Multi-Tier Subsystem Management Using SLOs

Kaladhar Voruganti
Technical Director, CTO Office
NetApp, Sunnyvale
January 29th, 2013
Talk Outline

- Part 1: Discuss Storage Architecture Trends due to NVM

- Part 2: Impact of the above trends on Storage Management Paradigm (Need for management by SLOs)

- Part 3: How can Standards Help?
Talk Outline

- Part 1: Discuss Storage Architecture Trends due to NVM

- Part 2: Impact of the above trends on Storage Management Paradigm (Need for management by SLOs)

- Part 3: How can Standards Help?
Trend 1: NVM Accelerating Multi-Layered Storage Architectures

Traditional Network Storage

- DAS Based Network Storage
  - Tries to Optimize For Both Capacity & IOPs

Emerging World

- Host Side Flash
- Network Flash
- All-Flash Array
- Cloud Storage
- SSD/HDD Array

IOPs Value
Capacity Value

© 2012 NetApp, Inc. All rights reserved.
# NVM Solution Alternatives

<table>
<thead>
<tr>
<th>Category</th>
<th>Solutions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Server Flash Cache</strong></td>
<td></td>
</tr>
<tr>
<td>(Dedicated/Clustered)</td>
<td>NetApp Flash Accel,</td>
</tr>
<tr>
<td></td>
<td>Fusion-IO/IO-Turbine, LSI,</td>
</tr>
<tr>
<td></td>
<td>VFCache, Flashsoft,EMC Lightning</td>
</tr>
<tr>
<td><strong>Network Flash Cache</strong></td>
<td></td>
</tr>
<tr>
<td>(Near-server / in-network)</td>
<td>NetApp CacheIQ</td>
</tr>
<tr>
<td></td>
<td>EMC Thunder,ION,</td>
</tr>
<tr>
<td></td>
<td>Avere, GridIron</td>
</tr>
<tr>
<td><strong>All-Flash Array</strong></td>
<td>Violin</td>
</tr>
<tr>
<td></td>
<td>WhipTail</td>
</tr>
<tr>
<td></td>
<td>XtremIO (EMC)</td>
</tr>
<tr>
<td><strong>Storage Flash Cache/Tier</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NetApp Flash Cache</td>
</tr>
<tr>
<td></td>
<td>NetApp Flash Pools, EMC FAST, HDS</td>
</tr>
</tbody>
</table>
Trend 2: NVM Enabling Load/Store Persistence Programming Model

Applications

SNIA NVM API
- open/close/mmap/munmap
- read/write/fsync
- load/store/msync

SNIA NVM libc
- memcpy

VFS
- page fault

OEM NVM filesystem driver

NVM

user

kernel

hardware
Talk Outline

- Part 1: Discuss Storage Architecture Trends due to NVM

- Part 2: Impact of the above trends on Storage Management Paradigm (Need for management by SLOs)

- Part 3: How can Standards Help?
Impact 1: Service Offered by Multiple Layers and Vendors

Today

- Gold
- Bronze

Service Offered By
- Single Layer
- Single Vendor

Emerging

- Gold
- Bronze

Host Side Flash

Service Offered By a combination of
- Multiple Layers
- Multiple Vendors

HDD Array

Cloud Storage
Impact 2: Service Level To Configuration Binding is Not Fixed

Today
- SSD
- SATA

Emerging
- Time T1: Flash at Host
- Hybrid SSD/SATA (use SATA)
- Time T2: No Flash at Host
- Hybrid SSD/SATA (Use SSD)

- Fixed Binding between Configuration and Service Level
- Dynamic Binding between Configuration and Service Level for Better resource utilization
Impact 3: In band Policy Binding

**Today**

- Volume
  - Created using Gold Service (policy)
  - Map file/LUN to Gold Volume
  - Application reads/writes to the file/LUN that is bound a priori to Gold Storage

**Emerging NVM Load/Store Model**

- Application asks for Gold Policy
  - In band Policy Mapping

- SNIA NVM API library
  - Requests Kernel for Gold NVM region
  - Software Layer on top of SNIA NVM API Layer
    - Provides Gold Level Semantics to Gold NVM region

- Application specifies Policy
  - In band

Application does not specify Policy

In band Policy Mapping

© 2012 NetApp, Inc. All rights reserved.
Impact 4: Data Structure Level Policies

Today

Volume
Created using
Gold Service (policy)

Map file/LUN to
Gold Volume

Application
reads/writes
to the file/LUN that
is bound a priori to Gold Storage

File, Object, LUN, Volume
Level Policy Management

Emerging
NVM Load/Store Model

Application
Asks for Gold Policy
For data structure

malloc(size,GoldService);
new(KVStore, GoldService);

Software Layer on top of SNIA NVM API Layer
Provides Gold
Level Semantics
To Gold Data Structure
In Gold NVM Region

Data Structure Level
Policy Management
Talk Outline

- Part 1: Discuss Storage Architecture Trends due to NVM

- Part 2: Impact of the above trends on Storage Management Paradigm (Need for management by SLOs)

- Part 3: How can Standards Help?
Following Standards Efforts are Required

- Service Definition using SLOs needs to be standardized
  - For coordination between different layers
  - For policy specification in Load/Store access model

- Inter-Layer Control Protocol needs to be standardized

- Inter-Layer I/O Protocol needs to be standardized
**SLO Standards for Service Definition**

<table>
<thead>
<tr>
<th>SLO Dimension</th>
<th>Old way of describing Service Attributes using Vendor technology attributes</th>
<th>SLO Way of describing Service Attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance</td>
<td>SATA, SSD etc</td>
<td>Latency Ceiling</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Latency Floor</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IOPs or Bandwidth Ceiling</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IOPs or Bandwidth Floor</td>
</tr>
<tr>
<td>Operational Recovery</td>
<td>SnapVault, SnapMirror, NDMP</td>
<td>RPO</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Retention Time</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Backup Mode</td>
</tr>
<tr>
<td>Disaster Recovery</td>
<td>SnapMirror Async, MetroCluster</td>
<td>RPO</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Distance</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Replication Mode</td>
</tr>
<tr>
<td>Protection (Availability)</td>
<td>RAID-DP, HA</td>
<td>Media Failure Protection</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Shelf Failure Protection</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Head Failure Protection</td>
</tr>
<tr>
<td>Space Guarantee</td>
<td>Dedup, Compression, Fractional Reserve etc</td>
<td>Thin Provisioning</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Thick Provisioning</td>
</tr>
</tbody>
</table>

- Services Described using Quantitative, Vendor Technology Independent SLO Attributes
- Service Definition can encapsulate multiple layers managed by different Vendors
- We need to standardize both SLO dimensions and SLO attributes for Services
Control Protocol Standards Needed

Coordinate Control Actions Between Layers

<table>
<thead>
<tr>
<th>Migration</th>
<th>Workload Throttling</th>
<th>Caching</th>
<th>Encryption</th>
</tr>
</thead>
<tbody>
<tr>
<td>Which Layer Should Migrate data During SLO violation</td>
<td>Which layer should throttle Workload during SLO violation</td>
<td>Coordinate Cache Sizing across layers</td>
<td>Which layer should encrypt data</td>
</tr>
</tbody>
</table>

- Control Protocol works either directly between different vendor layers or between an Orchestrator and the different vendor layers.

- Is SMI-S the right place to work on this standard, or where do we standardize this?

- Need a Common Cost Model for Coordination amongst layers.
I/O Protocol Coordination Standards Needed for the following

- Sharing of QoS Class (Gold, Silver, Bronze etc) Information across Layers (tag each I/O)

- Share Caching hints across layers
  - E.g. NFS 4.2 fadvise(), SCSI Disable Page Out

- Share Encryption/Dedup Hashes across layers to get the best of storage efficiency and encryption

- Sharing storage efficiency info (e.g. a priori knowledge that there is a clone relationship and so prepopulate the sharing map at the host)
Conclusion

- NVMs are accelerating the bifurcation of Capacity and IOPs optimized storage architectures

- Need to take a holistic end to end data management view to manage both IOPs and Capacity layers

- SLO Based Management Notions hide the multi-layer, multi-vendor issues from storage subscribers

- Standards have a key role to play for service definition, control and I/O interactions between layers
Acknowledgements

- Steve Byan
- David Dale
- Fred Knight
- Madalin Mihailescu
- Bala Ramachandran
- Doug Santry