Persistent Memory Productization driven by AI & ML

Danny Sabour
VP Marketing, Avalanche Technology
Persistent Memory Usage from Cloud to Node

CLOUD

Big data processing
Data Warehousing

EDGE

Data Visualization
Data caching
Data filtering

IoT NODE

Realtime data processing

Compute Storage

Deep Learning

Training

Inference
Node Architecture: From Data Fusion to AI

- **The Node:**
  - Collects Data
  - Aggregates data
  - Sends data back to Edge
  - Make decisions per input
  - Learn based on local stimuli

- **Needs**
  - Persistence
  - Low Power
  - Smallest die size

**Inference: Local Decision Making**
IoT Roadmap

**Embedded IoT Memory**
- 28nm NVM MACRO 8-64Mb
- 28nm SRAM MACRO 8-64Mb
- 28nm NVM MACRO 64-256Mb
- 14nm SRAM MACRO 8-64Mb

**Discrete IoT Memory**
- Gen 1 SPI 8Mb
- Gen 2 QSPI 32Mb
- Gen 3 OSPI 256Mb

**Technology**
- 28nm NVM MACRO 8-64Mb
- 28nm SRAM MACRO 8-64Mb
- 14nm SRAM MACRO 8-64Mb
Edge/Cloud Architecture

- The Cloud and the Edge Architecture are similar, but with the possible variations in:
  - GPU Cores which can be used for Machine learning
  - PCIe Cache for performance
  - Communication to other domains in the edge for inference and learning

- Needs
  - Persistence
  - Speed

Training: Learning New Capabilities
Medium Density Persistent RAM (MD-RAM)
Universal Memory in the Cloud: NVDIMM

NVDIMM-N: Brought Persistence to DRAM
Controller + Battery + Flash

NVDIMM-P: Hides Persistent Memory Endurance
Controller + Battery + Cache

AVA-DIMM: Persistent DRAM replacement

High Density Persistent RAM (HD-RAM)
High and Mid Density Persistent RAM

- **High-Density Persistent RAM**
  - Gen 3 HD-AVRAM
    - DDR 8-32Gb
    - DDR 64Gb

- **Mid-Density Persistent RAM**
  - Gen 1 MD-AVRAM
    - DDR 1-4Gb
  - Gen 2 MD-AVRAM
    - DDR 4-8Gb
    - DDR 4-16Gb

- **Technology**
  - 1T1R
    - Regular
    - Small
    - Tiny
    - Micro
Where does MRAM fit into the picture

1 Magnetic structure

Embedded:
- High Performance
- High Temperature

Stand Alone:
- IoT
- Med Density RAM
- High Density RAM

Avalanche Specification for different MRAMs available at: www.avalanche-technology.com/products
• MRAM products are making inroads from the cloud to the IoT Node
  • It is not one product

• MRAM is unifying the Software defined world
  • Compute
  • Storage
  • Network

• Built on a Unified CPU & Memory platform based on a unified MTJ architecture for
  • Compute, Storage, Network
  • Across the Cloud, Edge, Node

• Phased product introduction has begun