SCI StorInt™ Dispatch – EMC080122
Announcing DMX SSD and Thin Provisioning

This Silverton Consulting (SCI) Storage Intelligence (StorInt™) Dispatch provides a summary of EMC's latest DMX SSD/Flash and Virtual (thin) Provisioning support coming out in March with release 5773 of DMX microcode.

Summary

A precursor to Symmetrix had SSD/DRAM drives and EMC seems now to be returning to these roots. Also, most of today's SSD devices use SSD/DRAM drives. EMC's use of SSD/Flash is an industry first and thankfully much cheaper than SSD/DRAM as well.

Virtual provisioning is an EMC DMX umbrella term that covers thin provisioning as well as increased usability and configurability to speed up defining LUNs on DMX. One EMC chart showed that the new microcode (release 5773) can create a 1TB LUN in under 10 minutes, ~ 6X faster than two levels back (release 5771). Thin provisioning allows a LUN to be defined but not reserve actual space until written to with only minimal space allocated to support blocks being written.

DMX SSD/Flash

There are some obvious advantages to SSD/Flash drives. Read response time is much faster than real disk. Write response time is not as good but is still better than real disk. SSD/Flash drives consume much less power. SSD/Flash costs 30 times more than disk capacity but merits this price by providing 30 times the IOPs performance.

EMC suggests that the optimum environment for SSD/Flash drives is one with a low latency, high transaction workload. SCI would include one more caveat, a high Read to Write ratio workload. Single level (SLC) NAND flash can only handle 100K erase/write passes before a memory block becomes unreadable and requires a complete memory block be erased before any of its data can be overwritten. Wear-leveling algorithms are used to manage blocks changes such that no one block is updated too often. A 1:1 (Read:Write) workload on SSD/Flash drives would cause a drive failure much sooner than a 5:1 workload. Of course with 146GB of NAND flash, effective wear-leveling, and small memory blocks it would take a long time to generate 100K erase-write cycles to every memory block. Furthermore, EMC and STEC, their drive OEM, have done extensive work to mitigate this problem by providing advanced wear-leveling algorithms.

On the other hand, SSD/Flash's sequential throughput drives larger memory block sizes. SSD/Flash has historically yielded poor sequential throughput but by using a larger memory block and multiple channels to access data in parallel, good performance can be obtained. The drive is rated at 200MB/s read and 100MB/s write. STEC's website did not provide detailed internals on the drive but their sequential performance indicates a relatively large (>64K byte) erase-write memory block with 4 or more parallel paths.

Silverton Consulting
Strategy, Storage & Systems
P.O. Box 492|Broomfield, CO 80038-0492|
720-221-7270|info@silvertonconsulting.com

© 2008 Silverton Consulting, Inc.
All Rights Reserved

What this means for DMX SSD/Flash drive lifetimes is that it is EMC's problem to solve. SSD/Flash drives have the same warranty/maintenance coverage as any other DMX component and any drive failures will be relatively transparent to the customer.

DMX Virtual (thin) Provisioning

EMC has followed the "multi-pool" approach to thin provisioning, where customers define multiple storage pools and assign thin devices to each pool. Thinly provisioned devices consume storage only when written. As such, storage pools can run out of space and must be monitored for space consumption. Once a storage pool is out of capacity, all new writes to LUNs in the pool will **fail -** this is true of any thin provisioning. EMC provides on-going monitoring of storage pools with warning/alert space thresholds and script automation. Nonetheless, space must be watched much more closely than before.

Thinly provisioned LUNs spread data across all the drives in a pool. This wide striping of data improves the performance of most workloads.

Also, EMC's thin provisioning improves the performance of data replication. For example, TimeFinder/Snap, and SRDF both synchronous and asynchronous, Data Mobility and Adaptive copy all now support "thin-to-thin" replication, meaning that only written data needs to be replicated. In addition, thin provisioning as well as SSD/Flash are now available to be used by all the other key features and facilities of DMX.

Announcement significance

Of the two announcements, the more significant one is the SSD/Flash drive availability on DMX. Flash technology is rapidly advancing in capacity but write performance/erase-block cycles have remained relatively flat. The erase before write will always limit SSD/Flash applicability to specific workloads. EMC is betting that these workloads exist with an overriding need for very low response times. For such workloads, the relative expense of SSD/Flash drives can be easily justified. Moreover, flash technology, like all semi-conductor technology, is becoming cheaper to manufacture on a capacity basis so this relative expense should shrink over time.

Some in the industry believe that SSD/Flash and thin provisioning are complementary features. In contrast, SCI feels that while thin provisioning can more effectively use the expensive/high performance SSD storage, it's not very conducive or necessary for SSD/Flash. The wide striping performance advantage of thin devices does little for SSD/Flash drives whose response times are measured in microseconds. However, thin provisioning for SSD/flash can also widen the effective "wear-leveling" from a single drive to a pool of devices. In the end each capability is worthwhile alone, as separate features that further advance EMC's DMX product line, and by coupling them together, one can gain additional leverage.

Silverton Consulting
Strategy, Storage & Systems
P.O. Box 492|Broomfield, CO 80038-0492|
720-221-7270|info@silvertonconsulting.com

© 2008 Silverton Consulting, Inc.
All Rights Reserved