Adaptive Power Optimization Techniques in NVMe® SSD

Session Name : SSDT-303-1: New Form Factors and Interfaces for SSDs
Session Date and Time : Aug 7, 2025 Santa Clara Convention Center
12:10 PM - 01:15 PM

Ram Edupuganti, Manager, Applications Engineering
Microchip Technology Inc.

Co-Authors: Shreya Jayaprakash, Anusha Kodimela, Michele Cirella, Jayanth Murthy , Pitamber Shukla, Niles Yang and Murthy Hari

Adaptive power optimization Techniques in NVMe® SSD

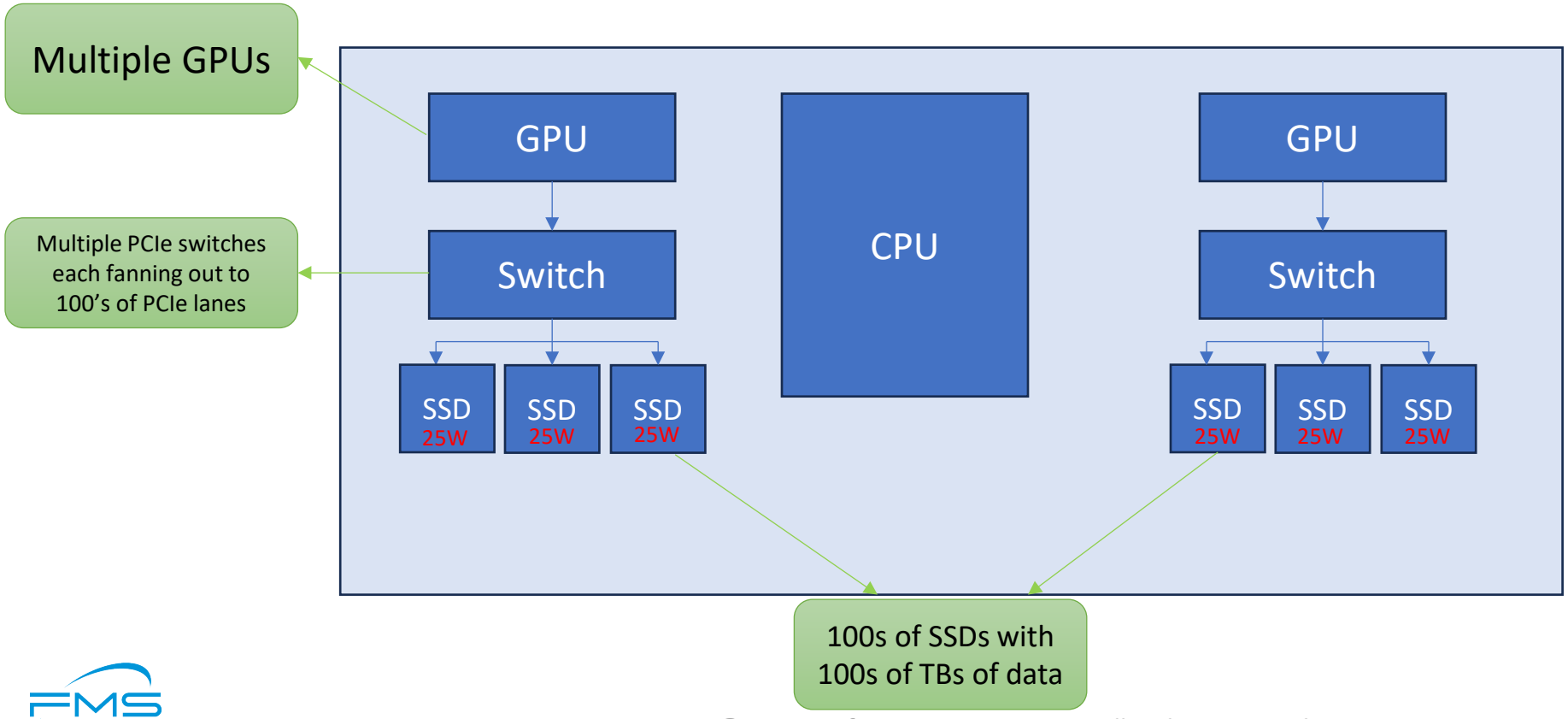
- Abstract: NVMe® SSDs have evolved from a standard passive storage device into a highly complex system with advanced SSD controller architectures which double the performance every generation, support multiple CPUs, HW accelerators, offload engines for computational storage, advanced error correction and machine learning engines in recent times. Beyond the controller, the components used in the SSDs such as SDRAM and NAND memory are also scaling in capacity and speed and consuming more power. Due to varying demands and application needs of the data center SSD, the operational workload also varies. This in turn drives the need for optimizing the power consumption of the SSDs.
- Considering the scale and lifetime of SSDs deployed in data centers, saving even a few milliwatts of power makes a big impact to TCO (Total Cost of Ownership) and the environment. It is imperative to apply innovative power optimization techniques and achieve greater savings.
- In this presentation, we will discuss various power optimization techniques that can be applied to an SSD from the design phase to the retirement phase, at the controller, component, and system levels. We will also share data on which technique provided the most cost-effective results.

Agenda

- Requirements
- What are the challenges ?
- Power optimization techniques
  - Recommended power optimization techniques
  - Advanced power optimization techniques

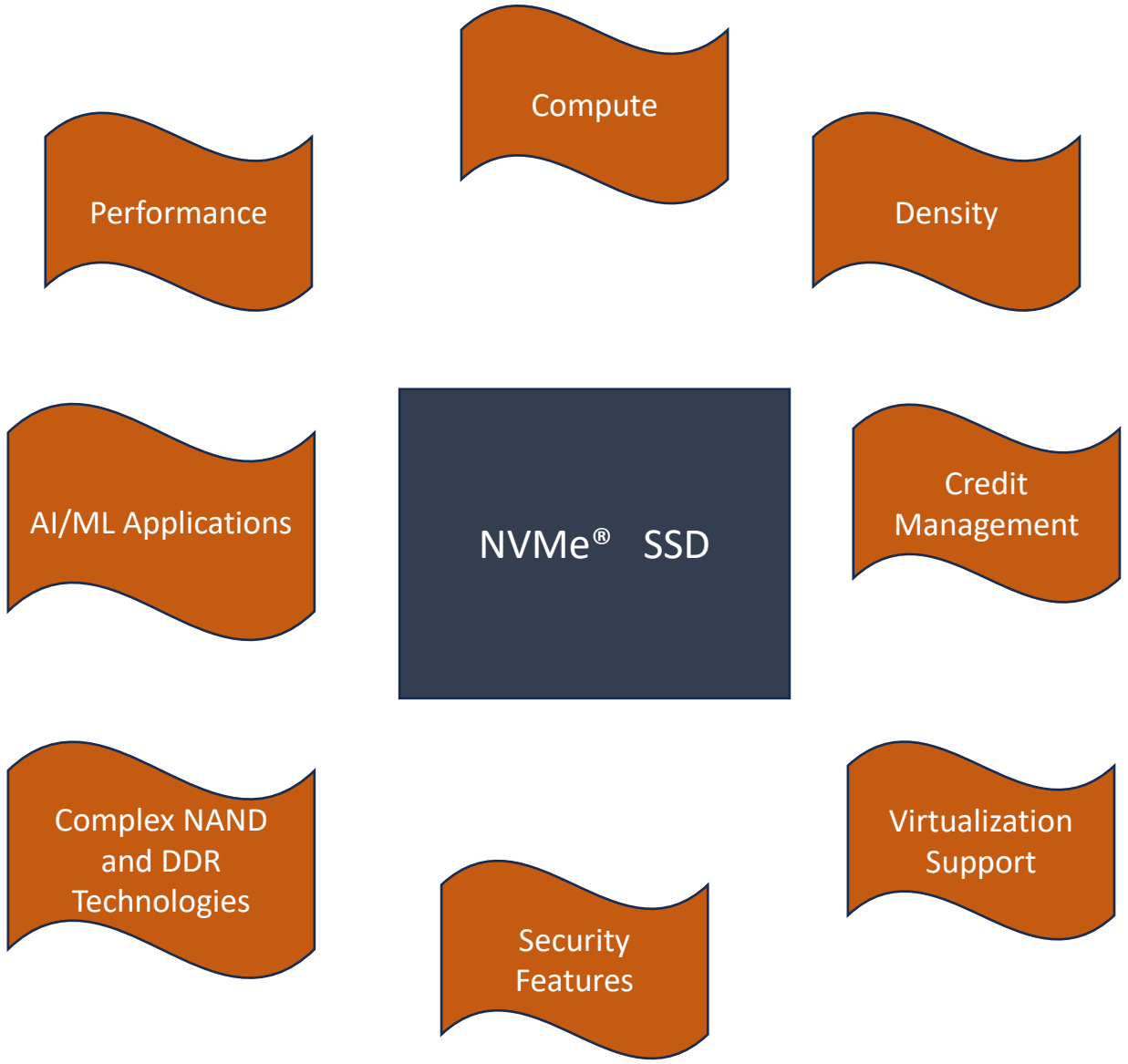
Requirements

- Active power specification varies for different form factors
  - PCIe® add-in card active power specification for SSD is 25W
  - Idle power specification for SSD is <5W, controller is much less



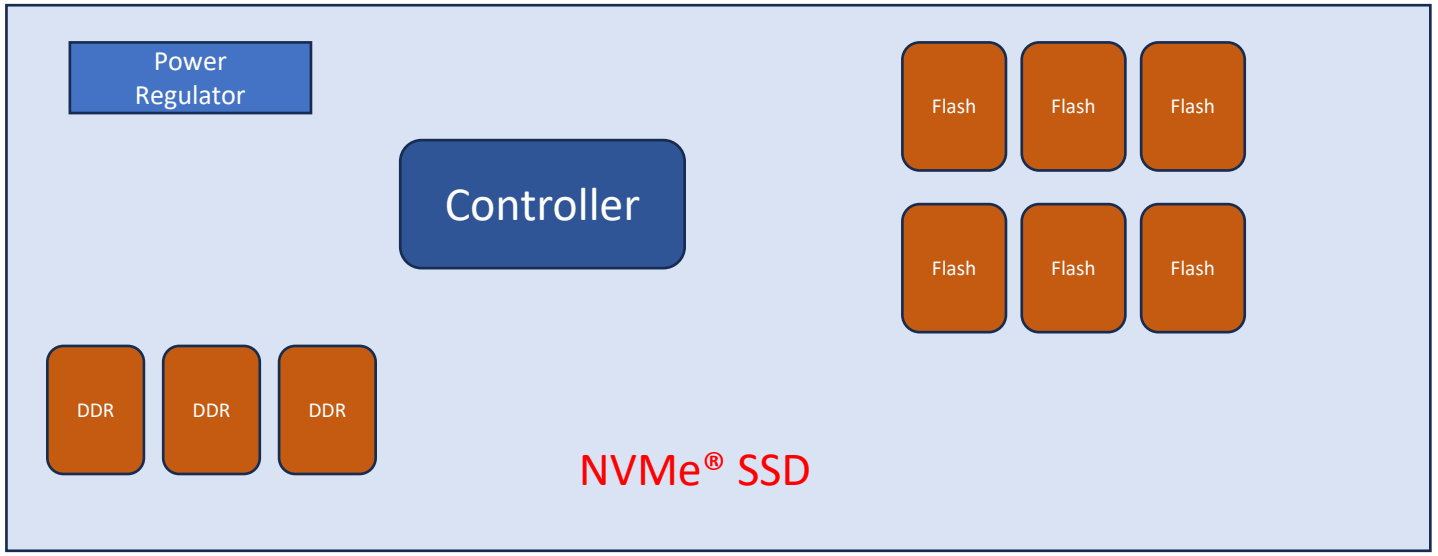
Challenges

- Some of the main challenges for power are,
  - Performance:
    - PCIe® bandwidth doubling every generation
      - From Gen-3, 8Gb/s =>16Gb/s=>32Gb/s
    - Increase in the speed of all the memory devices such as NAND and DDR
  - Density:
    - Increase in the density of the SSD, resulting in more NAND and DDR devices in the drive
      - Now 150TB+ drives are built
  - Features:
    - Support for various NAND, DDR and other devices
    - Security features
    - Credit management system for guarantee QoS
    - Virtualization support for Hyper scalars
    - AI/ML applications



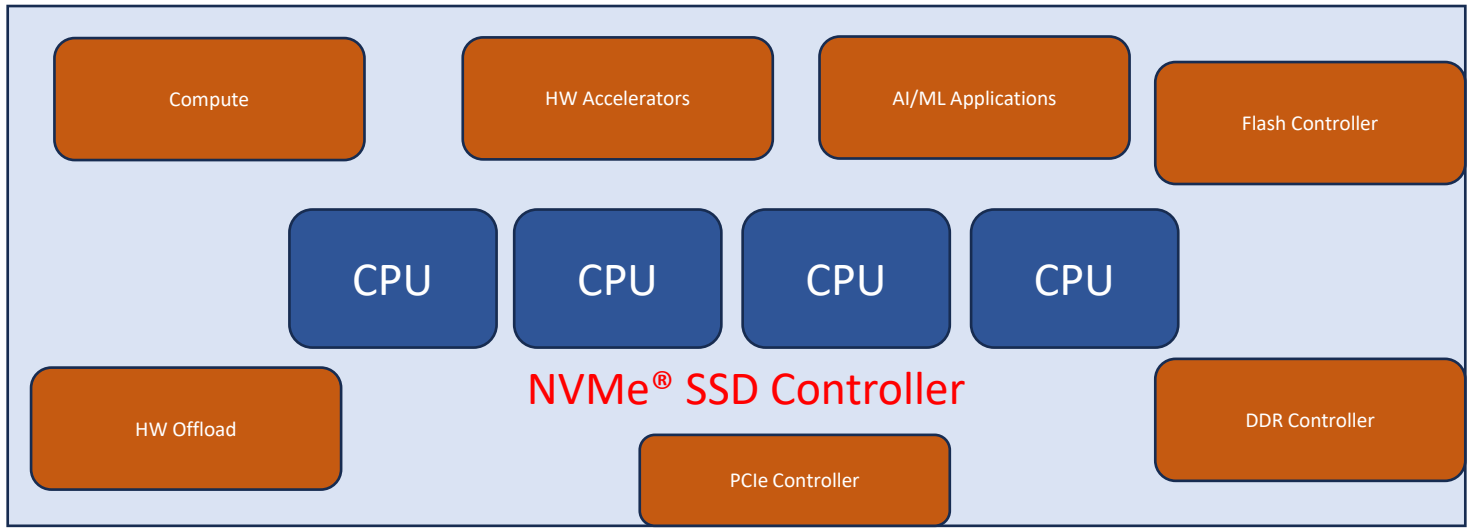
Addressing those challenges

- Power challenges have to be addressed by,
  - Utilizing low-power memory devices, such as DDR and NAND, and implementing energy-efficient techniques
  - Use pre-planning and pre-silicon power analysis tools with post silicon correlation
  - Using proper heat sink
  - Using efficient power regulators
  - Using correct process corner parts, etc.



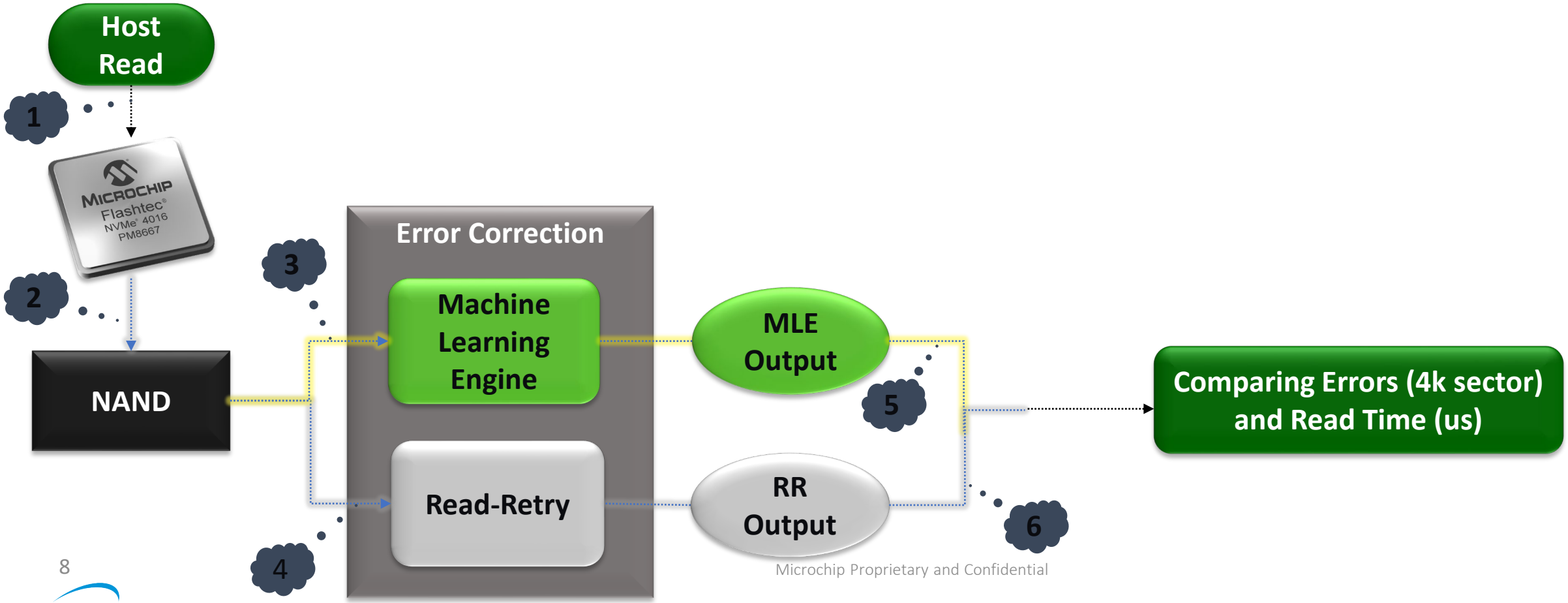
Optimizing Power in Controller

- Power gate unused HW blocks and memory banks
- Ability to throttle the DDR and NAND frequency based on the operation and I/O performance
  - Crucial to dynamically toggle the frequencies according to the performance requirements
- Run all the interfaces at appropriate speed instead of max speed
- Using only necessary number of CPUs
- Use the data path resources efficiently
  - Read and write path may use different resources!



Advanced Power Optimization Options

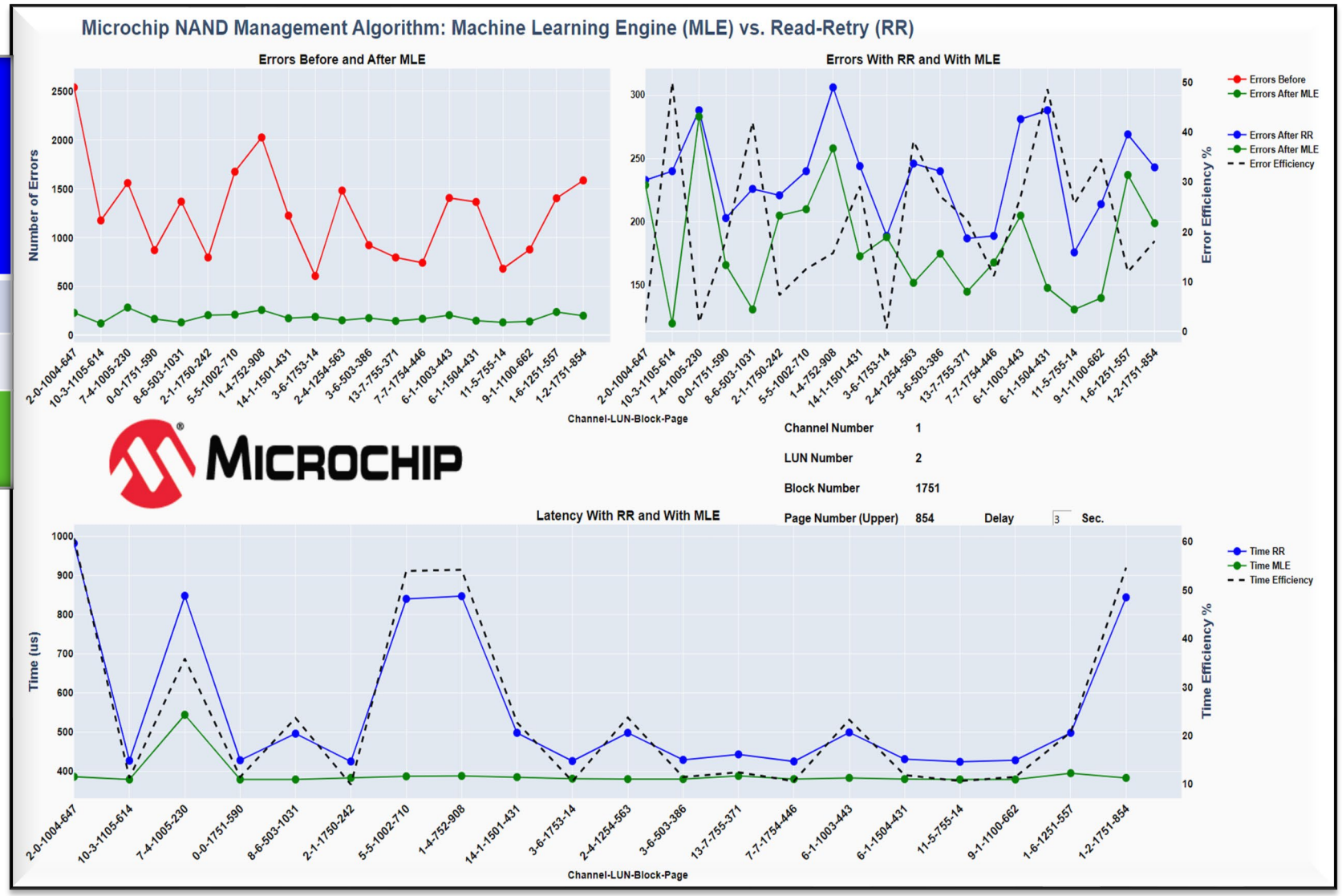
- Machine Learning Engine (MLE) can be used to recognize the IO request patterns, traffic, write amplification to dynamically apply the power management strategies
- MLE based block refresh and wear-leveling will reduce the overall system power usage
- Here is one of the example applications of MLE



Advanced Power Optimization Options

- MLE can be used for NAND and power optimization by correctly predicting the Vt voltage and correcting the NAND errors in less number of iterations

	Total Flash Power (milliwatts)	Number of Iterations (avg)	Time (us)
MLE	63.1	3	~675
RR	69.12	16	~1275
% Improvement	~9%		~47%



Idle Power Optimization

- Some specification requires the Total SSD power to be <5W and hence controller power even lower
- For idle power optimization,
  - Automatic idling of processor cores
  - Clock gating all the interfaces such as DDR, NAND, etc.,
  - Power gate all unused HW block
  - Put the device in one of the PCIe® low power states
  - Divide the clocks to run at lower speeds

- PCIe® Gen-5 power states
  - L0s
  - L1
  - L2
- PCIe® Gen-6 power states
  - L0s
  - L0p
  - L1
  - L2

NVMe® 5016

- Microchip Technology's Flashtec® NVMe 5016 controller is the engine behind the next wave of enterprise SSD innovation—delivering an industry-leading 2.5 GBps per Watt performance benchmark.

Thank you !

Visit Microchip Booth # 115