We propose optimal design methods of 3D NAND Flash to achieve high-density, high-bandwidth and low-power nvCIM. The proposed 3D NAND nvCIM can produce TOPS/W ~40 (4I4W), and achieve 74.6 frame/sec for a heavy VGG16 network. It is potential to provide high-density and low-power accelerator for deep-learning neural network.

The storage reliability of multi-bit flash is of vital importance for the flash based deep neural network (DNN). In this work, the critical concerns correlated with the storage reliability (I_d) distribution and data retention) of multi-bit flash and its impacts on the DNN are investigated for the first time.

A convolutional neural network (CNN) core that can be readily mapped to a 3D NAND flash array was demonstrated in a standard 65nm CMOS process. Logic-compatible embedded flash memory cells were used for storing multi-level synaptic weights while a bit-serial architecture enables 8 bit multiply-and-accumulate operation.

A novel synaptic architecture based on NAND cell strings is proposed as a high-density synapse capable of XNOR operation for binary neural networks (BNNs) for the first time. By changing the threshold voltage of NAND flash cells and input voltages in complementary fashion, the XNOR operation is successfully demonstrated.

A complementary memory cell based on field-programmable ferroelectric diode for ultra-low power current-SA free BNN applications.
We proposed a complementary memory cell with 2T4D XNOR structure for computing in memory (CIM) applications. Firstly, a field-programmable diode (FPD) was demonstrated with a W/HZO/W structure. For BNN application, a 2T4D XNOR cell-based CSA-free BNN architecture is proposed with small cell area (16 F²) and superior efficiency (387 TOPS/W).

3:40 PM  38.6  A 2TnC Ferroelectric Memory Gain Cell Suitable for Compute-in-memory and Neuromorphic Application
Stefan Slesazeck, Taras Ravsher, Viktor Havel, Evelyn Breyer, Halid Mulaosmanovic, Thomas Mikolajick, NaMLab gGmbH, TU Dresden

A 2TnC ferroelectric memory gain cell is proposed, that can be operated either in FeRAM or FTJ mode. The internal gain of the cell mitigates the need for 3D integration of the FeCAPs, making the concept very attractive for embedded memories, compute-in-memory and neuromorphic applications.

4:05 PM  38.7  Memory-Logic Hybrid Gate with 3D-Stackable Complementary Latches for FinFET-based Neural Networks
Chieh Lee, Yue-Der Chih, Jonathan Chang, Chrong Jung Lin, Ya-Chin King, National Tsing Hua University, Taiwan Semiconductor Manufacturing Company

A memory-logic hybrid gate with complementary resistive switching pairs on vias in BEOL FinFET technologies with an area-efficient, 3D-stackable structures is proposed. Stable output logic stages enabled by the complementary states on the RRAM pair have been demonstrated. Through stacked-vias architectures, logic operations based on multiple non-volatile states are achieved.