ioScale: The Holy Grail of Hyperscale

Introduction

The workload requirements for hyperscale datacenters are to have a 40% Compound Annual Growth Rate (CAGR) over the next three years. More people every day are accessing the same datacenters from more devices, using more applications. The majority of hyperscale datacenters meet these workload demands with distributed, scale-out architectures that are designed for efficiency. Open-source applications and databases eliminate expensive software licensing and allow code to be tuned and modified for a company’s specific needs. Similarly, workloads are distributed over commodity server farms, eliminating expensive Storage Area Network (SAN) hardware, management, and support costs. However, these companies face significant challenges addressing the performance and cost impact of slow I/O.

Unlike disk drive-based and SSD-based systems, ioScale minimizes the hyperscale footprint by making servers more powerful, resilient, and reliable. IoScale gives hyperscale environments the performance headroom they need to handle spiking holiday traffic, content gone viral, or the slash-dotted article, with far less infrastructure than disk drive- or SSD-based systems.

Low-Latency Design Speeds Transactional Processing and Application Response

SAS and SATA hard disk- and SSD-based systems place unnecessary controllers, embedded processors, and storage protocols between applications and the storage. This adds latency and slows application response times, forcing hyperscale companies to over provision performance for the highest expected traffic spike. The ioScale cut-through architecture minimizes latency to maximize throughput and application response times. It distributes workload across many more channels of flash than SSDs so each transaction processes faster.

---

1 ISI Group, November 1, 2012
Figure 1. Fusion-io Virtual Storage Layer gives applications direct access to flash memory.

Figure 2. RAID controllers narrow stripe access to just 2-4 channels of flash. ioScale wide-stripes access to many channels, so transactions process faster and applications are more responsive.
Consistent and Predictable Performance for Unpredictable Workloads

The end result of the ioScale cut-through architecture and large single device capacity is more predictable and consistent performance. This enables hyperscale companies to deliver high levels of performance to customers during peak hours. The following diagram illustrates how ioMemory performance compares to SSD performance under a real-world transactional workload.

Figure 3. ioMemory architecture delivers consistently high and predictable performance, while SSD performance is erratic.

Industry-Leading Capacity and Performance Density Lowers Costs

The more capacity a 1U to 2U server can hold, the more transactions that server can process, and the more value that server provides. ioScale delivers from 410GB to 3.2TB of low-latency, high-performance flash memory per PCI Express slot. With up to four times the capacity of a 2.5” disk form factor SSD, ioScale solutions make each server node far more powerful, which translates into lower hardware, power and cooling, and maintenance costs.

Figure 4. Four high-capacity SSDs are needed to match the capacity of one ioScale 3.2TB.
Reliability through Simplicity

Distributed hyperscale architectures are inherently resilient to server failure, but ioScale makes each server even more reliable which, by extension, improves the reliability of the entire system. The ioMemory platform includes powerful features that guarantee reliability. For example, Adaptive Flashback provides complete fault tolerance at the erase block level, which enables ioScale products to self-heal NAND errors or failures without affecting the customer experience. Companies can achieve the same level of reliability on fewer servers, with fewer failure points and less complexity.

The following diagrams show how SSDs introduce more failure points that reduce server uptime, while the ioScale design protects server uptime.

When one SSD or a RAID controller in a RAID 0 fails, the entire server drops out of the system. The SSD approach requires more servers to achieve any given service level, and adds costs of SSD backup inventory and IT overhead to swap devices and restore servers.

Figure 5. SSDs in RAID 0 add five failure points, reducing server uptime.
Adaptive Flashback self-heals NAND errors and failures at the erase-block level, eliminating unnecessary redundancy and failure points.

*Figure 6. Adaptive Flashback protects data without adding failure points, increasing server uptime.*

---

### Optimize Applications to Exploit Flash Memory Potential

Hyperscale companies tune applications and databases to increase performance and efficiency. The ioMemory SDK enables developers to optimize applications for ioScale flash. Examples include the ability to make applications write more efficiently to extend flash endurance, or reclaim unused capacity for more efficient capacity use. For example, Percona, a leading MySQL support and solutions company, posted an Atomic Writes extension to the XtraDB storage engine using the Atomic Writes API. Because the MySQL Atomic Writes feature can replace double-write buffering for most MySQL updates, this extension can effectively double the write endurance storage media.

*Figure 7. Atomic Writes can bypass the MySQL double-write buffer due to inherent atomicity.*
High Endurance and Simple, Comprehensive Flash Management

ioScale implements advanced wear-leveling technology that delivers endurance measured in Petabytes Written (PBW), and lasts longer than most components in the datacenter. The ioMemory technology has proven itself in the world’s most demanding environments. Fusion ioSphere software provides comprehensive monitoring and management of all ioMemory devices in the data center, allowing customers to see all device health and life expectancy at a glance.

*Figure 8. ioSphere simplifies flash management*

<table>
<thead>
<tr>
<th>SSD CHALLENGES</th>
<th>IOSPHERE BENEFITS</th>
</tr>
</thead>
<tbody>
<tr>
<td>• No Centralized Management</td>
<td>• Web-Based Centralized Management</td>
</tr>
<tr>
<td>• No Real-Time Monitoring or Alerting</td>
<td>• Live Performance Monitoring and Alerts</td>
</tr>
<tr>
<td>• No Remote Configurability</td>
<td>• Remote Format and Firmware Updates</td>
</tr>
<tr>
<td>• No Remote Access Across Networks</td>
<td>• Remote Access and Configurability</td>
</tr>
</tbody>
</table>
ioScale Increases Hyperscale Efficiency

The market is filled with competing SSDs that promise more performance and lower costs. But value is about more than straight dollars per gigabyte. The SSD array approach requires more servers, more devices, and more complexity that far outweigh any device savings. The following diagrams illustrate the efficiency and economy of ioScale-based environments.

*Figure 9. ioScale minimizes server and infrastructure costs.*

<table>
<thead>
<tr>
<th>SSDs</th>
<th>ioScale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Server</td>
<td>2U</td>
</tr>
<tr>
<td>1U</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Devices for 3.2TB Capacity</th>
<th>SSD SSD SSD SSD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spare Device Inventory</td>
<td>SSD SSD</td>
</tr>
<tr>
<td>Not needed</td>
<td></td>
</tr>
</tbody>
</table>

| Power and Cooling |  |
|-------------------|  |
| RAID Controller   | None |
| Management Tools  | None |
| RAID Array Maintenance | None |

As datacenter workloads grow, so do savings on the cost to scale out.

*Figure 10. ioScale cost savings propagate throughout the datacenter.*
Summary

As this paper shows, ioScale is flash that has been optimized for hyperscale environments. A perfect complement to the efficiencies hyperscale architectures are designed for, ioScale is the efficient answer to slow I/O. It offers the most performance and capacity per server, while minimizing infrastructure costs. Its low-latency design guarantees consistently high levels of performance for all workloads; it doesn’t require aggregating devices, which increase system complexity; and it offers an SDK so companies can optimize applications to run natively on flash, and truly exploit the media’s potential.