All Flash Array Data Protection Schemes

Sean R. Barry
XtremIO System Engineer
All Flash Arrays – How Did We Get Here?

- 30+ years of HDD and distributed data and parity
- Built for spinning media
  - Data Protection built into the distribution schema

Source: gabrielchapman.com
All Flash Arrays – How Did We Get Here?

- Gen 1 - EPROM
  NAND 35+ Years old

- Gen 2 – Flash
  NAND 18+ Years of innovation

Source: spectrum.ieee.org
All Flash Arrays – How Did We Get Here?

- Gen 2 – Flash NAND 18+ Years of innovation:
  - Substrate
  - Gates
  - Bit Density
  - Manufacturing

Source: www.eetimes.com
Hard Disk Drive Performance

Flash NAND Technology has an opportunity!
THE EVOLUTION OF “FLASH” ARRAYS

Flash Optimized Hybrids
- Dual Controller
- Limited Scale
- Single Workload
- Some Data Services

Scale Up All Flash Arrays
- Dual Controller
- Limited Scale
- Single Workload
- Data Services? - Maybe
- Choice Between
  - Performance
  - Efficiency - Data Svcs

Scale Out All Flash Arrays
- Multi Controller
- Petabyte Scale
- Mixed Workloads
- In-Line Data Services
- The All Flash Data Center

Gen 1 Flash Systems
Legacy Disk Arrays

Gen 1
Gen 2
Gen 3
Gen 4

Source: Wikibon December 2014
ON LOAN FROM THE DISK ERA

Log Structuring

Garbage Collection Nightmare

Post-Processing

Erratic Performance

Legacy RAID

Write Amplification

Scale Up Architecture

Front End Bottleneck
Data Protection Methods on Flash

- Flash-optimized data protection without compromise

<table>
<thead>
<tr>
<th></th>
<th>XDP</th>
<th>RAID 1</th>
<th>RAID 5</th>
<th>RAID 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capacity Overhead</td>
<td>8%</td>
<td>50%</td>
<td>25%</td>
<td>20%</td>
</tr>
<tr>
<td>Write I/O Overhead</td>
<td>1.22</td>
<td>2 (64%)</td>
<td>2 (64%)</td>
<td>3 (146%)</td>
</tr>
<tr>
<td>(stripe update)</td>
<td></td>
<td>(64%)</td>
<td>(64%)</td>
<td>(146%)</td>
</tr>
<tr>
<td>Read I/O Overhead</td>
<td>1.22</td>
<td>0</td>
<td>2 (64%)</td>
<td>3 (146%)</td>
</tr>
<tr>
<td>(stripe update)</td>
<td></td>
<td></td>
<td>(64%)</td>
<td>(146%)</td>
</tr>
</tbody>
</table>

- SUPERIOR
  - USABLE CAPACITY
- UP TO 4X
  - BETTER ENDURANCE
- UP TO 4X
  - BETTER PERFORMANCE
Data Protection Innovation Shift - Hardware

- Lessons learned from enterprise storage field deployments
- Innovation at controller and HDD levels
- When RAID failed, manual data stripe rebuilding

Source: [www.tomshardware.com](http://www.tomshardware.com)
Data Protection Innovation Shift - Hardware

- Hardware Architecture completely different because of Flash NAND Media
- Dual storage controllers
- Software driven data protection
Data Protection Innovation Shift – Metadata and Database Transaction-like I/O Operations

Emptiest Stripe Always Known

<table>
<thead>
<tr>
<th>Volume LBA to Fingerprint (first stage metadata)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LBA offset</td>
</tr>
<tr>
<td>------------</td>
</tr>
<tr>
<td>100</td>
</tr>
<tr>
<td>104</td>
</tr>
<tr>
<td>108</td>
</tr>
<tr>
<td>112</td>
</tr>
<tr>
<td>116</td>
</tr>
<tr>
<td>120</td>
</tr>
<tr>
<td>124</td>
</tr>
<tr>
<td>128</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fingerprint to Physical Location (second stage metadata)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fingerprint</td>
</tr>
<tr>
<td>-------------</td>
</tr>
<tr>
<td>20147A8</td>
</tr>
<tr>
<td>AB45CB7</td>
</tr>
<tr>
<td>F3AFBA3</td>
</tr>
<tr>
<td>963FE7B</td>
</tr>
<tr>
<td>0325F7A</td>
</tr>
<tr>
<td>134F871</td>
</tr>
<tr>
<td>CA38C90</td>
</tr>
</tbody>
</table>

Emptiest Stripe Always Known
All Flash Array Take-Aways

• It’s not about the Flash/SSDs—it’s all about the architecture of NVM management
• All-Flash Arrays are more memory intensive to manage I/O and provide data services
• System design is key to All Flash Array longevity and datacenter utilization
Thank you!

Sean.Barry@emc.com