# *5G/6G Projects*

## Accelerated Millimeter Wave (mmW) Phased Array Prototyping (AmmP3) - $15.7M, Lead: CA DREAMS-University of Southern California (USC) and Northrop Grumman-

California Defense Ready Electronics and Microdevices Superhub (CA DREAMS) Accelerated mmW Phased-Array Prototyping Project (AmmP3) aims to develop a hub-centered mmW phased-array toolkit, accessible through CA DREAMS MOSIS 2.0, to accelerate prototyping and reduce barriers to mmW adoption. Currently, DoD and commercial markets are not fully utilizing the available mmW spectrum due to the limited availability of requisite phased-array technology. CA DREAMS AmmP3 focuses on the development of 5G/6G-relevant prototypes to accelerate availability of high-performance front ends, including phased array antennas, beamforming ICs, and broadband amplifiers, integrated using SOTA advanced (2.5D and 3D heterogeneous integration [HI]) packaging for leading-edge size, weight, power, and SWaP-C systems. By exercising the CA DREAMS operational model, AmmP3 will demonstrate process-module, component, and system-relevant prototypes for dual-use applications that will be made available for users/collaborators through the MOSIS 2.0 Storefront.

## Wideband Multifunctional Software Defined Radio (WMSDR) - $8M, Lead MMEC -Lockheed Martin-

The Wideband Multifunction Software Defined Radio (WMSDR) project seeks to provide a 5G/6G solution for adaptive wideband transceiver architectures and components. Adaptive wide band, reconfigurable, multifunction RF transceivers that meet DoD 5G/6G needs are neither readily available, nor produced in the U.S. Current approaches require *multiple* radio solutions that are narrow-band, high-SWaP-C, and negatively impact performance/effectivity and sacrifice flexibility/resistance to interference. This results in dedicated solutions which lack flexibility for multifunction. Transceiver instantaneous bandwidths are limited to ~1GHz and consume significant power for performance. DoD has unique reliability, environmental, security, and other non-foreign source needs for 5G. These constraints limit SWaP-C such as: (1) multiple nonreconfigurable fixed narrow band filters to address spurious non-linearities/interferers, (2) high-power consumption when increasing bandwidth and sample rates (affecting Nyquist bandwidth) due to current quantizer solutions for the Rx/Tx functions, (3) additional filtering required when incorporating mixers to convert microwave signals into a near zero IF band for digitization, (4) high-power consumption and data bandwidth when moving high speed quantized digital data between processors, (5) large SWaP-C to achieve increased sampling rates and wider instantaneous Nyquist bandwidths through low phase noise clock sources. WMSDR will enable a low power small form factor software define radio enabling: (1) simultaneous real-time spectral wideband field of view, (2) chip-scale wideband frontend with self-tuning filters for interferer mitigation and performance optimization, (3) electronical steerable array integrated with advanced Intel transceiver with 16 Rx/Tx channels for beam steering and RF MIMO solutions and is designed to scale to higher element counts (4) real-time, low-latency DSP array leveraging a massively parallel architecture for high instantaneous bandwidth and low power. The domestically produced WMSDR solution will combine a Wideband Adaptive RF Front End Assembly (WAF) with an Intel 18A-based transceiver designed for 5G/6G application in the FR1, FR2 and FR3 bands. This concept will enable a low SWaP, highly integrated, wideband Software Defined Radio that can successfully operate in contested and congested environments. The WMSDR project is led by Lockheed Martin who will act as prime integrator for the low-SWaP, SOTA, autonomous controlled RF front end with the reconfigurable transceiver prototype. Intel will develop, fabricate, and deliver a new wideband reconfigurable transceiver built on 18A technology. The University of Dayton and CapV will design Barium Strontium Titanate (BST) varactor tuning elements. Indiana Microelectronics will design Wideband Adaptive RF Front End Assembly (WFA) filters based on the BST varactors and Purdue University will design a new ASIC autonomous filter controller to be fabricated at Skywater. 3D Glass will design and deliver a Heterogenous Integrated (HI) filter package incorporating WFA filters and filter controller ASIC. This effort is expected to mature heterogeneous RF reconfigurable filters and multi-function transceiver architecture from a current TRL/MRL of 4 to a TRL/MRL of 6 by the end of the project.

## Wideband, Scalable multiple-input-multiple-output (MIMO) arrays for NextG Systems: From Antennas to Decoders (NextG) - $6M, Lead: NEMC-Sivers Semiconductors -

Frequency Range 3 (FR3) allocations (spanning 7.125 GHz to 24.25 GHz) and established DoD bands require low SWaP-C wideband radios from 7GHz to 15GHz. Since commercially available SOTA decoders are limited, flexible universal decoders that enable extensive code mechanisms and on-the-fly code-switching would greatly augment the complexity of embedded security. Cellular networks currently use Frequency Range 1 and Frequency Range 2 bands (FR1 belongs to Sub6GHz range of 450 - 6GHz and FR2 belongs to mm-wave 24.2 -52.6 GHz). There is now a significant uptake of 5G Massive-MIMO solutions (M-MIMO) is a reference to massive multiple-input multiple-output wireless communications technology in which base stations are equipped with a very large number of antenna elements to improve spectral and energy efficiency) globally for Frequency Range 1, and it is expected that the M-MIMO share of the network will continue to grow, motivated by increased network coverage, capacity, and user throughput. Despite increased capacity with M-MIMO in FR1 bands, additional spectrum will be required to support traffic growth and demand for new services. This project aims to develop scalable wideband transmit/receive MIMO systems, including 7GHz to 24 GHz wideband MIMO antennas, ICs, and novel decoder back-ends for 5G/6G and DoD deployments. At the end of this project, the goal is to provide engineering samples to potential 5G and DoD customers and prepare for volume production in GF 22nm FDX for MIMO TRX ICs. The project also aims to provide engineering samples for the MIMO array tile, expect to achieve a TRL of 6 for the scalable 256-element 32 T/R MIMO tile, and engage with assembly houses capable of volume production. This project will translate IC and antenna technology developed at Sivers and UML into commercial and DoD application-ready prototypes.

## Linearized Frequency Multiplier Transmitter Technology Transition (LIFTTT) - $2.2 M, 5G/6G Lead: ARL/AFRl-

project will expand warfighter capability to utilize mmW spectrum for communications and EW applications. ARL and AFRL will collaborate on (1 fabricating a packaged, micro-scale transmitter array with throughput beyond SoA (ARL IP), reducing its SWaP-C and (2 extending current capability above 100 GHz to cover 6G and EW applications in the D-band, providing a technology solution where commercial-off-the-shelf equipment cannot meet transmitter requirements.

## Integrated Radio Frequency (RF) GaN Technology to Support NextG Wireless Systems (Integrated RF GaN) - $5.1M, Lead: Southwest Advanced Prototyping hub (SWAP)-NXP -

The Integrated RF GaN Technology program aims to achieve significant advancements in efficiency, bandwidth, and size of power amplifiers for 5G/6G and DoD applications. This will be achieved through innovative power amplifier and built-in self-test architectures, utilizing package and wafer level hetero-integration. Current wireless infrastructure equipment, limited by sub-6 GHz spectrum allocations, uses a mix of amplifier technologies, including LDMOS, SiGe, and GaN. These radios often use a banded approach with specific amplifiers for optimal efficiency and performance within each band. Additionally, current infrastructure relies on conventional chip and wire packaging and assembly. 5G/6G infrastructure requires higher frequencies, wider bandwidths, and smaller amplifier form factors. Higher frequencies necessitate increased efficiency to mitigate self-heating, and GaN technology enables wideband operation with increased instantaneous bandwidths. However, GaN efficiency is limited as bandwidths exceed 400 MHz. This project addresses these challenges by combining reconfigurable/built-in self-test power amplifier architectures with heterogeneous integration of GaN and CMOS technologies. This will be the first implementation of reconfigurable GaN power amplifiers with built-in-self test hardware and algorithms, using heterogeneous integration of GaN on Si in commercially scalable module/hetero-integration technologies. A successful outcome will establish new benchmarks for reconfigurable power amplifier integration and built-in self-test hardware and algorithms, leading to more efficient and compact MIMO arrays with lower cost and energy usage. The project will provide essential learnings and prototypes for lab-to-fab migration of these technologies, extending U.S. technology leadership in 5G/6G and benefiting DoD systems with stringent SWaP-C requirements.

## Scalable Modular Architecture for RF Transceivers (SMART) - $5.7M, Lead: SWAP-Alphacore-

A team led by Alphacore Inc. is developing a ground-breaking Joint Communication and Sensing (JCAS) transceiver for use in 5G/6G telecommunications, radar, and related applications for both Defense and Commercial systems. This project will lead to an IC that performs at unprecedented levels of capability, compactness, low power, and economy. Working under the auspices of the Arizona State University-led SWAP, Alphacore will integrate its own *Scalable Modular Architecture for RF Transceivers* (SMART) technology with cutting-edge contributions from its partners at Arizona State University, Rice University, Auburn University, and Lockheed Martin. Key benefits of the JCAS transceiver include increased resource efficiency, optimized spectrum utilization, reduced latency and improved reliability while simultaneously reducing the size and cost of transceivers. The IC, which will reach TRL 7/MRL 7 by the end of the program, will meet the challenges of robust operation in both terrestrial and space-based systems.

# *Secure Edge/Internet of Things (SE IoT) Projects*

## Validated GPU Based Secure Processing Module (GUARD)- $4.8M, Lead: MMEC-Northrop Grumman-

The Validated GPU Based Secure Processing Module (GUARD) project seeks to provide a solution for defense in diversity for IoT systems. The processing needs envisioned to enable the capabilities our future warfighters need at the tactical edge far exceeds what is available today and forward deployed tactical edge solutions demand a security posture significantly beyond what is commercially supportable. The GUARD project will develop a secure graphics processing unit prototype that is based on mature commercial GPU technology and proven protection concepts to enable a forward-looking architecture for a family of secure edge GPU ecosystems for both DoD and commercial applications. A low SWaP secure GPU module will enable users to rapidly deploy and update emerging AI algorithms to promote an asymmetric advantage in AI and sensor processing. GUARD will advance the technologies needed to significantly reduce SWaP-C required for future edge computing assets, develop a holistic security posture to protect the control flow of the processor from external influence, ensure the integrity of the system during execution, prevent exfiltration of critical program information, and provide a means to recover from and/or respond to threat events. The effort will achieve these goals by developing the architecture to pair proven root of security (RoS) technology with GPU specific enhancements and techniques for mitigating boot time side channel attacks with commercial high performance GPU devices from NVIDIA. NVIDIA recognizes the need for secure computing solutions in both the datacenter and the edge to protect sensitive data from unauthorized access. As a result, NVIDIA has implemented software and hardware solutions such as confidential compute, trusted execution environments, inline memory encryption, and secure boot. The integration of the NVIDIA GPU with a root of security will protect data at-rest and make the determination of a safe environment before enabling the GPU to process critical information. The GUARD team is led by Northrop Grumman which is responsible for overall prototype design, proof of concept demonstration, and testing. The University of Maryland supports the effort with vulnerability analysis and Battelle with security modeling. Future hardware integration will be performed by NHanced. Commercial GPUs are being sourced from NVIDIA. The GUARD project expects to mature the Root of Security integration with NVIDIA's GPU ecosystem from a TRL/MRL of 2-3 to a TRL/MRL of 5.

## Modular Digital Direct Waveform Synthesizer w/ Integrated Analog-to-Digital Converter (ADC) Functionality (DDWS) - $9.2M, Lead: SCMC -HII Mission Technologies-

Frequency-agile analog synthesizers or digital direct synthesis (DDS) components today are integrated with other components to support DoD needs in expensive subsystems. DDS drawbacks include phase noise, limited-precision tuning words, dependence on clock quality, complex schemas to overcome limitations, and significant power consumption and size. This project seeks to develop a broadband Digital Direct Waveform Synthesizer (DDWS) that is highly programmable, offering high-speed communication links, direct data acquisition for flexible processing, secure storage, and the ability to create a root-of-trust processing environment. Successful development of the end-state, TRL-8 Multi-Chip Module (MCM) enables a scalable digital signal processor environment featuring secure computing, compliance with zero-trust and root-of-trust, and support for AI/ML applications at the edge, all with the lowest overall footprint (reduced SWaP, increased efficiency, improved thermal management and signature control, intermittent independence from cloud). It will provide a ruggedized system for cognitive radar/EW with AI/ML and high-speed signal processing. The MCM with DDWS facilitates sensor integration into smaller platforms (autonomous vehicles, small drones, buoys) and empowers ISR and EW systems with new capabilities. The DDWS project will benefit from SCMC hub-supported EDA access and expanding into dual-use markets/applications, as this system is also currently being evaluated for integration into 5G communications systems/base stations. HII is engaging SCMC for continuing hub support for additional, externally funded projects to build off DDWS.

## ARC-V Secure Processor (ARC-V) - $7.8M, Lead: SWAP-Idaho Scientific-

Idaho Scientific is developing a security-hardened processor ASIC chiplet over a three (3) year program. Developing highly secure systems that can protect important assets from cyberattacks starts at the silicon level. Idaho Scientific has established an agreement to license Synopsys' ARC-V™ Processor IP and integrate it with Idaho Scientific's Helios Memory Guard (HMG) technology. This combination creates a trusted processing solution that addresses the stringent performance and security requirements of SoCs targeting aerospace and defense applications. The project's primary objective is to create a secure, low-power processor chiplet that allows companies to confidently deploy advanced systems, which rely on powerful but vulnerable commercial electronics. While this initiative is initially focused on Aerospace and Defense use, the technology has far-reaching implications for the commercial sector. When the secure processor chiplet becomes available for commercial use, it will be a game-changer for industries that depend on confidential computing and protecting sensitive data, such as data servers, healthcare, banking, and telecommunications. Hackers can exploit imperfect software and weak system security to take control of devices or reverse-engineer them, threatening national security. The Helios Memory Guard technology that will be integrated with the Synopsys ARC-V processor prevents a broad range of threats, including code injection attacks, Call-oriented programming, jump-oriented programming, planting foreign code in or tampering with boot, memory corruption attacks on non-owned memory, cold boot or data remanence attacks, and reading or altering memory via Input/Output, Direct Memory Access or Rowhammer-style attacks. Year One (1) is focused on front end design and demonstration in an FPGA simulation.

## An Efficient Hardware Crypto Engine based on In-Memory Computing (IMCRYPTO) - $2.8M, Lead: SCMC -University of Notre Dame-

Current commercial offerings of widely adopted AES cryptographic engines, associated with von Neumann processing hardware such as CPUs, require numerous clock cycles to initiate via Round Key and Key Scheduling operations involving firmware, software, storage and logic. Furthermore, limitations such as subsequent cryptographic operations involving significant data transfers require overhead of hundreds of sequential instructions, and information radiated from data transfer paths during boot sequences require additional technologies to limit transmission vulnerabilities. Sensor systems on missile platforms have historically been processor limited—the volume of sensor data (e.g., EO/IR or RF sensors)—exceedingly outpaces the means of von Neumann architectures to extract sufficient information; thus, system requirements are strained (i.e., the acquisition range can be limited by a constrained number of track files due to hardware bandwidth, latency and power consumption). The goal of this project seeks to address these limitations with the use of a non-von Neumann architecture. IMCRYPTO allows for data to be computed where it resides—in memory. IMCRYPTO leverages the unique features of In-Memory Compute to (1 overcome the inherent data transfer bottleneck in typical crypto engines, (2 offer high parallelism, (3 support multiple block ciphers, (4 achieve encryption/decryption within a single unit, and (5 reduce information leakage. The end state of this project is a TRL-6 ASIC IP that provides lower SWaP and more secure crypto operations than what is available today. Potential performance impacts include,

1) 5X increase in throughput and 60X reduction in energy, for the popular AES-128 encryption
2) Will support multiple cryptographic algorithms including SHA-256, MD5, and DES
3) Will provide more resilience to power and timing side channel attacks than ASIC-based AES accelerators

IMCRYPTO will leverage Commons supported Core/Foundry partners in Global Foundries and SCMC - supported EDA access to aid in their development and prototyping efforts

### Porous Silicon Particle-Based Ink for Device Protection (PoSiP)- $0.60M Lead: ARL-

Current anti-tamper countermeasures rely on software solutions, and there is a need for physical destruction of electronics. ARL has developed reactive porous silicon (pSi) technology as a physical anti-tamper solution, in the form of a printable ink. This effort is focused on improving the ink in 3 areas: printability, long shelf-life packaging, and consistent effects.

### State of the Art Microelectronics Security Evaluations (SAMSE)- $0.99M Lead: Army-

Army strives to increase the Department of Defense's readiness for evaluating next generation microelectronics. To this end, the project is developing a next generation device survey describing current SOTA for heterogeneously integrated/advanced packaged (HI/AP) microelectronics designed for secure edge processing, which includes security features, implementations, and known open-source vulnerabilities for these components. In addition, the effort compiles available analytical tools and testing techniques necessary for HI/AP microelectronics when performing security evaluations for selected SOTA microelectronics components.

## *Artificial Intelligence Hardware (AI HW) Projects*

### Energy-Efficient and Scalable AI HW Systems through Heterogeneous Integration of Specialized Chiplets (ScaleAI) - $6.7M, Lead: NW AI - Stanford-

This project will drastically improve energy consumption and performance of AI HW systems for DoD applications, reaching the TRL 6 and the MRL 7 upon completion. Existing AI systems use large off-chip memories and spend enormous time and energy shuttling data back and forth between compute and memory chips—the memory wall. This memory wall gets worse as traditional 2D scaling via miniaturization gets increasingly difficult—the miniaturization wall. The deadly combination of the memory wall and the miniaturization wall severely limits the capabilities of existing AI HW for DoD applications such as the Connected Battlespace. This project will use innovations in semiconductor materials, monolithic 3D integration technologies, and AI system architectures to enable new AI capabilities through a dramatic 100-fold improvement in system-level energy efficiency compared to current commercial solutions (GPUs/CPUs). A variety of interconnected, heterogeneous, AI-specialized chiplets, built using leading-edge CMOS and 3D CMOS+X semiconductor technologies such as carbon nanotube transistors, resistive memory, and oxide semiconductors, form the foundation of the overall approach. The transition plans include: 1) AI HW transition via the DIB and 2) establishment of a new 3D CMOS+X technology in a U.S. foundry—a world-first capability—for sustained U.S. leadership in semiconductor innovation. Thus, this project will directly advance the Commons' mission of lab-to-fab translation of critical technologies.

## Lab-to-Fab Transfer of complementary metal oxide semiconductor (CMOS) +memristor Chips for Edge Intelligence (Edge-Intel) - $7.9M, Lead: NEMC-University of Massachusetts, Amherst-

This project aims to transfer analog memristor technology from research labs to commercial foundries, creating a thriving AI HW ecosystem for DoD mission-relevant applications. Energy and latency issues are acute at the network's edge and on autonomous and wearable platforms where extreme energy efficiency is mandatory for DoD applications. Commercial AI HW purely based on CMOS does not address these issues and faces challenges in terms of SWaP. Physical computing with analog memristors has been demonstrated as a promising solution as it avoids constant data shuttling between memory and compute chips, and instead performs analog computing where the data is stored using the laws of physics, substantially reducing latency and energy consumption. The greatest obstacle to adapting this technology to real-world applications is the lack of an analog memristor technology in the IC industry. In this project, multi-level memristor devices and processes from universities and startup labs will be transferred to the industrial fabs and integrated with current CMOS technology, offering a commercially viable analog memristor device and system technology on U.S. soil at TRL 7 by the end of the project. Our project will enable compact, fast, intelligent electronic systems for unmanned aerial vehicles, autonomous sensing systems, hypersonic weapons, etc. The AI chips will have substantial SWaP advantages and will address challenges in data-to-decision and autonomy requirements in DoD missions. The proposed project fully engages the NEMC ecosystem and leverages local and national facilities and infrastructure, including academia, startups, defense contractors, R&D, and commercial foundries. By bringing each team member's expertise together, we will enhance the NEMC capabilities in AI HW development and implement the mission of the Commons in lab-to-fab transfer.

## Energy-Efficient, Scalable, and Self-Learning AI HW with 3D Electronic-Photonic-Integrated-Circuits (3DEPIC_AI) - $5.7M, Lead: Northwest-Artificial Intelligence hub (NW AI) -UC Davis-

The project "Energy-Efficient and Scalable, and Self-learning AI HW with 3D Electronic-Photonic-Integrated-Circuits (3D EPIC-AI)" addresses the significant limitations of contemporary AI HW solutions reliant on GPU, CPU, and traditional memory technologies such as HBM, DRAM, and SRAM. By taking the 'best of both worlds' approach, the project innovatively combines photonics and electronics into a compact and advanced 3D integrated circuit module, this initiative aims to create a new class of AI HW that effectively overcomes the innate shortcomings of existing architectures. Today's AI HW systems are hindered by substantial drawbacks, including excessive power consumption, bulkiness, weight, and inadequate processing speed—all resulting from limitations inherent in electronic processing and the von Neumann bottleneck. Furthermore, these systems lack the scalability and adaptability to seamlessly learn new AI models. The proposed innovations in 3D photonic-electronic integrated circuits (3D EPIC) present an unprecedented opportunity to enhance energy efficiency, throughput, and self-learning capabilities, facilitating the development of commercially viable AI HW made in the United States. Key advancements within this project revolve around the design of a novel photonic-electronic AI neural network architecture that incorporates cutting-edge technologies, such as silicon photonics, III-V photonics, Resistive Random Access Memory (RRAM) crossbar arrays, and 3D photonic-electronic integration techniques. The forecasted results promise dramatic improvements in both throughput and power efficiency on a resilient 3D platform, alongside the implementation of self-learning capabilities that significantly enhance learning speeds. The implications of this technology for the DoD are extensive and multifaceted, supporting a wide array of platforms including Unmanned Aerial Vehicles (UAVs), robotics, soldiers in the field, missiles, and naval vessels, among others. With the anticipated performance characteristics of beyond 1 Peta Operations Per Second (POPS), exceeding 10 Tera Operations Per Second per Watt (TOPS/W), and a self-learning capability at the edge, the project targets a solution with more than 10x throughput power density compared to current systems. Achievement of these ambitious goals entails successful transitions from laboratory to fabrication readiness (TRL5/MRL5), demanding engagement with U.S.-based manufacturing foundries. The lab-to-fab transition will leverage collaborative efforts with NW-AI partners, encompassing innovative prototyping in the areas of photonics, electronics, RRAM, and 3D electronic-photonic circuits, ultimately guiding these advancements toward commercial manufacturing at specialized foundries partnered with DoD labs and contractors. Through 3D EPIC-AI, we aim to redefine the landscape of AI HW, delivering systems that are not only high-performing but also significantly more energy-efficient, compact, and capable of evolving in real time to meet the pressing demands of modern defense applications.

## CMOS+X: Integrated Ferroelectric Technologies for Ultra Efficient AI HW (FerrAI) - $4M, Lead: NW AI -UC Berkeley-

This project aims to substantially better energy efficiency for future AI HW compared to SOTA. Our team includes UC Berkeley as the lead and Stanford University and Georgia Institute of technology among the academic partners. In addition, the team includes MIT-LL who will be primary fab for lab-to-fab transition, and Intel who will drive system level functions. Further, we have Northrop Grumman Corporation and Raytheon Technologies who will drive defense related applications. The project will exploit unique properties of Ferroelectric materials to lower the power supply voltage of computing hardware as well as achieve non-volatile memory that can be directly integrated with the microprocessor. As the AI workload requires highly memory-centric and parallel operations, low voltage operation together with access to tightly integrated large memory can have a multiplicative effect on the energy efficiency of the AI HW. The impact on future defense systems can be enormous. Essentially, with significantly improved energy efficiency, it will be possible to process data at the edge improving efficiency and speed of decision making, leading to improved safety for defense use cases.

## Fab-Accessible BRain-Inspired Chips for Sensors (FABRICS) - $2.1M, Lead: Naval Research Laboratory (NRL)-

FABRICS aims to develop ferroelectric devices made from materials that are compatible with FEOL and BEOL fab processes and that implement synapse- and neuron-like functions for neuromorphic computing applications. To this end, a novel rapid laser annealing process is used to transform hafnium- and zirconium-oxides into crystalline phases that exhibit strong ferroelectric polarization and are stabilized for their final device environment. Simulations of processing conditions and of the device physics, derived from the characterization of prototype devices we are fabricating, will be employed to generate process flows and circuit designs to enable full PDK support for ultimate fab integration of these devices.

## CMOS+MRAM Hardware for Energy-EfficienT AI (CHEETA) - $8.7M, Lead: SCMC -Purdue University-

The current landscape of edge AI HW is dominated by mobile SoCs, GPUs and digital accelerators. These platforms achieve processing efficiencies in the range of 1-10 TOPS/W. However, they are all limited by the time and energy required to move data between processing units and memory (the von Neumann bottleneck) and the large leakage of on-chip memories. CHEETA seeks to leverage CMOS+MRAM to address the challenges. To alleviate the von Neumann bottleneck, CHEETA will deploy in-memory computing (IMC) primitives that leverage the non-volatility of MRAM to realize Matrix-Vector Multiplications (MVMs) – the key computational kernel in AI workloads – in situ within MRAM arrays. In addition, the design and fabrication of IMC-Plus primitives will extend the capabilities of the basic IMC primitives in two important ways: 1) ROM-overlays in order to realize operations beyond MVMs such as non-linear activations, Softmax, etc., and 2) stochastic neurons that replace area- and power-demanding ADCs in order to improve the efficiency of IMC. The proposed CMOS+MRAM hardware primitives will enable a new generation of energy-efficient AI HW that push the boundaries of the current state-of-the-art. The end state of CHEETA seeks to develop a TRL-6 technology that drives greater than 100X improvement in energy efficiency compared to current AI HW. Furthermore, the high endurance of MRAM compared to other NVM technologies and the intrinsically radiation-tolerant nature of MRAM will offer compelling advantages in DoD applications. This project is leveraging the SCMC lab-to-fab transition by the incorporation of the Honeywell radiation-hardened CMOS FEOL, Everspin BEOL, with an end-state transition partner in Northrop Grumman. CHEETA will also be provided with SCMC-supported EDA access during the development process.

## Spaceborne Low-Energy AI Computing (SLEAC) - $6M, Lead: SWAP-Arizona State University –

Spaceborne Low Energy AI Computing (SLEAC) aims to extend the power of AI to satellites orbiting our planet. Advancing the performance of satellites through AI has the potential to unlock advances that will yield a major advantage for national defense. Computer hardware designed for spacecraft undergoes a manufacturing process called radiation hardening. It enables the equipment to survive exposure to harsh radiation levels in orbit, but it also compromises performance compared to state-of-the-art computing on Earth. If it were possible to directly integrate a highly efficient, radiation hard AI chip with focal plane array image sensors used in space, satellites could track objects that are too faint or too fast to be detected by current systems. SLEAC project will overcome this challenge by using radiation hard metal oxide-based analog-optimized resistive memory (ReRAM) arrays integrated with foundry-supplied CMOS to create an AIMC prototype processor capable of more than 10 TOPS/W in extreme environments. This would be several times more efficient than modern unhardened systems and over 100 times more efficient than the current state of the art radiation hard systems. ReRAM is a robust form of nonvolatile memory that is being developed at Arizona State University, University of Southern California and Sandia National Laboratories starting from a Sandia baseline process. The project will also leverage resistive switching materials and devices developed for extreme environments at the Center of Neuromorphic Computing under Extreme Environment (CONCRETE) sponsored by the Air Force. It is being optimized and electrically conditioned to have near-ideal properties for AIMC, including stable resistance tunability with a high dynamic range and extremely low noise, making it ideal for scaling to large arrays with high accuracy and efficiency for AI computing under extreme environments. SLEAC plans to integrate ReRAM on the back-end-of-line of commercial radiation tolerant CMOS wafers, using Arizona State University's recently established SWAP capabilities. This will enable demonstration of large-scale ReRAM arrays with high yield, endurance, and low variability, validated across millions of devices. Ultimately this technology will enable demonstration of a radiation hard spaceborne remote sensing systems capable observing phenomena that are currently hidden.

## RAVENS 2 chip Demonstration platforms for Deployment (R2D2) - $1.9M, Lead: NRL-

The RAVENS 2 chip Demonstration platforms for Deployment (R2D2) program aims to demonstrate advanced performance in AI tasks through advanced packaging and systems-level integration of the RAVENS2 ASICs developed under the NeuroPipe ARAP into prototypes of applied systems such as a sensory data processors and autonomous vehicle navigation systems. These will be tested at the respective laboratories and warfare centers to demonstrate advances in performance and reductions in SWaP above the established baseline of existing AI HW. The effort will demonstrate the feasibility of memristor-based circuits to dramatically shift computational loads associated with cognitive functions away from digital hardware towards fully analog neuromorphic devices.

## Ultra Efficient In-Hardware Prototype Using Hyperdimensional Computing (ENERGY) - $1.6M, Lead: MMEC-Intel Federal-

The Ultra Efficient In-Hardware Learning Prototype Using Hyperdimensional Computing (ENERGY) seeks to provide an AI HW solution for in-hardware learning. AI proliferation faces a number of technical challenges. Current AI HW is large and power hungry which negatively impacts SWaP. In addition, SOTA implementations are sample/training inefficient and susceptible to noise and perturbations. The ENERGY project will develop a hardware prototype that is capable of learning in-hardware directly from the raw image and video data in an incremental fashion that is SWaP efficient in noisy and dynamically changing environments. The effort will build on a breakthrough neural architecture based on Hyper-Dimensional Computing (HDC) which has been developed that enables revolutionary few-shot learning capabilities. HDC, a non-von Neumann computing framework that mimics human memorization and cognition, uses a high-dimensional representation of data called hypervectors (> 1 Kbits), that mimic the representation in the brain which are robust, holistic, and sparse. Recent results have shown that good, few-shot learning performance does NOT require pretraining on a large dataset of labeled examples if the neural representations are defined by tight geometric regions in high dimensions akin to higher-level brain areas. To obtain data hypervectors from complex images for image and video processing, a novel and compact unsupervised/supervised image segmentation module inspired by the early stages of visual processing in the brain will be developed. Hypervectors or symbols for each region of interest will be extracted using streaming encoding techniques. This data-to-symbol converter will be trained offline once with image data in an unsupervised/supervised fashion and then compressed for low SWaP requirements. Initial algorithms will be demonstrated in an FPGA. A longer-term goal is to combine an FPGA with a ferroelectric diode array tailor-made for compute-in-memory to greatly reduce sample, training, and power demand compared to CPU/GPU implementation. The ENERGY project is led by Intel who is responsible for AI and hyperdimensional computing models, FPGA implementation, and program management. AI-Sensation is working with Intel on overall design of algorithm hyperdimensional computing algorithms for joint learning and reasoning. Future activity will include University of Pennsylvania who will work on ferroelectric diode-based compute-in-memory. The AFRL and the ARL will support characterization of the ferroelectric devices. ENERGY is expected to mature Hyperdimensional Computing from a TRL/MRL of 4 to enable breakthrough neural architectures of a TRL/MRL 7.