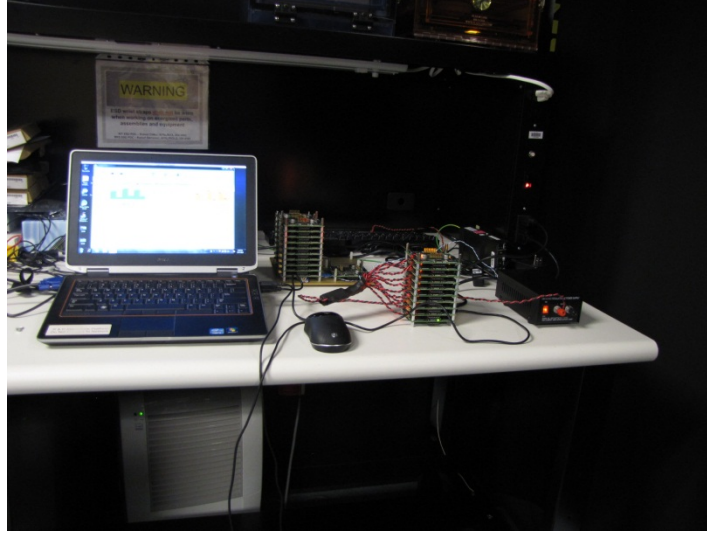
Figure 29. Current Research Setup Using CogniMem Technology Hardware

Starting with Prof. Rennie's and Dr. Cain's earlier work, the use of solid state neural networks was examined to extract secondary results from typically-measured wind-tunnel control and status parameters, thereby effectively leveraging other AFOSR funded programs. This collaboration area is relatively new for AFRL and still being defined, but recent e-mail exchanges have reinvigorated the relationship. Some preliminary results are given below.

Compression of Feature Data for Input into Hardware Neural Memory

The memory that each otherwise identical neuron has for signature feature storage is limited to 256 8-bit registers per neuron. Thus all input data variables must be preprocessed to conform to the available 8-bit representation. Such a limited amount of storage essentially means that each variable must be restricted to fall within 256 bins that represent the continuous spectrum of feature values. This compression can result in ambiguous signature representations within the neural memory. That means data representation must be carefully considered when providing both training and test data to the neural network. In order to maximize the dynamic range of the limited storage, a linear algorithm was used to convert the data to a form acceptable by the memory register.

$$data' = \left\lfloor \left(\frac{data - data_{min}}{spread_{data}} \right) spread_{register} + register_{min} \right\rfloor, \qquad (2)$$

where *data'* was the data after conversion to fit within the 8-bit (256 bin) memory, *data* was the unconverted raw data and the symbols $\lfloor \ \rfloor$ represent the floor function. The min and max values for *register* are 0 and 255, thus *spread_{register}* is the difference or 255. This compression algorithm

was used to try and maximize the dynamic range of the available neuron memory, but additional preprocessing on features with relevant statistical outliers or nonlinear distributions was needed such that careful preparation of the data was occasionally required.

Table 2 illustrates feature data from a signature set that required logarithmic scaling prior to conversion. The use of scaling on two highly variant feature sets reduced conflicting data conversions from over 39,000 within a data set containing 260,000 signatures to 0 conflicting categories. Close analysis of the dataset statistical make-up with the appropriate preprocessing reduced degenerate neuron counts and increased distinguishability between signatures. The table compares the original data (that had no contradictions), to the scaled data (that then had two categories with the same signature), to the same data scaled with a logarithmic function so that the categories are again distinguishable. The table shows a representative sample of the data showing how the two compression algorithms affected the distinguishability of the signatures, where the red indicates identical values.

Table 2. A Representative Data Sample

	Category	Feature Value 1	Feature Value 2	Feature Value 3	Feature Value 4	Feature Value 5
Original Signature 1	Cat 1	147.04	-3.1311	33.397	98,786	98,281
Original Signature 2	Cat 2	149.6	-3.1319	33.396	98,786	98,283
Signature 1 after linear scaling	Cat 1 or 2	1	148	11	72	222
Signature 2 after linear scaling	Cat 1 or 2	1	148	11	72	222
Signature 1 after log scaling	Cat 1	18	148	11	74	226
Signature 2 after log scaling	Cat 2	19	148	11	74	226

The limited memory associated with each neuron created the non-trivial requirement to pre-process or compress the input data for optimized signature representation. This compression results in reduced feature resolution and more ambiguity between different signatures.

Simulations of the CogniMem hardware were used to train and test the network on signature analysis and general classification tasks using wind tunnel simulation data supplied by Notre Dame and Innovative Technology Applications Co. Network training was accomplished using subsets of signatures taken from the overall database of 260,000. Training sets as large as 160,000 were examined but became difficult to manage due to the large memory requirements of the simulations. A representative cluster of 10,000 signatures, each with 36 feature variables, was designated to train the network for the recognition tasks. The simulated signature data was previously generated using physical models of wind tunnel performance as outlined in [7]. The idea was to use neural networks as a means to generalize the models and to act as predictive systems for wind tunnel control and optimization with the ability to extract secondary results such as the model drag area from the data. Hardware implementation of the ANNs brought such capabilities into size, weight and power constrained areas where the addition of large computer networks was not practical.

Performance Comparisons between Software and Hardware Neural Network Implementations

Much of the appeal for biologically inspired neuromorphic hardware is in its potential for increased performance on specific tasks with reduced power requirements. In this section, the results of a signature recognition task conducted in hardware and software platforms are compared. As part of the implementation of hardware-based neural networks for the management of wind tunnel performance data, recognition times were recorded and compared for equivalent tasks on both platforms.

The test utilized wind tunnel performance data consisting of 262,144 signatures made from 36 individual feature variables and describing 18 category bins (model drag area). A selected population of signatures (9,600 in total) representing the 8 categories (1,200 each) was used as the training set. Both the software and hardware based platforms were trained on these signatures prior to the timed recognition test. In both instances the committed neuron count after training was identical at 8,897.

The software version of the radial basis function neural network was run on a high performance dual Intel Xeon processor with eight 3.07 GHz cores and 48 Gigabytes of memory. The recognition task examined all 262,144 signatures of the original dataset against the training data, requiring ~29.0 minutes to complete. The hardware-based, zero instruction set computer radial basis function neural network was run on a stack of 40K identical CogniMem neurons and finished the task in ~9.06 minutes using less than 12 W of power. This equated to increased performance (3.1 times speed-up) with lower SWaP requirements (1.7% of the power), assuming a typical load of 692W for the Xeon system used [10].

Although the training set was not optimized to reduce recognition error, the results support the energy efficient performance of Application Specific Integrated Circuit hardware designs. The results also suggested that embedded neuromorphic chips have the potential to add autonomy and intelligence to the control of complex industrial processes like the T&E

environmental testing facilities. Further improvements to performance and efficiency may be available through self-optimization of neural circuits and through neural pruning as discussed in the next section.

SWaP Performance Enhancement through Neural Pruning and Optimization

The results presented in this section were made possible by aligning addition resources funded through the AFOSR Summer Faculty Fellowship Program (Contract Number FA8750-13-2-0116) in partnership with Professor Zhanpeng Jin and Simon Chen from Binghamton University (BU) - SUNY, Binghamton NY, whose work on neural optimization was extremely beneficial to this project. The general techniques and results are given below. Details of this collaborative work will be available through the final reports that are being prepared connected with the contract number given above.

Oversized neural networks that carry too many less-influential neurons, significantly increase the burdens on resources and power consumption, as well as weaken the network's generalization capability [9]. To address this problem, researchers at BU and AFRL investigated adaptive neural pruning and a heuristic self-optimization technique to optimize the neuromorphic hardware from a comprehensive perspective. Energy-aware neuronal pruning can increase the energy efficiency of neuromorphic hardware by deactivating the least functionally significant neurons (i.e., "least recently fired (LRF)"). The self-optimization not only focused on the recognition accuracy but also considered the recognition singularity and energy efficiency. Professor Jin established a hybrid objective function for the heuristic multi-objective optimization.

$$\min R(W) = \alpha\varepsilon_a(W) + \beta\varepsilon_s(W) + \gamma\varepsilon_c(W) \tag{3}$$

R(W) is the system cost, $\varepsilon_a(W)$ is the accuracy cost (i.e., the distance between the input and the reference pattern), $\varepsilon_s(W)$ is the singularity cost (i.e., the number of fired neurons), $\varepsilon_c(W)$ is the complexity cost (i.e., the number of activated neurons) and α, β and λ are the regulatory parameters. The rationale of this approach is that, a neuromorphic system needs to be co-optimized towards the goals of minimizing the system cost, reducing recognition errors, avoiding the ambiguity caused by concurrent firing of multiple neurons and reducing the resource and energy consumption. [11]

Professor Jin's experiments showed that, by strategically selecting the neurons with more significant impacts, the neural network can maintain high recognition performance with only a limited number of neurons (e.g., up to 90% recognition rate with only about 15% of the neurons). The multi-objective optimization provides a way of achieving a balanced tradeoff between recognition accuracy and resource utilization, and thus enables the possibility of further improving system overhead and energy efficiency. When combining these two approaches, it was observed that both the recognition accuracy and neuron efficiency are significantly improved, while still maintaining the recognition rate at a reasonable level. This study laid the

foundation for future research on highly flexible, scalable and energy-efficient neuromorphic architectural frameworks for mission-critical defense applications that demand a rather high level of massively parallel computing, especially in SWaP constrained environments. [11]

4.5 Transitions Resulting From This Project

The accomplishment of technology transfer was a significant goal of this project. Five patents were awarded with two pending in the areas of neuromorphic hardware and memristor based circuit designs as a result of this and other complementary in-house efforts. Numerous papers and presentations were achieved and are listed in Appendix A

5.0 CONCLUSIONS

The in-house research efforts reached several critical milestones which were recognized through the award of five patents with two pending for work related to the invention of autonomous reconfigurable memristive technologies for reconfigurable electronic circuits and electronic memristive synapses. The synaptic computing element in the reconfigurable circuits is autonomously programmed whenever the circuits perceive new information and learning is enabled by the application. These technical accomplishments will ensure the government a potential return on investment an intellectual property protection. Refer to the patents descriptions in Appendix A for more detail.

This effort provided a great opportunity to pursue fundamental knowledge regarding hardware-based neuromorphic computing technology and its potential to address the needs of the Air Force. Basic building blocks, models and testing techniques were developed from which new computing architectures were create based upon memristive device technology. Commercial technology was identified that will allow understanding of system integration and architectures issues. This understanding will not only facilitate pursuit of opportunities to move emerging neuromorphic computing technology to the field faster, but also guide the development of far term technologies. Much of the success under this effort will be used in a new 6.2 in-house effort entitled, "Solid State Neural Technologies (SSNT)," and other efforts within the Information Directorate. Application areas include: addressing shortfalls of the processing, exploitation and dissemination chain, enhancing computer architectures for the T&E community, bringing computational intelligence closer to the sensor and a variety of cyberspace needs. Under SSNT technology will be explored that can benefit the developmental T&E community with whom there was little interaction prior to this LRIR. Results suggest that embedded neuromorphic chips have the potential to add autonomy and intelligence to the control of complex industrial processes like the T&E environmental testing facilities.

With that said there is still more basic research that needs to be pursued to properly develop practical hardware-based neuromorphic computing architectures for the Air Force. Models for memristive technologies require more refinement for accurate and efficient design and simulation of new architectures. A methodology is needed to move from sound device models to accurate system models. Understanding is needed of the issues of preparing data for processing in advanced hardware-based neural networks and other neural processors. Advanced computing architectures should have the agility to address a variety of Air Force applications, but the proper techniques are needed to work with a variety data types. The technology will need to interact with different systems including both legacy and emerging concepts that are land, air and space based. Future architectures must work with time dependent input, real time data and tap into a variety of databases. Given current budget issues and limits on available manpower trying to tackle these issues will be difficult. Fortunately, as a result of working on this LRIR, numerous collaborative opportunities have arisen with academia, industry and other government agencies that allow this burden to be shared with a larger community with similar interests.

6.0 REFERENCES

[1] Thiem, C., Wysocki, B., Bishop, M., McDonald, N., and Bohl, J., *Foundations of Neuromorphic Computing*, AFRL-RI-RS-TR-2013-125, AFRL Information Directorate, Rome NY, 2013.

[2] Pino, R., Bohl, J., McDonald, N., Wysocki, B., Rozwood, P., Campbell, K., Oblea, A. and Timilsina, A. "Compact Method for Modeling and Simulation of Memristor Devices: Ion Conductor Chalcogenide-Based Memristor Devices," *IEEE/ACM International Symposium on Nanoscale Architectures (NANOARCH)*, 2010, pp. 1-4.

[3] McDonald, N. "Al/Cu_xO/Cu Memristive Devices: Fabrication, Characterization, and Modeling." Master of Science Thesis, University at Albany, State University of New York, 2012, URL: http://www.dtic.mil/dtic/tr/fulltext/u2/a567650.pdf. Accessed 30 Nov 2013

[4] Rose, G., Kudithipudi, D., Khedkar, G., McDonald, N., Wysocki, B. and Yan, L. "Nanoelectronics and Hardware Security," **Network Science and Cybersecurity**, Springer, New York, 2014, pp. 105-123.

[5] Pino, R., Patent US8275728. URL: http://patft.uspto.gov/netacgi/nph-Parser?Sect2=PTO1&Sect2=HITOFF&p=1&u=/netahtml/PTO/search-bool.html&r=1&f=G&l=50&d=PALL&RefSrch=yes&Query=PN/8275728. Accessed 29 Mar 2013.

[6] CogniMem Technologies, Inc. "Disruptive Performance for DSP WHITE PAPER," URL: http://www.cognimem.com/_docs/White-Papers/WP_CM1K%20disruptive%20performance%20for%20DSP.pdf. Accessed 1Dec 2013.

[7] Rennie, R., and Cain, A. "Management of Wind Tunnel Performance Data Using Neural Networks," *Proceedings of the 50th AIAA Aerospace Sciences Meeting*, 2012-0321.

[8] Lawrence, J. and Luedeking, S., "Introduction to Neural Networks," California Scientific Software, Grass Valley, CA 1991.

[9] Lawrence, S., Giles, C., and Tsoi, A. "What Size Neural Network Gives Optimal Generalization? Convergence Properties of Back-Propagation," Technical Report CS-TR-3617, University of Maryland, College Park, MD, 1996.

[10] "80 PLUS Verification and Testing Report," www.plugloadsolutions.com/psu_reports/DELTA_DPS-1050DBA_SO-035_1250W_Report.pdf, last accessed 14 January 2014.

[11] Jin, Z. "Towards Energy-Aware Self-Optimizing Neuromorphic Hardware," Final Technical Report under FA8750-13-2-0116, to be published.

APPENDIX A – Patents, Publications and Presentations

Technology transition was a large component of this effort. Following is a list of Patents, publications and presentations that were a direct result of this effort.

Patents Granted:

The first two patents reflect concepts that evolved out of both AFOSR and AFRL 6.2 investment and are being leveraged by this particular effort. The third patent represents work that has taken place since this effort was first proposed and was accomplished by this effort.

1. U. S. Patent No.7,902,857, "Reconfigurable electronic circuit," by Dr. Robinson Pino. Issued 8 March 2011. This patent provides an apparatus and method for designing reconfigurable electronic computing systems. The invention relies on the ability to change the resistive state of a memristor device to achieve an optimal voltage at a specific circuit node. Whereby this dynamically and autonomously causes the circuit to reconfigure itself and produce a different output for the same input relative to the circuit's initial state. The circuit's state remains constant until the memristor's resistance is changed, at which point the circuit's function is "reprogramed".

2. U. S. Patent No. 8,249,838, "Method and Apparatus for Modeling Memristor Devices," by Dr. Robinson Pino and Mr. James Bohl was issued 21 August 2012. "The invention provides methods and an apparatus for accurately characterizing the linear and non- linear Lissajous current-voltage behavior of actual memristor devices and incorporating such behavior into the resultant model. The invention produces a model that is adaptable to large scale memristor device simulations."

3. U. S. Patent No. 8,275,728, "Neuromorphic Computer," by Dr. Robinson Pino was issued 25 September 2012. "The invention provides variable resistance circuits to represent interconnection strength between neurons and a positive and negative output circuit to represent excitatory and inhibitory responses, respectively. The invention provides advantages over software-based neuromorphic computing methods."

4. U. S. Patent No. 8,274,312, "Self-Reconfigurable Memristor-Based Analog Resonant Computer," by Dr. Robinson Pino and Mr. James Bohl was issues 25 September 2012. The invention provides, "An apparatus which provides a self-reconfigurable analog resonant computer employing a fixed electronic circuit schematic which performs computing logic

operations (for example OR, AND, NOR, and XOR Boolean logic) without physical re-wiring and whose components only include passive circuit elements such as resistors, capacitors, inductors, and memristor devices. The computational logic self-reconfiguration process in the circuit takes place as training input signals, which are input causing the impedance state of the memristor device to change. Once the training process is completed, the circuit is probed to determine whether the desired logic operation has been programmed."

5. U. S. Patent No. 8,427,203, "Reconfigurable Memristor-Based Computing Logic," by Dr. Robinson Pino and Dr. Youngok Pino. Issued 23 April 2013. An apparatus for reconfigurable computing logic implemented by an innovative memristor based computing architecture. The invention employs a decoder to select memristor devices whose ON/OFF impedance state will determine the reconfigurable logic output. Thus, the resulting circuit design can be electronically configured and re-configured to implement any multi-input/output Boolean logic computing functionality. Moreover, the invention retains its configured logic state without the application of a current or voltage source.

Patents Pending:

A. Patent Application Pending: "One Bit hardware-based Random Number Generator" by Inventors: Nathan McDonald and Bryant Wysocki, Apparatus and method for generating a truly random binary bit value in which the resistance of memristive devices serve as the analogue bit values. Said resistance values and the randomness by which they are generated stems from the inherently non-uniform, irreproducible variations in the materials of which the system is composed. This patent application is a direct spinoff of memristive research funded by this program.

B. Patent Application Pending: "A Write-Time Based Memristive Physical Unclonable Function" Inventors: Garrett S. Rose, Nathan McDonald, Lok-Kwong Yan and Bryant Wysocki, Impact: Novel use of memristive device physics to address difficult supply chain issues with refurbished electronics and counterfeit chips, as well as methods to potentially enhance hardware identification, forms the basis for new methods in hardware trust and cyber security. This patent application is a direct spinoff of memristive research funded by this program.

Published in Peer Reviewed Journals, Books, etc.:

Yakopcic, C.; Taha, T.M.; Subramanyam, G.; Pino, R.E.; Rogers, S.; , "A Memristor Device Model," *Electron Device Letters, IEEE* , **Vol. 32**, no.10, Oct. 2011, pp.1436-1438

Kozma, R.; Pino, R.; G.E. Pazienza, G., (Eds.), **Advances in Neuromorphic Memristor Science and Applications**, Springer Series in Cognitive and Neural Systems, Heidelberg, Springer, 2012.

Zhang, L., Chen, Z., Joshua Yang, J., Wysocki, B., McDonald, N., Chen, Y., "A Compact Modeling of TiO 2-TiO 2–x Memristor", *Applied Physics Letters*, **Vol. 103**, no. 15, 2013, pp. 153503-153503-4.

Liu, B., Chen, Y., Wysocki, B., and Huang, T. "Reconfigurable Neuromorphic Computing System with Memristor-Based Synapse Design." *Neural Processing Letters*, Aug. 2013, pp. 1-9.

Wysocki, B., McDonald, N., Thiem, C., Rose, G., & Gomez II, M., "Hardware-Based Computational Intelligence for Size, Weight, and Power Constrained Environments", **Network Science and Cybersecurity**, 2014, Springer, New York, pp. 137-153.

Liu, B., Chen, Y., Wysocki, B., & Huang, T., "The Circuit Realization of a Neuromorphic Computing System with Memristor-Based Synapse Design", **Neural Information Processing**, LNCS, **Vol. 7663**, Jan 2012, Springer, New York, pp. 357-365.

Peer reviewed conference publications:

Wysocki, B., McDonald, N., Thiem, C., "Hardware-based Artificial Neural Networks for Size, Weight, and Power Constrained Platforms", accepted for presentation at the SPIE Defense and Security Symposium (2014)

Ji, F., Li, H., Wysocki, B., Thiem, C., McDonald, N., "Memristor-based Synapse Design and a Case Study in Reconfigurable Systems" *2013 International Joint Conference on Neural Networks IEEE*, (To be published in IEEE Xplore)

Liu, B., Chen, Y., Wysocki, B., Huang, T., "The Circuit Realization of a Neuromorphic Computing System with Memristor-based Synapse Design" *Proceedings of the 19th*

International Conference on Neural Information Processing ICONIP2012, Doha Qatar, Nov, 2012, pp. 12-15.

Sun, Z., Bi, X., & Wysocki, B., "Spintronic devices: From Memory to Memristor," *IEEE 11th International Conference on Solid-State and Integrated Circuit Technology ICSICT*, Xi'an China, 2012, pp. 1-4.

Rose, G.; Rajendran, J.; Manem, H.; Karri, R; Pino, R., "Leveraging Memristive Systems in the Construction of Digital Logic Circuits," *Proceedings of the IEEE*, **Vol**. **100**, no. 6, 2012, pp. 2033-2049.

Wang, H., LiH., Pino, R., "Memristor-Based Synapse Design and Training Scheme for Neuromorphic Computing Architecture," *The 2012 International Joint Conference on Neural Networks (IJCNN)*, Brisbane Australia, 2012, pp.1-5.

Pino, R., Li, H., Chen, Y., Hu, M., Liu, B., "Statistical Memristor Modeling and Case Study in Neuromorphic Computing," *49th ACM/EDAC/IEEE Design Automation Conference DAC*, San Francisco, 2012, pp. 585-590.

McDonald, N., Pino, R, Rozwood, P., Wysocki, B., "Analysis of Dynamic Linear and Non-Linear Memristor Device Models for Emerging Neuromorphic Computing Hardware Design," *The 2010 International Joint Conference on Neural Networks (IJCNN),* Barcelona Spain, 2010, pp. 1-5.

Pino, R., Bohl, J., McDonald, N., Wysocki, B., Rozwood,P., Campbell, K., Oblea, A., Timilsina, A., "Compact Method for Modeling and Simulation of Memristor Devices: Ion Conductor Chalcogenide-Based Memristor Devices." *IEEE/ACM International Symposium on Nanoscale Architectures NANOARCH*, Anaheim California, 2010, pp. 1-4.

Invited Presentations

B. Wysocki served as an Invited expert panel member for conference wide presentation/discussion at the 3rd International Conference on Computing for Geospatial Research and Applications in Washington DC, 1-3 July 2012. His topic was "Neuromorphic Computing for Big Data."

B. Wysocki and C. Thiem briefed IARPA visitors on 26 July 2012, on the integrated technologies developed in our group during a tour of the Nanotechnology/Computational Intelligence facility. AFRL and IARPA leadership were in attendance. Briefing highlighted AFRL/RI's commitment to 6.1 basic research and highlighted recent RITB successes to high level officials.

C. Thiem presented our computational intelligence work at the Network Science and Reconfigurable Systems for Cybersecurity Conference (NSRSCC 2012, Beltsville MD), 28-29 Aug 2012 titled "ASIC-based Artificial Neural Networks for SWaP Constrained Applications."

B. Wysocki, N. McDonald, and G. Rose gave invited talks on 24 Sep 2013 at Sandia National Laboratories on memristive system design and neuromorphic architectures for autonomous systems.

N. McDonald, C. Thiem, and B. Wysocki gave invited talks on 18 Oct 2013 to faculty and students at SUNY Binghamton regarding "Nanotechnology for Neuromorphic Computing and Hardware-based Computational Intelligence" as part of the university's research seminar series. Much of the talks were based on research funded by this program.

LIST OF SYMBOLS, ABBREVIATIONS AND ACRONYMS

AEDC	Arnold Engineering Development Complex
AFOSR	Air Force Office of Scientific Research
AFRL	Air Force Research Laboratory
AFRL/RI	Air Force Research Laboratory Information Directorate
AIAA	American Institute of Aeronautics and Astronautics
ANN	Artificial Neural Network
BSU	Boise State University
CMOS	Complimentary Metal-Oxide Semiconductor
CNSE	College Of Nanoscale Science and Engineering
MOSFET	Metal–Oxide–Semiconductor Field-Effect Transistor
NFET	N-Channel Field Effect Transistor
PFET	P-Channel Field Effect Transistor
ReRAM	Resistive Random Access Memory
SIMD	Single Instruction Multiple Data
STTR	Small Business Technology Transfer Program
SWaP	Size, Weight and Power
T&E	Test and Evaluation
Ag	Silver
Al	Aluminum
Cu	Copper
Ge	Germanium
Hf	Hafnium
O	Oxygen

Se   Selenium

V   Volts

V_{app} Applied Voltage

W   Watts