AN ALTERNATIVE TO ALL-FLASH ARRAYS: PREDICTIVE STORAGE CACHING

THE EASIEST WAY TO INCREASE PERFORMANCE AND LOWER STORAGE COSTS

Bruce Kornfeld, Chief Marketing Officer, StorMagic
Luke Pruen, Technical Services Director, StorMagic
Peter Smith, System Admin, Harris Corporation
STORMAGIC SVSAN: BRIEF OVERVIEW

SvSAN turns the internal disk, SSD and memory of two or more servers into highly available shared storage.
ALL-FLASH ARRAYS ARE TEMPTING

- SSDs vs. HDD
  - 10x-20x more performance
  - 10x more expensive
- Lots of hype from flash array vendors
- Not all workloads need all-flash
- A balanced approach with advanced caching algorithms can meet your needs

Source: Statistica

SSDs are still 10x more costly than HDD.

SSD: $0.50 - $1.00 per GB
HDD: $0.05 - $0.10 per GB
PREDICTIVE STORAGE CACHING

Over 400% performance improvement with a patent-pending method for predicting when data will become ‘hot’

AUTOMATED
Two caches – system memory and SSD
Patent-pending predictive algorithms
Automation tracks ‘hot’ data and places in the right cache

BUILT FOR PERFORMANCE
Lower latency – less waiting for spinning disks
Data pinning
Solves the virtual server ‘I/O blender effect’

COST EFFECTIVE
Use fewer and less expensive disk drives (7.2K)
Lower CAPEX – server memory is inexpensive
Lower OPEX – power, cooling and maintenance
AN ALTERNATIVE TO ALL-FLASH ARRAYS: PREDICTIVE STORAGE CACHING

THE EASIEST WAY TO INCREASE PERFORMANCE AND LOWER STORAGE COSTS

Luke Pruen, Technical Services Director, StorMagic
TODAY’S STORAGE OPTIONS

- The performance gap between CPU and storage
- Disk only
  - High capacity, low performance workloads
- All-flash
  - Performance at any cost workloads
- Hybrid
  - Most workloads fit here

<table>
<thead>
<tr>
<th></th>
<th>Disk only</th>
<th>All-flash</th>
<th>Hybrid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latency</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Capacity</td>
<td>High</td>
<td>Low/Medium</td>
<td>High</td>
</tr>
<tr>
<td>Performance</td>
<td>Low/Medium</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Cost per GB</td>
<td>$</td>
<td>$$$</td>
<td>$$</td>
</tr>
<tr>
<td>Cost per IOP</td>
<td>$$$</td>
<td>$</td>
<td>$</td>
</tr>
</tbody>
</table>
THE IMPORTANCE OF CACHING

• Virtualized environments suffer from the ‘I/O blender effect’

• Working sets of data change over time

• Advanced caching can solve both problems
WRITE CACHING

Stage 1
- All new data written to SSD
- Data marked ‘dirty’ – not committed to HDDs

Stage 2
- Write operation is acknowledged immediately to the server

Stage 3
- ‘Dirty’ data is reordered and grouped based on disk locality
- Data de-staged and written to HDD, sequentially as possible

Stage 4
- SSD cache notified when data successfully written to HDD
- Cached data marked ‘clean’, remains in cache until space needed
WRITE CACHING

Stage 1
• All new data written to SSD
• Data marked ‘dirty’ – not committed to HDDs

Stage 2
• Write operation is acknowledged immediately to the server

Stage 3
• ‘Dirty’ data is reordered and grouped based on disk locality
• Data de-staged and written to HDD, sequentially as possible

Stage 4
• SSD cache notified when data successfully written to HDD
• Cached data marked ‘clean’, remains in cache until space needed
READ AHEAD & DATA PINNING

Read ahead mode
• I/O blender effect aware!
• Identifies sequential interleaved I/Os
• Detects sequential read streams
• Pre-fetches data into memory

Data pinning mode
• Pin specific data/workloads in memory
• Delivers most efficient read performance
• DBs, VDI, frequently repeated operations
• Manage multiple pin groups

Diagram:
- Read Operation
- Data Blocks
  - Data read from disk
  - Data pre-fetched from disk
  - Data not yet requested

Chart:
- Time to Boot VSAs (machines booting simultaneously)
- No caching (max)
- No caching (avg)
- Read data pinned (avg)
Predictive approach
- Tracker module has 7 levels
- All read I/Os monitored and analyzed
- Most frequently used – higher levels
- Cache placement based on levels

Flexible storage options
- RAM: Most frequently accessed data
- SSD/flash: Next most frequently accessed data
- HDD: Infrequently accessed data – ‘cold’ data

Sizing
- Assign cache sizes to meet requirements
- Grow caches as working sets change
- Use any combination of memory, SSD/flash and disk

Play to the strengths
- Play to the strengths of all mediums
- Memory highest IOPS
- SSD/flash magnetic drives providing low price per GB
REAL CUSTOMER DATA (BEFORE CACHING)

Performance characteristics of their existing workloads – measured by StorMagic

Workload

- 12 Virtual Machines
- 78 applications

<table>
<thead>
<tr>
<th></th>
<th>Read</th>
<th>Write</th>
</tr>
</thead>
<tbody>
<tr>
<td>Read/Write %</td>
<td>77%</td>
<td>23%</td>
</tr>
<tr>
<td>Sequential %</td>
<td>49%</td>
<td>39%</td>
</tr>
<tr>
<td>Average Per Day</td>
<td>991 GB</td>
<td>294 GB</td>
</tr>
<tr>
<td>Average Block Size</td>
<td>58 KB</td>
<td>54 KB</td>
</tr>
<tr>
<td>Average IOPS</td>
<td>212</td>
<td>138</td>
</tr>
</tbody>
</table>

Throughput IOPs

Block Size Distribution

Locality of access
REAL CUSTOMER DATA (WITH CACHING)

Impact of Predictive Storage Caching to increase performance

HDD Only (no caching)
- 1 x RAID5 = 3 x 1.2TB 10K SAS disks

HDD & Memory (with caching)
- 12GB of memory per host for caching
- 1 x RAID5 = 3 x 1.2TB 10K SAS disks

HDD Only

HDD & SSD (with caching)
- 1 x 200GB Samsung SSD
- 1 x RAID5 = 3 x 1.2TB 10K SAS disks

HDD, SSD & Memory (with caching)
- 12GB of memory per host for caching
- 1 x 200GB Samsung SSD
- 1 x RAID5 = 3 x 1.2TB 10K SAS disks

HDD & SSD

HDD, SSD & Memory

With Caching

<table>
<thead>
<tr>
<th></th>
<th>HDD Only</th>
<th>HDD &amp; Memory</th>
<th>HDD &amp; SSD</th>
<th>HDD, SSD &amp; Memory</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total IOPS</td>
<td>2400</td>
<td>12341</td>
<td>6165</td>
<td>13061</td>
</tr>
<tr>
<td>AVG Latency (ms)</td>
<td>21.86</td>
<td>4.94</td>
<td>9.37</td>
<td>5.21</td>
</tr>
</tbody>
</table>
SYNTHETIC ‘HERO’ NUMBERS

High performance achievement in a controlled lab environment

**HDD Config**
- 1 x RAID5 = 3 x 1.2TB 10K SAS disks

**Tiered Config**
- 1 x RAID5 = 3 x 1.2TB 10K SAS disks
- 1 x 200GB Samsung SSD
- 32GB of memory per host for caching

**Test**
- vSphere 6.5
- 2 x Windows VMs
- IOmeter
- VMDK on VMFS

### Total IOPS

<table>
<thead>
<tr>
<th>Workload</th>
<th>HDD  IOPS</th>
<th>Tiered IOPS</th>
<th>HDD  AVG (ms)</th>
<th>Tiered AVG (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4KB 100% Random Read</td>
<td>4263</td>
<td>207117</td>
<td>60.04</td>
<td>1.24</td>
</tr>
<tr>
<td>4KB 100% Random Write</td>
<td>2004</td>
<td>38196</td>
<td>127.17</td>
<td>6.70</td>
</tr>
<tr>
<td>4KB 70/30 Read Write 80% Random</td>
<td>1832</td>
<td>79738</td>
<td>147.36</td>
<td>3.21</td>
</tr>
<tr>
<td>32KB 100% Sequential Read</td>
<td>7729</td>
<td>85803</td>
<td>33.62</td>
<td>2.98</td>
</tr>
<tr>
<td>32KB 100% Sequential Write</td>
<td>8728</td>
<td>15651</td>
<td>29.33</td>
<td>16.36</td>
</tr>
</tbody>
</table>

### AVG Latency (ms)

<table>
<thead>
<tr>
<th>Workload</th>
<th>HDD AVG (ms)</th>
<th>Tiered AVG (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4KB 100% Random Read</td>
<td>1.24</td>
<td>60.04</td>
</tr>
<tr>
<td>4KB 100% Random Write</td>
<td>6.70</td>
<td>127.17</td>
</tr>
<tr>
<td>4KB 70/30 RD/WR 80% Random</td>
<td>3.21</td>
<td>147.36</td>
</tr>
<tr>
<td>32KB 100% SEQ RD</td>
<td>2.98</td>
<td>33.62</td>
</tr>
<tr>
<td>32KB 100% SEQ WR</td>
<td>16.36</td>
<td>29.33</td>
</tr>
</tbody>
</table>
SUMMARY: PREDICTIVE STORAGE CACHING

Increased performance – without the high cost of all-flash arrays

Flexible configurations – we’ll help ‘right-size’ the server hardware

Lower CAPEX – use slower drives for capacity, low-cost system memory and some SSD for performance

Lower OPEX – less power, cooling and maintenance
Q&A – HARRIS CORPORATION

Peter Smith
Systems Administrator
Network, Platforms and Security Group
Harris Corporation
**NEXT STEPS**

**Further reading:**


---

**SvSAN Product Information**

<table>
<thead>
<tr>
<th>Product</th>
<th>Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>SvSAN license</td>
<td>2, 6, 12 and unlimited TBs</td>
</tr>
<tr>
<td>License entitlement</td>
<td>2 mirrored servers</td>
</tr>
<tr>
<td>Maintenance and support</td>
<td>Platinum - 24x7 / Gold - 9x5</td>
</tr>
</tbody>
</table>

For further information, please contact: [sales@stormagic.com](mailto:sales@stormagic.com)