NVMe™ Namespaces
Micron Storage Solutions Engineering
SDC EMEA 2020
Agenda

- NVMe™ Namespaces Overview
- NVMe™ Namespaces different use cases
- Micron Study: VMware vSAN + NVMe Namespace Magic:
  - Split 1 SSD into 24 devices for great storage performance
<table>
<thead>
<tr>
<th><strong>Founded 40 years ago on October 5, 1978</strong></th>
<th><strong>Headquartered in Boise, Idaho, USA</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>$23.4B FY2019 annual revenue</td>
<td>4th Largest semiconductor company in the world</td>
</tr>
<tr>
<td>18 Countries</td>
<td>105 On the 2019 Fortune 500</td>
</tr>
<tr>
<td>13 Manufacturing sites and 13 customer labs</td>
<td>40,000 Patents granted and growing</td>
</tr>
<tr>
<td>34,000 Team members</td>
<td>Jan '20 New office opened in Israel</td>
</tr>
</tbody>
</table>
The Only Portfolio Spanning the Data Center Hierarchy

Compelling mix of memory, storage & persistent memory solutions.
### NVMe™ Namespaces

<table>
<thead>
<tr>
<th>What They Are</th>
<th>Namespaces divide an NVMe SSD into logically separate and individually addressable storage spaces. Namespaces may each have their own IO Queue.</th>
</tr>
</thead>
<tbody>
<tr>
<td>What They Do</td>
<td>Configured namespaces are used by a physical or virtual server. A namespace appears as a separate SSD to the connected host. Hosts interact with namespaces like they would local or shared NVMe targets.</td>
</tr>
<tr>
<td>Different From LUNs</td>
<td>NVMe namespaces have dedicated queues (submission, completion); they are SSD local and OS independent</td>
</tr>
<tr>
<td>Uses</td>
<td>Applications and workloads that benefit from multiple devices or where NVMe is shared across multiple systems</td>
</tr>
</tbody>
</table>
What are Namespaces?
Like partitions but not really…

**Partitions**
- Host level
- Shared hardware
- Single queue
- Serial operations
- Not presented as separate devices
- Cannot be used with vSAN

**Namespaces**
- Controller level
- Dedicated hardware
- Dedicated queues
- Parallel operations
- Presented as separate devices
- Can be used with vSAN
NVMe™ Namespaces
Dedicated Acceleration

A single NVMe SSD used as a Dedicated, non-interfering acceleration for groups of slow, in-server storage (HDD or SSD).

Configure the namespaces to provide the level of acceleration needed

Each namespace has its own IO queues

Example: Red Hat Ceph Storage
NVMe™ Namespaces
Shared Acceleration

High bandwidth fabrics enable shared NVMe. Namespaces share NVMe across platforms.

Sharing single NVMe SSD via namespaces: amortize cost, broaden NVMe benefit.

Configure namespace accessibility by host.

Example: Higher bandwidth, low latency Ethernet.
Fixed Thread Count Per Storage Device Application

Namespaces advantage: Make more devices

Application architecture assigns 2 threads per storage device, which under-utilize the NVMe™ SSD performance.

Each namespace appears as a discrete storage device to the host & application and get # of threads assigned to it.

More threads = higher performance!
Tiered Storage Application

Tiered storage, one NVMe™ SSD

Each namespace appears as a discrete storage device to the host & application

Combine namespaces just like physical drives – even on the same physical NVMe SSD
Each namespace appears as a **discrete storage device** to the host & application.

Namespaces appear as 8 physical NVMe storage devices.
Tiered Storage Application

Tiered storage, one NVMe™ SSD

HCI platforms such as vSAN creates storage groups out of physical disks.

Instead of using multiple physical disks for capacity tier and Caching tier –

Using namespaces, you can make disk groups from a single NVMe SSD
Micron Study – VMware vSAN + NVMe™ Namespaces Magic

Micron Storage Solutions Engineering
Test System Overview

Two Dell R730xd servers

- Dual Intel Xeon 2690v4 processors
- 256GB RAM
- Single 15.36TB Micron 9300 NVMe SSD, Divided into 24 namespaces
- Dual 25GbE NICs
Testing Overview
- The results were obtained by running HCIBench configuration for each test.
  - Configurations
    - 4x VMs per node
    - 8x VMDKs per VM (100GB each)
    - 4x threads per VMDK
    - 128 outstanding IOs
    - vSAN Default Storage Policy, deduplication and compression disabled

Key Metric
- How vSAN performance scaled with the number of disk groups, capacity drives per disk group, storage profiles, etc.
Disk Group Scaling – Write Test

<table>
<thead>
<tr>
<th></th>
<th>Average of IOPS</th>
<th>Average of Latency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>16,051</td>
<td>15.93</td>
</tr>
<tr>
<td>2</td>
<td>26,337</td>
<td>9.70</td>
</tr>
<tr>
<td>3</td>
<td>28,972</td>
<td>8.86</td>
</tr>
</tbody>
</table>
Disk Group Scaling – Read Test

Three-disk groups show almost 3x the performance over one-disk group

<table>
<thead>
<tr>
<th></th>
<th>Average of IOPS</th>
<th>Average of Latency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>112,044</td>
<td>2.28</td>
</tr>
<tr>
<td>2</td>
<td>208,262</td>
<td>1.22</td>
</tr>
<tr>
<td>3</td>
<td>319,182</td>
<td>0.80</td>
</tr>
</tbody>
</table>
2U 4-node Supermicro Big Twin (SYS-2029BT-HNC0R)
- Dual Intel Xeon Gold 6142 Processors
- 384GB RAM
- Single Micron 9300 NVMe™ SSD 15.36TB Divided to 24 namespaces:
  - 3x 600GB namespaces for cache drives
  - 21x 594GB namespaces for capacity drives
3x 25GbE Networking
- 1 for management, vMotion, etc.
- 2 dedicated for vSAN
1x 15.36TB 9300 per node
32 namespaces per 9300
- 3 disk groups
  - 1x 600GB cache
  - 7x 549GB capacity
Mellanox 25 GbE switch

Key Metric:
- Maximize 4k Rand Reads IOPS and 128K Seq Reads throughput (GB/s), while also maintaining a reasonable latency for each
- There is a certain point where your performance will no longer increase, but your latency will, and that is where we stopped adding threads
Namespaces with vSAN

2U 4-node Supermicro “Big Twin”
Dual Intel 6142 Gold CPUs
3x 25GbE Networking
- 1 for management, vMotion, etc.
- 2 dedicated for vSAN
1x 15.36TB 9300 per node
32 namespaces per 9300
- 3 disk groups
  - 1x 600GB cache
  - 7x 549GB capacity
Mellanox 25 GbE switch
VMworld 2019 Demo

Configuring a 4 node vSAN cluster into 24 namespaces

*No special tool required to Create or attach namespaces, simply use esxcli:
Create: `esxcli nvme device namespace create -A vmhba3 -c 1258291200 -p 0 -f 0 -m 0 -s 1258291200`
Attach: `esxcli nvme device namespace attach -A vmhba3 -c 1 -n 1`
VMworld 2019 Demo

750K IOPS (4K random reads)

Over **11.5GB/s** (128K sequential reads)!
Takeaways

- Namespaces allow scaling and drive more performance with a single NVMe™ SSD than most solutions do with 20+ physical drives.

- Using multiple Disk Groups and namespaces on a single NVMe™ SSD generate almost 3x than single DG.

- NVMe™ namespaces reduce the system form factor for higher density.

- Power savings, Creating less heat & Higher Reliability of NVMe™ SSDs - Reducing your TCO...
Additional Collateral


- https://www.micron.com/about/blog/2019/october/not-all-nvme-ssds-are-the-same-choosing-the-right-one-can-be-easy


- HCIBench

- Feel free to reach out:  
  Or Lapid, FAE  
  olapid@micron.com  
  +972-54-7716676  
  Esther Spanjer, Dir. Business Development  
  espanjer@micron.com  
  +972-50-8867576