Benefits of Storage Capacity Optimization Methods (COMs)

And

Performance Optimization Methods (POMs)

Herb Tanzer & Chuck Paridon

Storage Product & Storage Performance Architects

Hewlett Packard Enterprise

Source:
The Green Grid Forum 2016
The Economics of Sustainable IT
"Software Defined Storage Energy Efficiency Features"

Topics

- Green Storage Metrics
- Capacity Optimization Methods (COMs)
  - Features Overview
  - Test Characterization
    - Test vehicle description
    - Space and power savings
    - Performance impact
  - Benefits: CAPEX, OPEX
- Performance Optimization Methods (POMs): brief

Green Storage Metrics

- Metrics are based on SNIA Emerald™ and EPA ENERGY STAR® definitions of storage power efficiency
  - IDLE: Capacity (Idle GB/Watt)
  - ACTIVE: Transactional Performance (IOPs/Watt) and Streaming Performance or Bandwidth (MBPS/Watt)

    Note: Active workloads use the Oracle VdBench™ IO generator
- Data is collected in 1-minute intervals for the measurement period (2 hours idle & 30 minutes active) and averaged
- These metrics are used for “Acquisition” decisions
  - Manufacturer runs Emerald “benchmark” workloads and measures metric values in the lab
  - Customer compares per taxonomy type and by peak metric value
- TGG’s DCsP uses the same metrics, but measure operational datacenter workloads
  - Operational effectiveness (productivity) is measured as Capacity or Performance per kWh

Capacity Optimization Methods (COMs)

COMs are software techniques that allow more data to be stored on less capacity (i.e., use fewer disks and less power)

COM	Description	Potential Estimated Space Savings
RAID 5/6, Erasure Encoding	Data integrity w/o using full copies	When replacing Mirroring (RAID 1): ~40%
Thin Provisioning	Virtualize allocation of Storage	Take from 30% utilization (legacy) to ~ 80%
Deduplication	Eliminate duplicate files or data chunks	Can be 25-40% primary, up to 50% secondary
Compression	Reduction of data sets within files or blocks	2:1 compression is done routinely
Delta Snapshots	Reconstruct different versions of a file w/o multiple full copies	Large savings possible when change delta is small

The status of COM testing

- Currently, the SNIA Emerald Specification defines simple pass/fail tests that verify the existence and activation of particular COMs
- Currently, Storage ENERGY STAR requires that for the eligible categories:
  - Parity RAID is enabled in the System Under Test (SUT)
  - 1 out of 4 specific COMS is available as a selectable feature and needs to be verified for existence, but they are to be disabled during idle & active metric testing
- There has been some estimation of active COM benefits, but more test characterization is needed to establish valid methods to accurately account for their capacity-saving efficiency in idle & active testing
  - *Some early characterization (tested and estimated) of COM benefits is presented here*

Test Vehicle: COM

- High-mid storage array
  - Dual controller node pairs
  - Integrated disk devices
  - Expansion disk enclosures
  - We used 160 SFF SAS drives 450GB x 15k

COM Testing Procedure

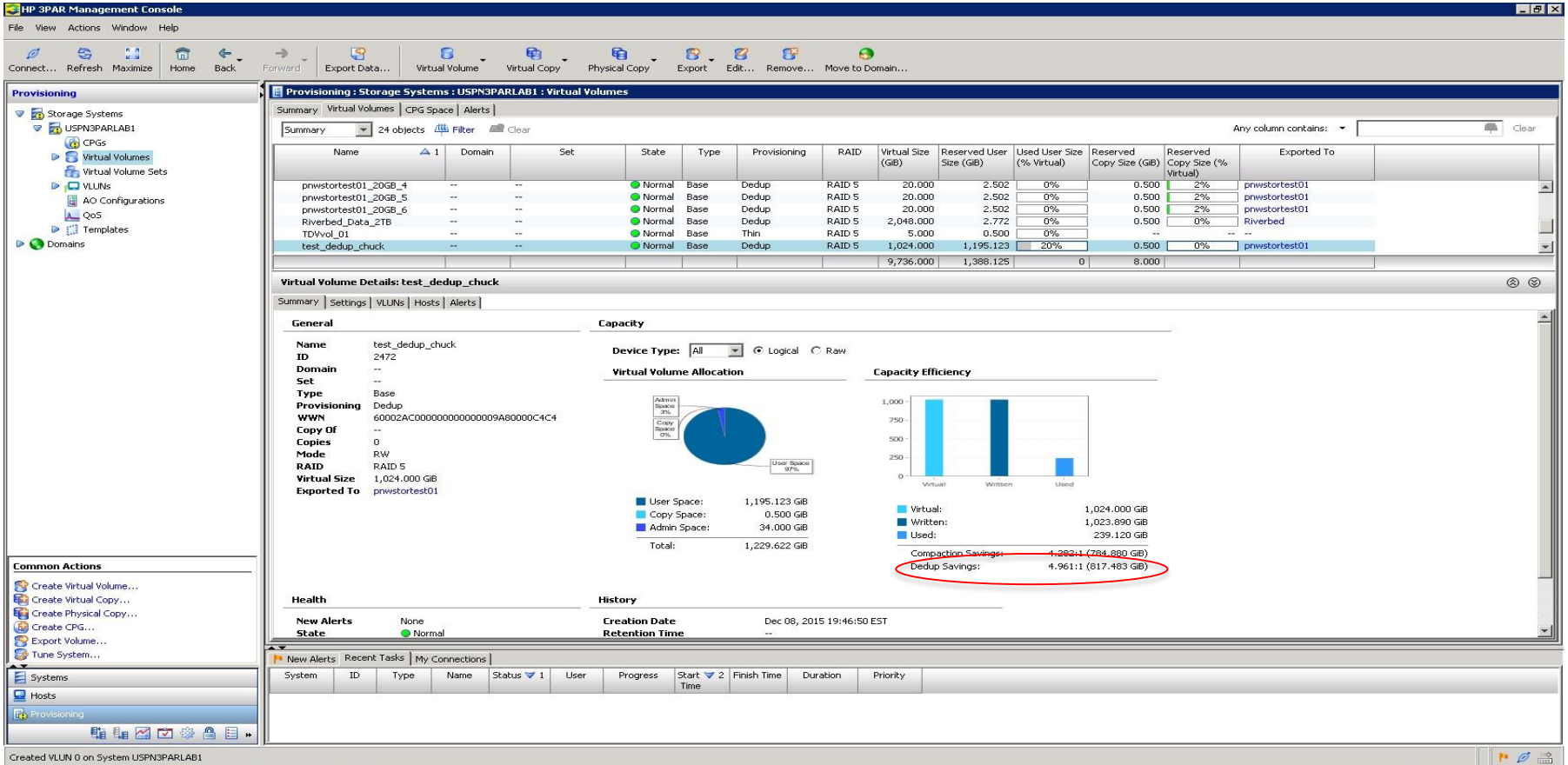
1. Create Baseline Configuration
   - Raid 1 - basis for comparing Parity Raid, Compression (2:1 ratio)
   - Raid 5 - basis for comparing Data dedup and Thin Provisioning (2:1 dedup, OP)
2. Record the Baseline Characteristics
   - Usable Capacity
   - Power Consumption
   - Active IO rate
3. Observe/Record the Optimized Characteristics from Enabling the COM
4. Adjust the Physical Capacity to Create Equivalent Usable Capacity w/Baseline

   Note: could also hold physical capacity constant & optimize the logical capacity
5. Record the Physically Optimized Characteristics

Dedupe Enabled Emerald Script File (Example)

```
###################
#
#  Single Lun Dedup Tests - fill
###################

dedupratio=5
dedupunit=16k

# concatenatesds=yes

sd=sd1,lun=\\.\PhysicalDrive10

#Hotband workload

wd=HOTwd_uniform,skew=6,sd=sd*,seekpct=rand,rdpct=50
```

Dedupe Enabled Storage Array Screen Shot (Example)



Energy Efficient Storage: SNIA Emerald Program Specific COM Examples

(70/30 R/W random workload) Based on **Physical Capacity Reduction**

COM Type	Raid Parity	Data Compression
Baseline Config	Raid 1 Mirrored and Striped	Uncompressed
Baseline Capacity (GB)	36,000 (160 – 450 GB disks)	36,000 (160 – 450 GB disks)
Baseline Power (W)	2360(I) 3134(B)	2360(I) 3134(B)
Baseline (GB/W)	15.25	15.25
Baseline Perf. (IOPs)	36,010	36,010
Baseline (IO/W)	11.49	11.49
Optimized Config	Raid 5 (3+1 Data + Parity)	Compressed (2:1)
Optimized Capacity (GB)	36,000 (107 – 450 GB disks)	36,000 (80 – 450 GB disks)
Opt. Power (W)	1564(I) 2200(B)	1464(I) 2030(B)
Optimized (GB/W)	23.02	24.59
Delta (GB/W)	+51%	+61%
Opt. Perf (IOPs)	17,040	8,235
Optimized (IO/W)	7.75	4.06
Delta (IO/W)	(33%)	(64%)

Energy Efficient Storage: SNIA Emerald Program Specific COM Examples

(70/30 R/W random) Based on **Physical Capacity Reduction**

COM Type	Thin Provisioning	TP + Data Dedup
Baseline Config	**Raid 5 (3+1 Data + Parity)**	**Full Provisioned No Dedup**
Baseline Capacity (GB)	54,000 (160 – 450 GB disks)	54,000 (160 – 450 GB disks)
Baseline Power (W)	2360(I) 3134(B)	2360(I) 3134(B)
Baseline (GB/W)	22.88	22.88
Baseline Perf. (IOPs)	25,250	25,250
Baseline (IO/W)	8.06	8.06
Optimized Config	**Thin Provisioning (50% written)**	**TP+Deduped Patterns (2:1)**
Optimized Capacity (GB)	54,000 (50% written)	54,000 (50% written-dedup 2:1)
Opt. Power (W)	1464(I) 2030(B) (80 disks)	1016(I) 1478(B) (40 disks)
Optimized (GB/W)	36.88	53.15
Delta (GB/W)	+61%	+132%
Opt. Perf (IOPs)	12,560	4,710 (2:1 Dedup Ratio)
Optimized (IO/W)	6.19	3.19
Delta (IO/W)	(23)%	(60%)

Summary of COM Benefits

COM Feature	Physical capacity savings	Idle / active power savings	IOPs (70/30 R-W Workload)
Parity RAID	33%	34% / 30%	0.47 X
Compression	50%	38% / 35%	0.25 X
Thin Provision	50%	38% / 35%	0.50 X
Thin Provision + Deduplication	75%	57% / 68%	0.20X
Snapshot	not tested		

Note: capacity savings = fewer drives

CAPEX Savings Estimated

COM Feature	Fewer drives	Cost savings* Tested 3.1 kW	Cost savings Scaled to 10kW – 100kW
Parity RAID	53	$34.5K	$110K -- $1100K
Compression	80	$49.3K	$157K -- $1570K
Thin Provision	80	$49.3K	$157K -- $1570K
Thin Provision + Deduplication	120	$71.3K	$228K -- $2280K

***Estimated customer costs for disk drive hardware:**
$549 for a 450GB 15K SAS drive
$2710 for an empty drive enclosure

OPEX Savings Estimated

COM Feature	Less installed power	Cost savings* Tested 3.1kW	Cost savings* Scaled to 10kW – 100kW
Parity RAID	934W	$497/mo	$1585/mo -- $15850/mo
Compression	1104W	$587/mo	$1874/mo -- $18740/mo
Thin Provision	1104W	$587/mo	$1874/mo -- $18740/mo
Thin Provision + Deduplication	1656W	$881/mo	$2811/mo -- $28110/mo

***It costs $5,320 per month per 10-kW installation for data center power, cooling, and space (based on a first-order approximation for a typical data center by Patel and Shah at HP Labs)**

Summary

For the most part, the COMs can deliver capacity and power efficiency, but at a performance penalty
[I.e., the idle Capacity/Power metric will improve, but when subjected to an active workload, the performance/power metric will be reduced]

Note: these are COMs not POMs (Performance Optimization Methods)

References

1. SNIA Emerald™ program, *SNIA Emerald™ Power Efficiency Measurement Specification,* http://snia.org/tech_activities/standards/curr_standards/emerald
2. U.S. Environmental Protection Agency, *ENERGY STAR® Program for Data Center Storage Specification*, www.energystar.gov/index.cfm?c=new_specs.enterprise_storage
3. The Green Grid, White Paper #58 - The Green Grid Data Center Storage Productivity Metrics (DCsP) http://www.thegreengrid.org/en/Global/Content/white-papers/WP58-DCsP
4. C. Patel and A. Shah, Cost Model for Planning, Development and Operation of a Data Center, HP Labs (2005)

the green grid®

® get connected to efficient IT

POM Backup Slides

www.thegreengrid.org

Workload with Hot Band IO Patterns

- Is asymmetric; comprised of several different IO streams, some of which contain hot spots or regions of more intense IO demand
- Results in varying degrees of logical block address re-referencing in certain regions of the overall IO space that can be either contained with the cache of an array, or can be placed on storage devices that deliver a higher IO rate
- Hot bands concentrate 54% of the IO in 32% of the space
  - Therefore the opportunity exists to “speed up” a fraction of the total capacity

IO Profile	% of workload	Read/Write Percentage	IO Size (KiB)	Access Pattern	Usable Address Range
Write Stream 1	5	0/100	See Table 12	Sequential	0-100%
Write Stream 2	5	0/100	See Table 12	Sequential	0-100%
Write Stream 3	5	0/100	See Table 12	Sequential	0-100%
Read Stream 1	5	100/0	See Table 12	Sequential	0-100%
Read Stream 2	5	100/0	See Table 12	Sequential	0-100%
Read Stream 3	5	100/0	See Table 12	Sequential	0-100%
Read Stream 4	5	100/0	See Table 12	Sequential	0-100%
Read Stream 5	5	100/0	See Table 12	Sequential	0-100%
Uniform Random	6	50/50	See Table 12	Random	0-100%
Hot Band 1	**28**	70/30	See Table 12	Random	**10 -18%**
Hot Band 2	**14**	70/30	See Table 12	Random	**32-40 %**
Hot Band 3	**7**	70/30	See Table 12	Random	**55-63 %**
Hot Band 4	**5**	70/30	See Table 12	Random	**80-88 %**

Re. SNIA Emerald Specification

Hot Band enables the benefits of Storage Caching and Device Tiering

- The Hot Band workload, when run on High End Storage (typically) demonstrates the performance/power advantage of two product features
  - Array Based Cache
  - Device Tiering
- Although the initial goal was solely cache focused, there is also a benefit of implementing faster tiers in the product via Device Tiering
- The hot band workload is included in the Emerald Specification, thus enabling these two features to improve the active metrics

Hot Band Primary Active Metric Comparison

Configuration	Tier Type	Power Consumption (Modeled)	IOPS	IOPS/Watt
Large Array (Initial)	15k RPM	7491 Watts	18410	2.457
Large Array (Cache Assist)	15k RPM	7491 Watts	39900	5.326
Large Array (Cache + Tiered)	15k + SSD	7283 Watts	42870	5.886

Overall Primary Metric Improvement ~ 140% (2.40x)

the green grid®

® get connected to efficient IT

COM Backup Slides

www.thegreengrid.org

Snapshot COM Benefits

- Due to the lack of an active snapshot workload, we present the following idle capacity analysis
- Most snapshot products consist of two components
  - A base file or image created at a specific Point In Time (PIT)
  - A “delta” file - the enumeration of the changes to the base file until another PIT
- The combination of these enables the recreation of many versions of the file or image in a space-efficient manner.
- Consider the case where a 100 GB file has 50 % of it’s contents change in a week.
  - Only 150 GB is required to store both “files” – 25% physical capacity savings

the green grid®

® get connected to efficient IT

Thank you

www.thegreengrid.org