Method to Establish a High Availability and High Performance Storage Array in a Green Environment

Dr. M. K. Jibbe
Director of Quality Architect Team, NetApp APG
mahmoudj@netapp.com

Marlin Gwaltney
Quality Architect, NetApp APG
marling@netapp.com

Abstract

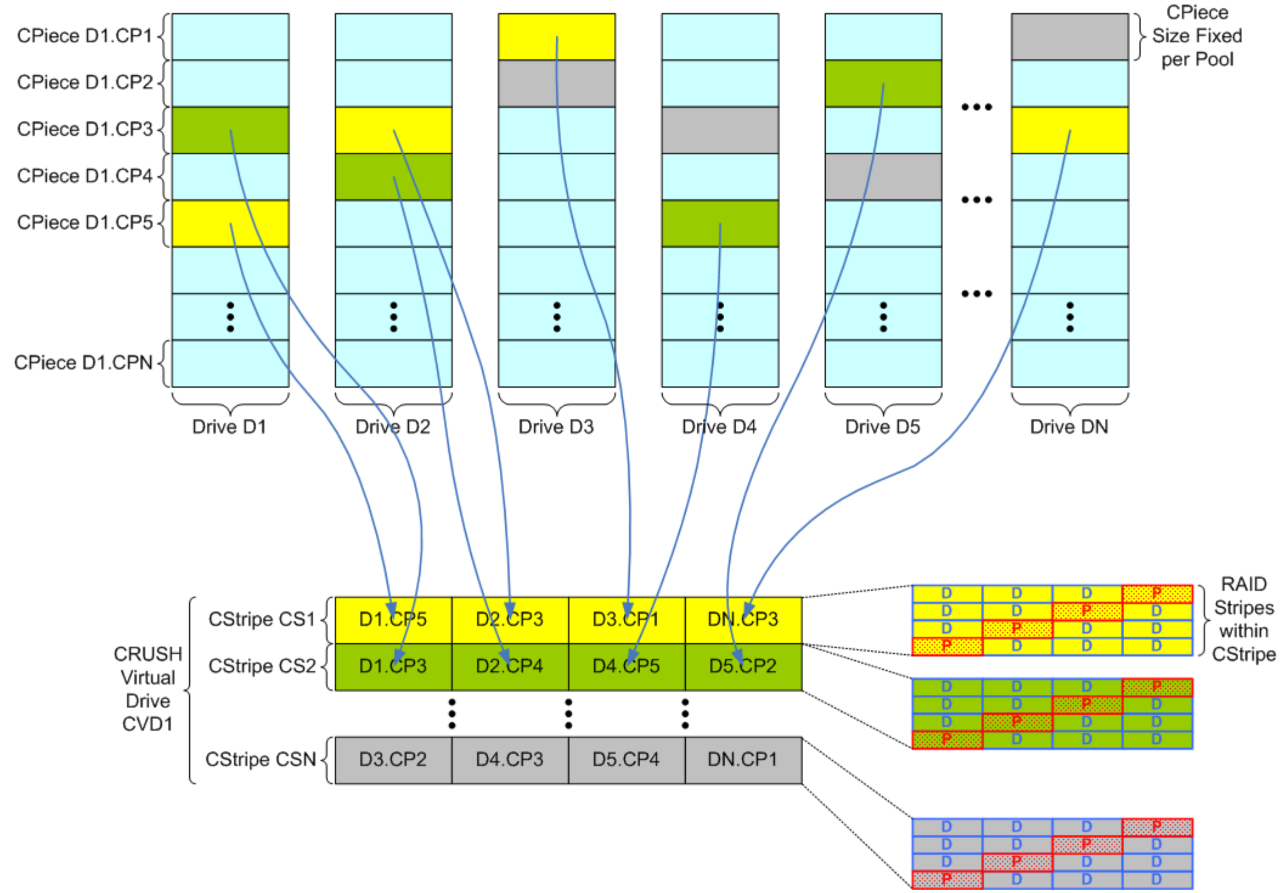
- Individually, the following features offer the user flexibility, performance, etc.:
  - Dynamic Disk Pools (DDP) using CRUSH
  - SSD Flash Read Cache
  - SSD Disk Groups
- The individual advantages of utilizing SSD Drives and/or Dynamic Disk Pools in a storage array are further magnified when they are integrated together to provide a higher availability/reliability, higher performance, and more environmentally-friendly system.
- These 3 technologies can be leveraged together to build an extremely flexible, cost effective, scalable storage solution.

Dynamic Disk Pools Using CRUSH

- CRUSH (Controlled Replication Under Scalable Hashing)
  - Algorithm developed by University of California at Santa Cruz as part of the Storage System Research Center
  - Uniformly distributes data/parity across pool of storage in a pseudo-random, but repeatable, fashion
    - Takes traditional RAID volume & distributes across larger pool of drives
    - Algorithm is deterministic, so independent nodes can access data in the same location to support N-way mirroring/failover
  - For drive failures, the reconstructed data is written into unused/reserved space on drives across the entire disk pool instead of Hot Spare drive(s)

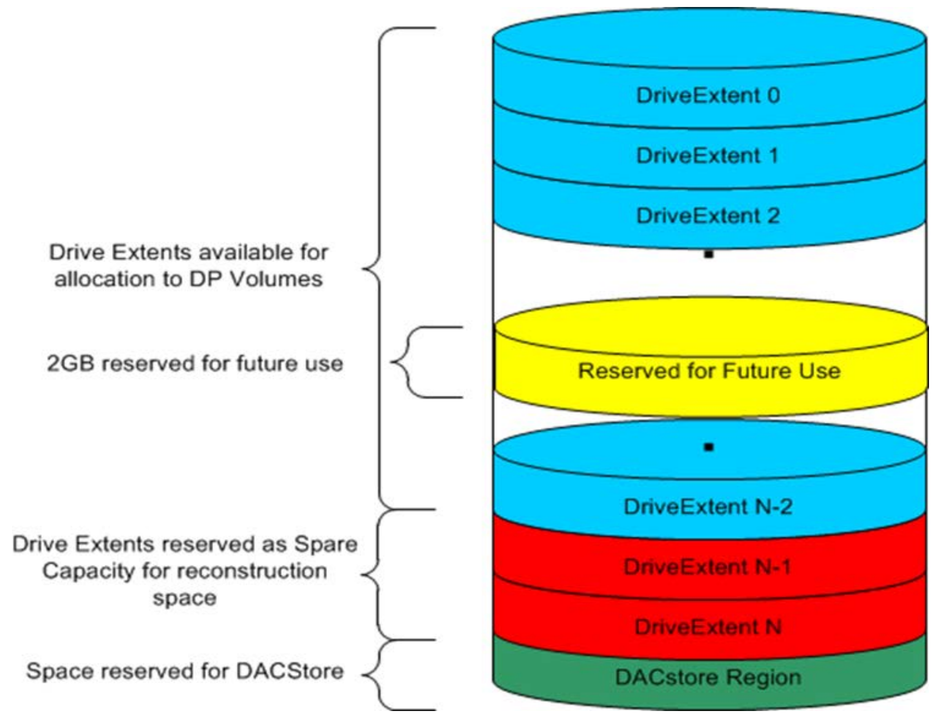
DDP Data/Parity Mapping

- “Chunks” of User Data and Parity information distributed across subset of drives in disk pool
  - Instead of sequentially striped across all drives with standard RAID algorithms



DDP Drive Extent Utilization

- DDP drive mapping includes extents for Disk Pool volumes and reserved space for spare capacity used during reconstruction after drive failure(s)



Benefits of DDP

- Simplified Configuration Management
  - Dynamic Disk Pools include flexible configuration options (from only one pool with all drives to multiple pools in a system) to optimize the system for end user requirements
  - Removes complicated management for :
    - Determining volume group RAID level and drive width
    - Managing hot spares for handling drive failures
- Increased I/O Performance
  - Every LUN is spread across every drive in the pool
  - System throughput is consistent because whether one LUN or 100 LUNs are active, every drive is contributing
    - Not limited by number of spindles in drive group

Benefits of DDP

- Increased Storage Capacity Flexibility
  - Thin Provisioning functionality allows DDP volumes to dynamically grow with demand
- Increased Fault Tolerance
  - With the distributed data/parity mapping and reserved capacity for reconstructing data, DDP can withstand ***n*** number of drive failures without the need to replace the failed drives

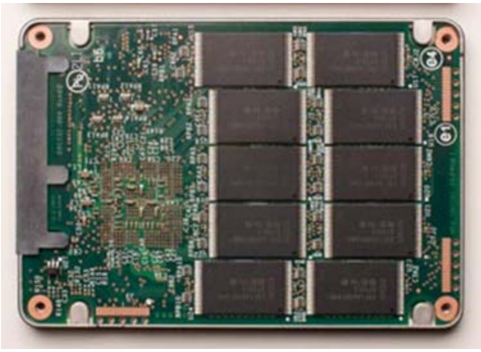
DDP Benefits – Faster Rebuilds

- Standard RAID
  - Drive failure recovery can take days/weeks with 2TB/3TB drives
  - With roadmaps from drive vendors showing 24TB drives, recovery could literally take months
  - RAID I/O performance may drop 30% to 50% during rebuild
  - The system is susceptible to additional drive failures
- Faster reconstruction/rebuild times with DDP
  - Parallel operations - uses every drive in the pool for the intensive process of rebuilding data due to drive failure(s)
  - Up to 8 times shorter than standard RAID reconstruction
  - Improved data protection and system less vulnerable since in degraded state for shorter period of time
  - Lower I/O performance penalty with shorter rebuild periods

Drive Technology Comparisons

- Hard Disk Drive – HDD
  - Uses mechanical components
  - Rotates platters on a spindle
  - Data read from or written to the platters performed by the actuator & magnetic head

- Solid State Drives - SSD
  - Uses Solid State electrical components instead of mechanical components
  - Solid State NAND-based Flash memory components mounted on the circuit board



Drive Technology Comparisons

- SSD vs. HDD drives
  - Electrical vs. Mechanical components
    - Faster startup time (no mechanical spin up)
    - More durable/shock-resistant & reliable
    - “Greener” - Lower power/heat/cooling requirements
    - Lighter in weight
    - Quieter
  - Significant I/O Performance Improvement
    - Lower access time than HDDs, <u>especially reads</u>
    - Write speeds vary among SSDs, but Enterprise-level SSDs are much faster than HDDs
  - Cost Efficiency
    - Still cost more per GB than HDDs
    - But cost less for IOPs than HDDs

SSD Flash Read Cache

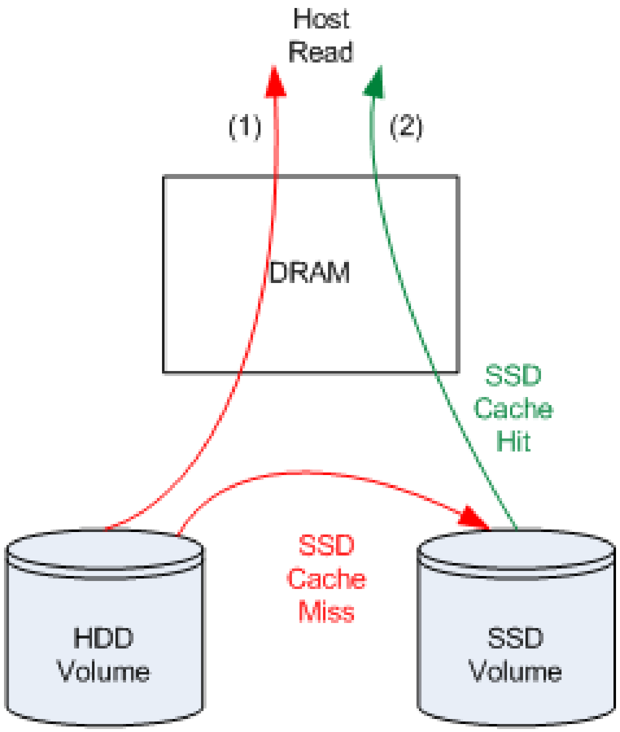
- Use SSD Flash Read Cache feature to
  - Optimize the investment in the usage of solid state drives
  - Simplify the management to achieve tiered performance
  - Manage the "hottest" read requests in the system dynamically for continuous optimal performance

SSD Flash Read Cache

- Uses Solid State Drives to improve read performance for selected volumes that reside on Hard Disk Drives
  - Secondary cache vs. primary cache (DRAM)

- The SSD cache moves data from an HDD volume to the SSDs following a host read or write

- Subsequent host reads of the same LBAs can be read directly from the SSDs with a <u>much lower response time</u> than re-reading the data from the HDD volume



SSD Flash Read Cache

- Workload characteristics that benefit from SSD Cache:
  - Performance limited by HDD IOPs
  - High percentage of read commands relative to write commands
  - Large number of read commands that are repeated to the same or adjacent areas of the drive
  - The working size set that is repeatedly accessed is smaller than the SSD cache capacity.

SSD Flash Read Cache

- Workload characteristics that benefit from SSD Cache:
  - Performance limited by HDD IOPs
  - High percentage of reads relative to writes
  - Large number of reads that are repeat reads to the same or adjacent areas of the drive
  - The working size set that is repeatedly accessed is smaller than the SSD cache capacity.

SSD Flash Read Cache

- Workload characteristics that benefit from SSD Cache:
  - Performance limited by HDD IOPs
  - High percentage of reads relative to writes
  - Large number of reads that are repeat reads to the same or adjacent areas of the drive
  - The working size set that is repeatedly accessed is smaller than the SSD cache capacity.

All SSD/Flash Storage

- Using the increased performance, environmental, and cost effective benefits of SSD drives, systems needing the highest I/O performance solution can be configured using the following options:
  - All SSD/Flash Volume groups in mixed SSD/HDD system
  - All SSD/Flash Array with all SSD-based volume groups

Combining DDP & SSD Cache/Drives

- Individual advantages of using Dynamic Disk Pools & SSD Drives further magnified when integrated together:
  - Simplified configuration management with Disk Pools with all or mixed SSD drive usage
    - Pools are configured based on the drives' Quality of Service capabilities by putting drives with different architecture, drive speed, etc. into different Pools.
    - With elimination of Hot Spare drives, all of the expensive SSD drives are used in the distributed Disk Pool operation instead of some SSD drives being underutilized waiting in the spare standby mode for a drive to fail

Combining DDP & SSD Cache/Drives

- Higher I/O performance of Disk Pools further increased when integrated with faster SSD drives using SSD Flash Read Cache and/or all SSD/Flash volumes
- Shorter rebuild times of Disk Pools further increased with usage of higher performing & reliable SSD drives for reconstructing data
  - System restored back to optimal state quicker
  - Eliminate I/O performance drop associated with degraded/reconstruction mode sooner

Combining DDP & SSD Cache/Drives

- “Greener” and more environmentally-friendly system when using Disk Pools with SSDs in the same footprint as a system with the mechanical HDDs
  - SSDs use less power…
  - …So they generate less heat and require less system cooling

Ultimate Tiered Storage Use Case

These 3 technologies can be leveraged together to build an extremely flexible, cost effective, scalable storage solution:

- For the ultimate solution in tiered storage, a single array could be configured with 5 separate disk pools/groups of mixed drive types:
  - Lowest Performance/Cost: Disk Pool of SAS/SATA Near-Line drives
  - Lower Performance/Cost: Disk Pool of 10K SAS drives
  - Standard Performance/Cost: Disk Pool of 15K SAS drives
  - Drive Group with SSD drives for SSD Flash Read Cache to give a Read I/O performance boost for selected volumes in the previous three Disk Pools
  - Highest Performance: Pool including all SSD drives

Integrated Feature Performance

I/O performance comparison for Disk Pool systems using SSD vs. HDD drives:

- System Configuration:
  - SSD: 1 Disk Pool with 20 SSD drives & 8 volumes
  - HDD: 6 Disk Pools of 60 HDD drives & 6 volumes each
- Read I/O performance results for 4K random I/O:
  - SSD with Read caching ~3 times better than HDD with read caching
  - SSD without Read caching ~4 times better than HDD with read caching (*faster bypassing the RAID caching & going directly to drive*)
  - SSD I/O Average Response Time is ~10-14 times better than HDD drives with read caching

Integrated Feature Performance

I/O performance comparison for Disk Pool systems using SSD vs. HDD drives:

- System Configuration:
  - SSD: 1 Disk Pool with 20 SSD drives & 8 volumes
  - HDD: 6 Disk Pools of 60 HDD drives & 6 volumes each
- Write I/O performance results for 4K random I/O:
  - SSD with Write Caching/Mirroring ~2 times better than HDD with write Caching/Mirroring

Integrated Feature Performance

Rebuild comparison for Disk Pool systems using SSD vs. HDD drives:

- HDD Disk Pool rebuild times are ~2-4 times faster than comparable RAID 6 volumes depending on single or dual drive failure scenarios and 4K random I/O or 512K sequential I/O workloads
- SSD Disk Pool rebuild times are ~2-4 times faster than comparable HDD Disk Pool rebuild times depending on single or dual drive failure scenarios and 4K random or 512KB sequential I/O workloads

Integrated Feature Environmental

Power comparison for Integrated DDP and SSD systems:

- By comparison, SSD drives use ~33% less power than HDD drives (~2W vs. ~6W)
- For comparable systems with the same footprint using SSD drives vs. HDD drives, the power difference for the following drive configurations would be:

Number of Drives	Power Consumption Difference (W)
12	~48
24	~96
60	~240

Summary

- Key benefits of individual DDP & SSD storage features:
  - Dynamic Disk Pools
    - Simpler Configuration Management
    - Increased I/O performance with distributed data management across all the drives in the larger groups
    - Likewise, quicker Rebuild times with distributed data/parity across all the drives in larger drive groups & no dedicated hot spare drives
  - SSD Drives vs. HDD Drives
    - SSD Drives are faster, more reliable/durable, "Greener" (less power/heat/cooling), more cost effective (IOPS), lighter, & quieter
    - SSD Flash Read Cache provides faster Read I/O performance for repeated Read commands
    - Use all-SSD/Flash drive groups or systems for storage solutions requiring the highest I/O performance and to take full advantage of SSD technology

Summary

- Key benefits when combining DDP and SSD-based storage features:
  - Increase I/O performance of Disk Pools using SSD Flash Read Cache or All-SSD drive groups
  - Reduce Reconstruction Rebuild Time of Disk Pools further by using All-SSD drive groups
  - Configure "ultimate" Tiered Storage system using mix of Disk Pools with different classes of HDD drives, SSD/Flash Read Cache to increase performance of these Pools, and all SSD-drive Disk Pool for the highest performance tier
  - Configure "Greener" systems using DDP & SSD drives to lower power consumption/heat and save on cooling in the same footprint area

Conclusions

- The individual Disk Pool and SSD drive storage features provide advantages in configuration management, I/O performance, and system environmental factors.
- However, when users combine these features together, the integrated solution extends and enhances these feature improvements.
- When the Dynamic Disk Pool, SSD Flash Read Cache, and/or all Flash/SSD drive group features are integrated together, they provide a scalable, higher availability/reliability, higher performance, and more environmentally-friendly system than systems using these features individually or not at all.

Thank you