



Abstracting the Flash Translation Layer for Enterprise-ready MLC

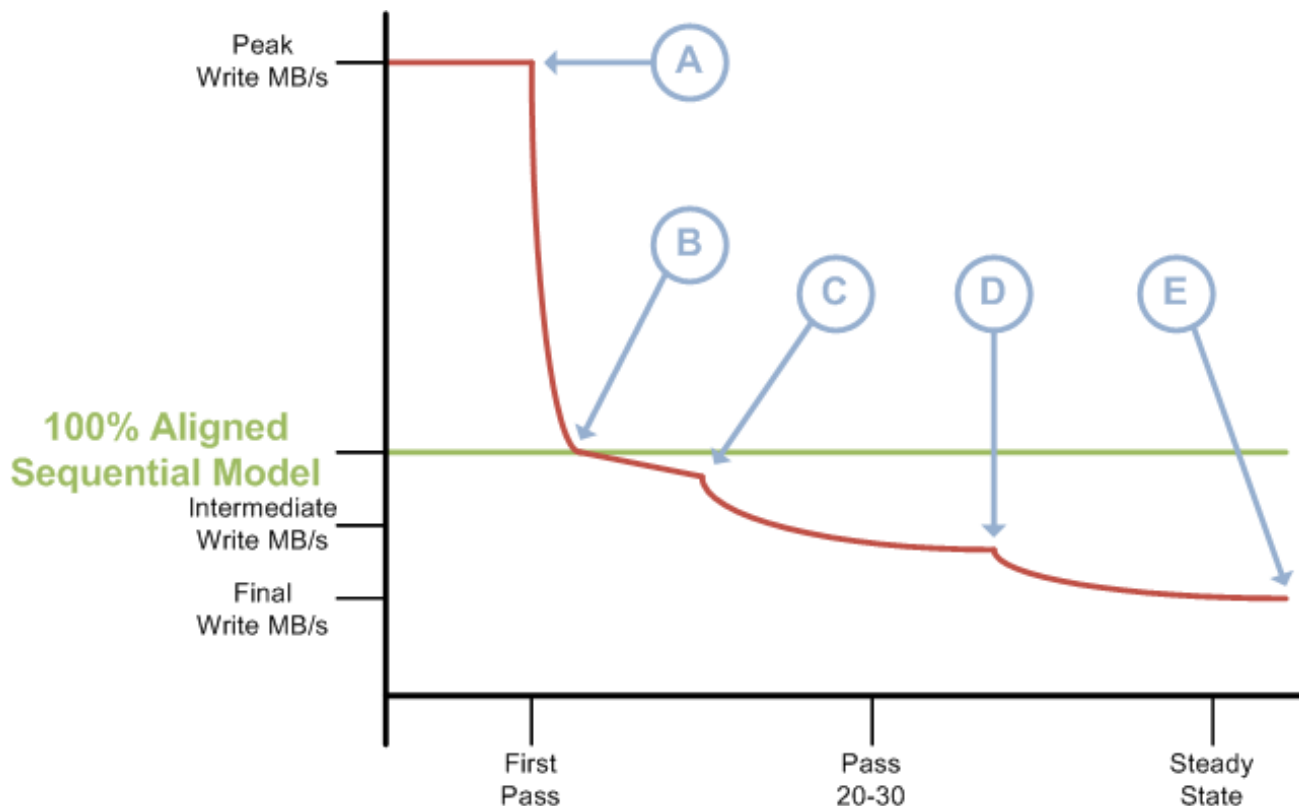
Erik de la Iglesia
GridIron Systems

Issues with SSDs in Arrays

- SSDs are not drop-in replacements for HDDs
- System optimizations for single SSD may not be portable to arrays
- Many complex issues (for 5 minutes)
 - Aligned Sequential Writing Trap
 - Wear Leveling Self-Destruct

Consider FTL as a System / Device Abstraction

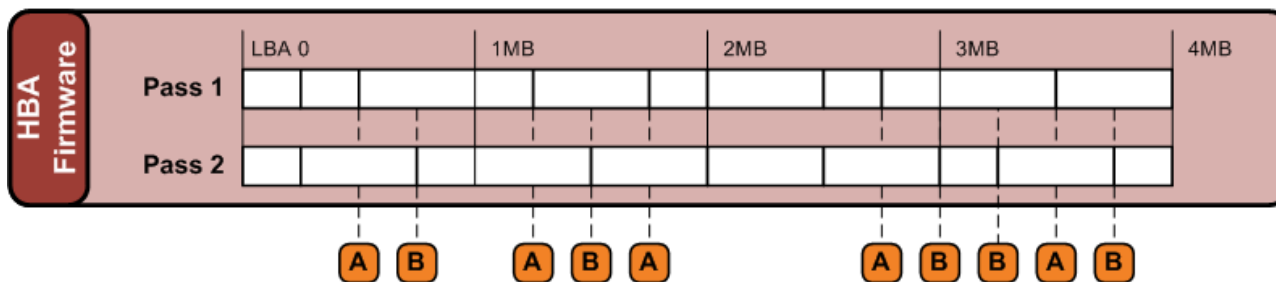
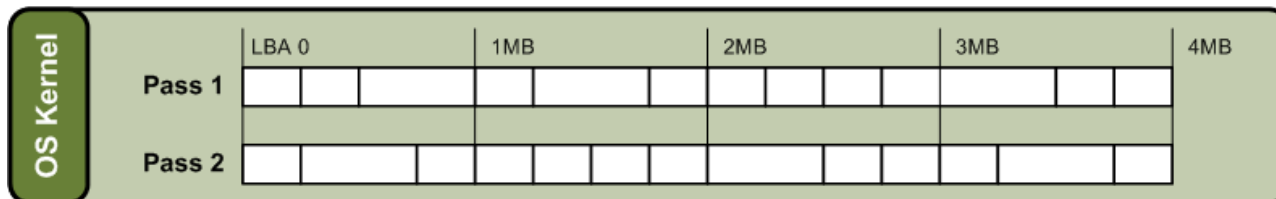
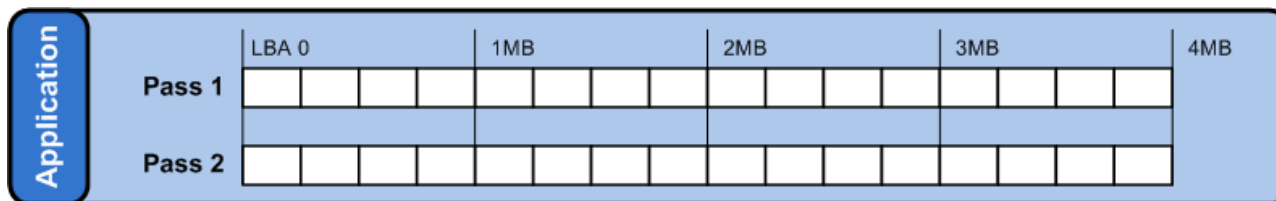
Write Performance over the life of an SSD Device



Theory: Minimize write amplification and controller overhead by using aligned, sequential, fixed-blocksize writing

256KB Aligned Sequential Writes

Two passes over same LBA space

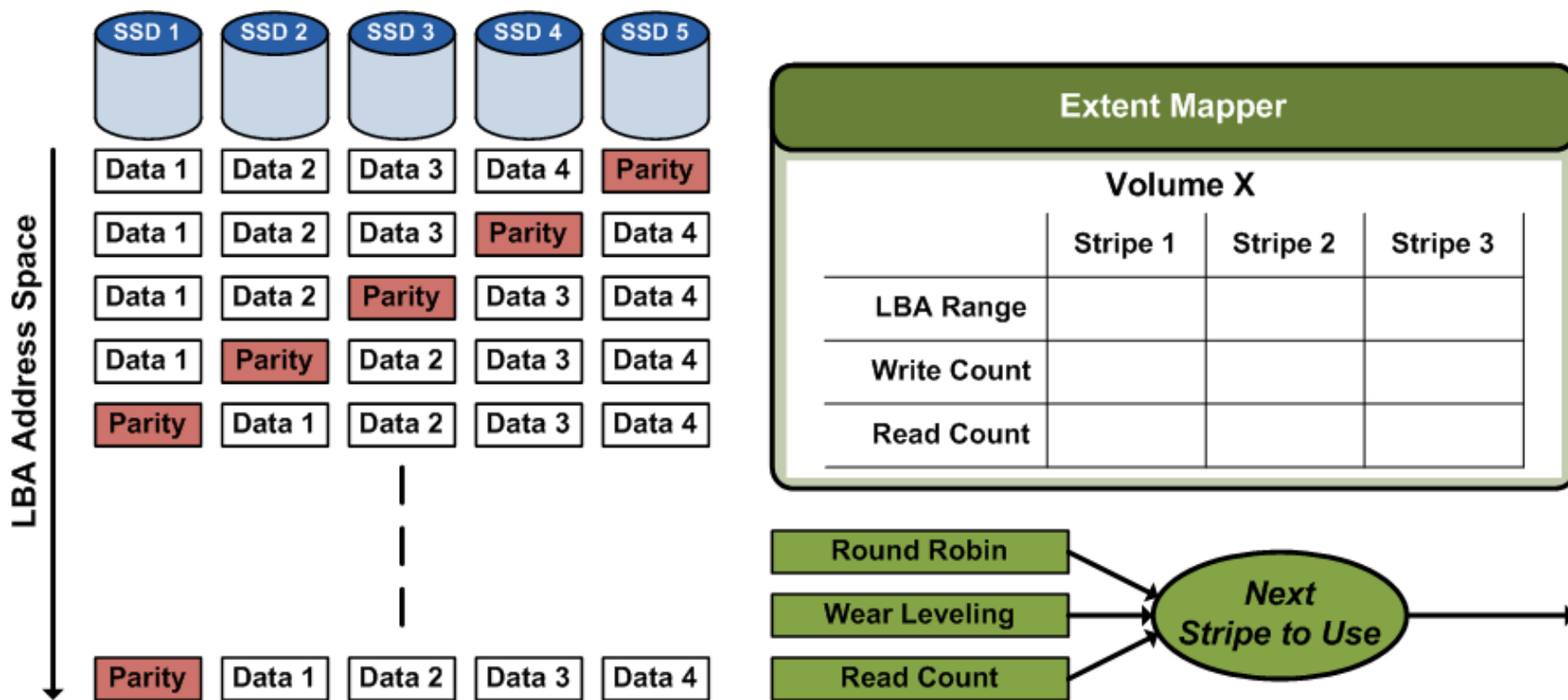


A Indirection Table Write Amplification

B Data Aggregation Write Amplification

Traditional Storage Array

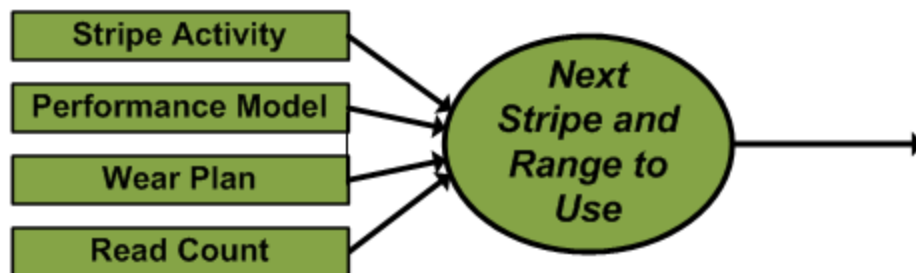
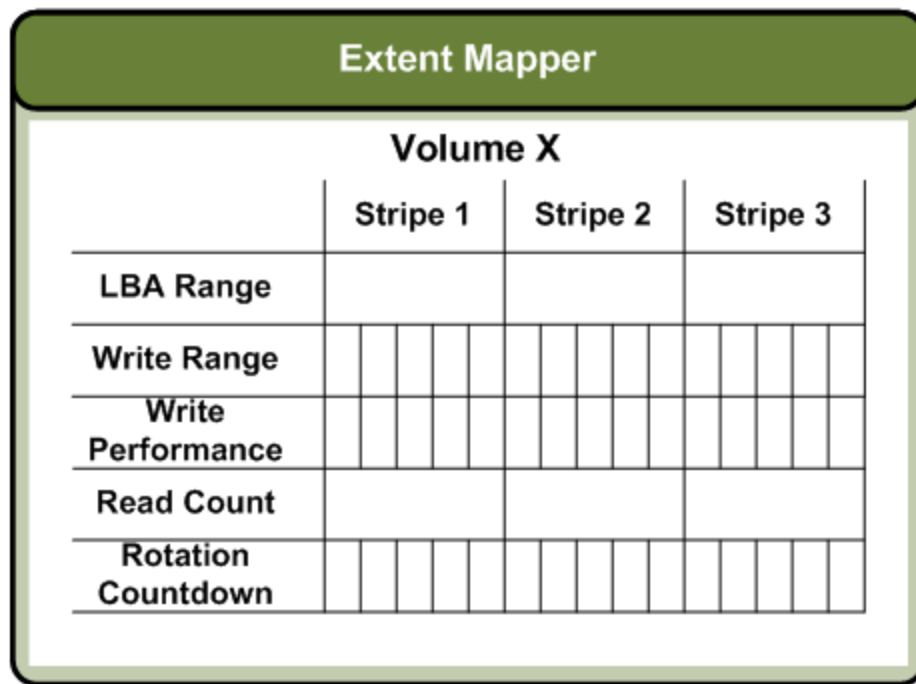
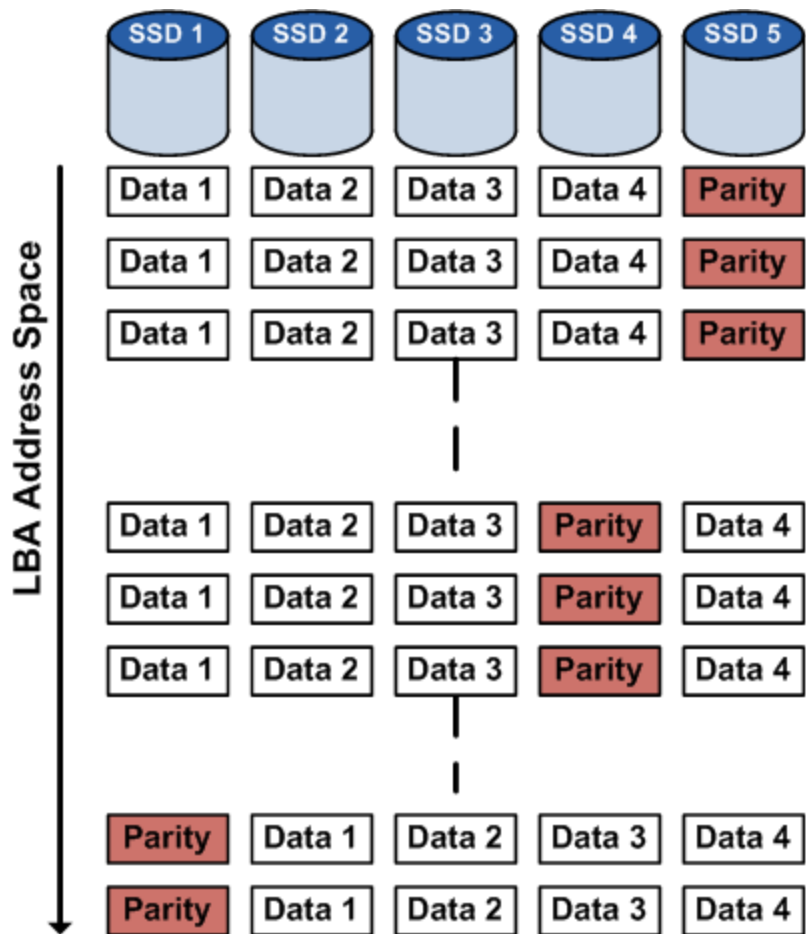
Volumes as Extents over Stripes



Theory: Global wear-leveling extends array life

Flash - Optimized Storage Array

Plan for wear and performance



Final Thoughts / Questions?

- Creating all-flash arrays is easy
- Maintaining performance and reliability over time requires new system architecture
- For performance and reliability, split FTL functions between array controllers and SSD devices
 - Block management and write buffering in SSDs
 - Global wear profiling in array controller
 - Lock down Kernel and HBA behavior