SNIA Hot Banding

Steven Johnson

SNIA Emerald™ Training

SNIA Emerald Power Efficiency Measurement Specification, for use in EPA ENERGY STAR®

July 14-17, 2014

SNIA Green Storage Initiative

SNIA Emerald™

Agenda

- Introduction to SNIA Emerald Hot Banding test
- Concatenated Work Space
- Discussion of the components of the workload
  - Random
  - Five Sequential read streams
  - Three Sequential write streams
  - Four hot bands
  - Complex mix of transfer sizes.

SNIA Green Storage Initiative

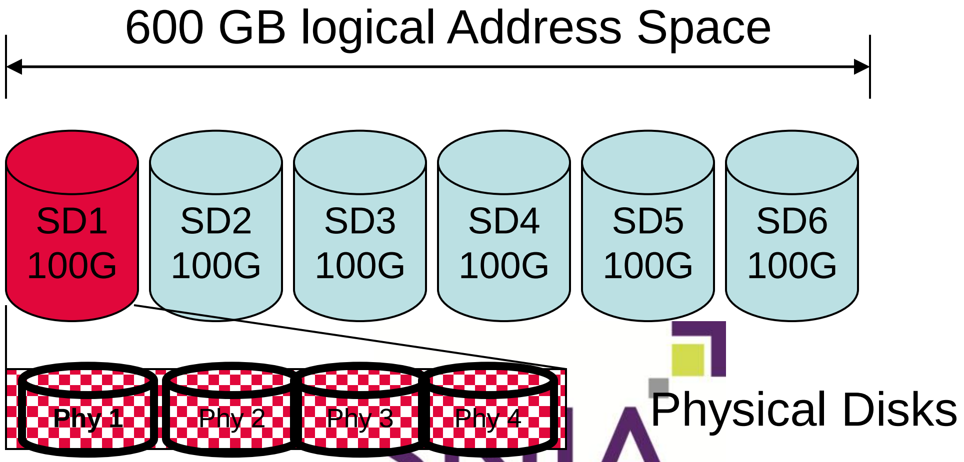
SNIA Hot Band

- Problem
  - Uniform access across storage does not address performance aspect of cache in customer-like environments
  - Customer environments do not have uniform access across storage unless they are doing a full back up
- End of the day, goal is to create a fairly complex workload to measure the power consumption – not a pure performance benchmark

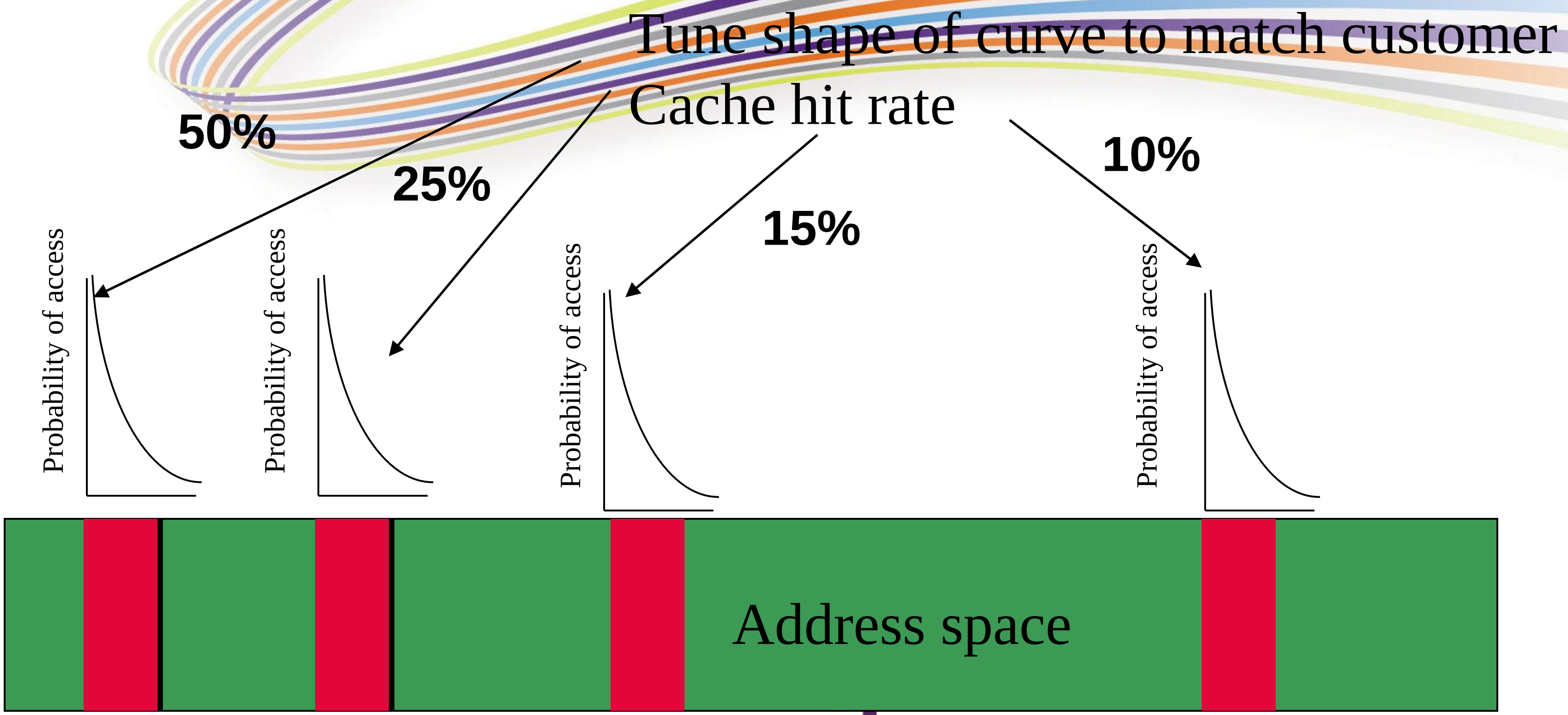
SNIA™
Green Storage Initiative

Concatenated Work Space

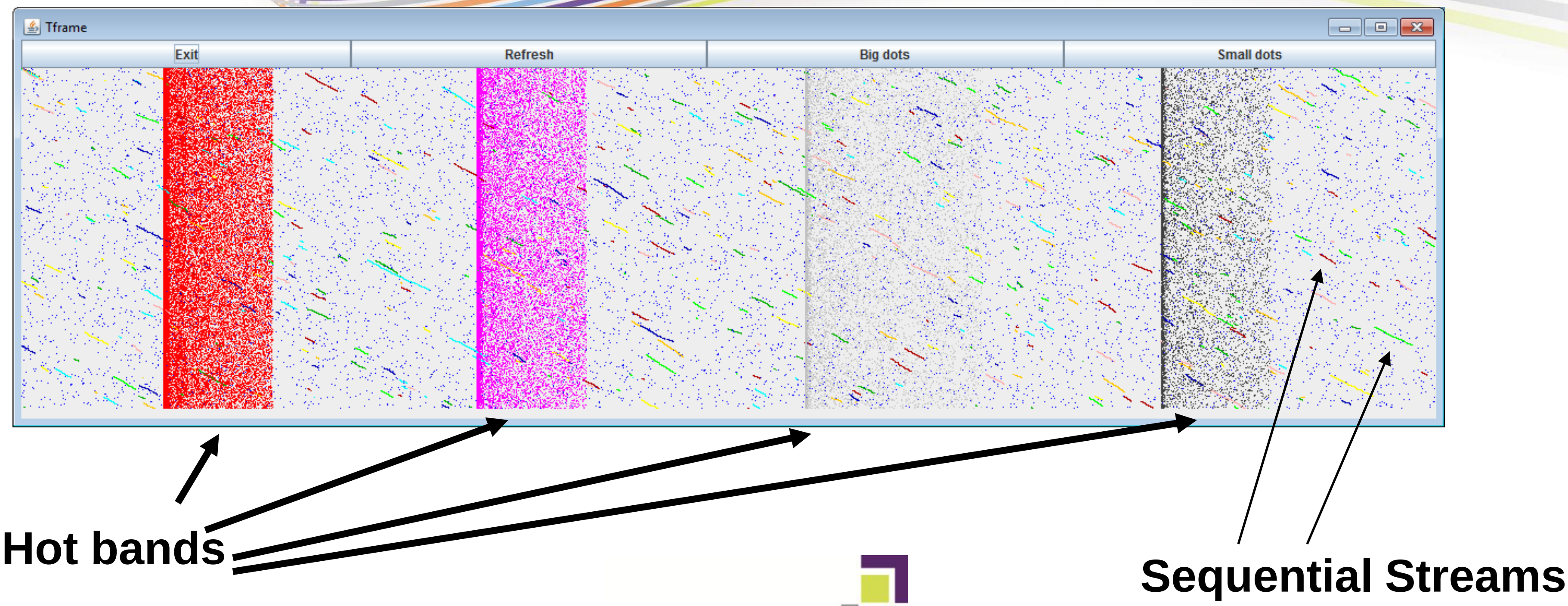
- First part of the process is the configure your storage.
- Determine your optimal configuration (Raid 1, Raid 5, ??, Stripe size, Volume Manager settings, etc)



How we could approximate Customer Workloads



Sample Scatter plot with 4 hot bands, random access and sequential streams

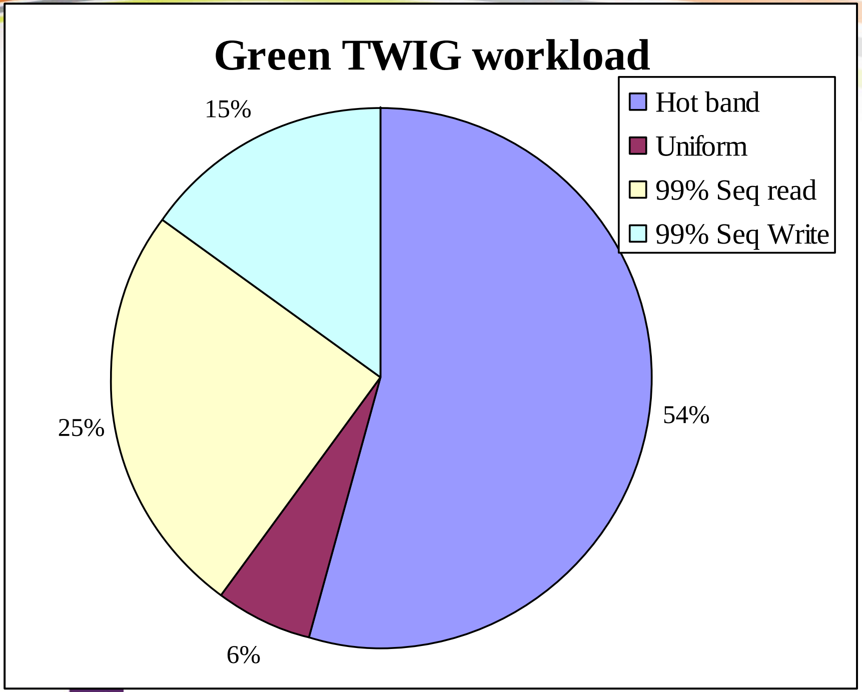


Rich mix of workload that is relatively cache friendly

- Broad mix of IOS:
  - Workload type % workload
  - Hot band 54%
  - Random 6%
  - 99% Seq read 25%
  - 99% Seq Write 15%

Workload type	% workload
Hot band	54%
Random	6%
99% Seq read	25%
99% Seq Write	15%

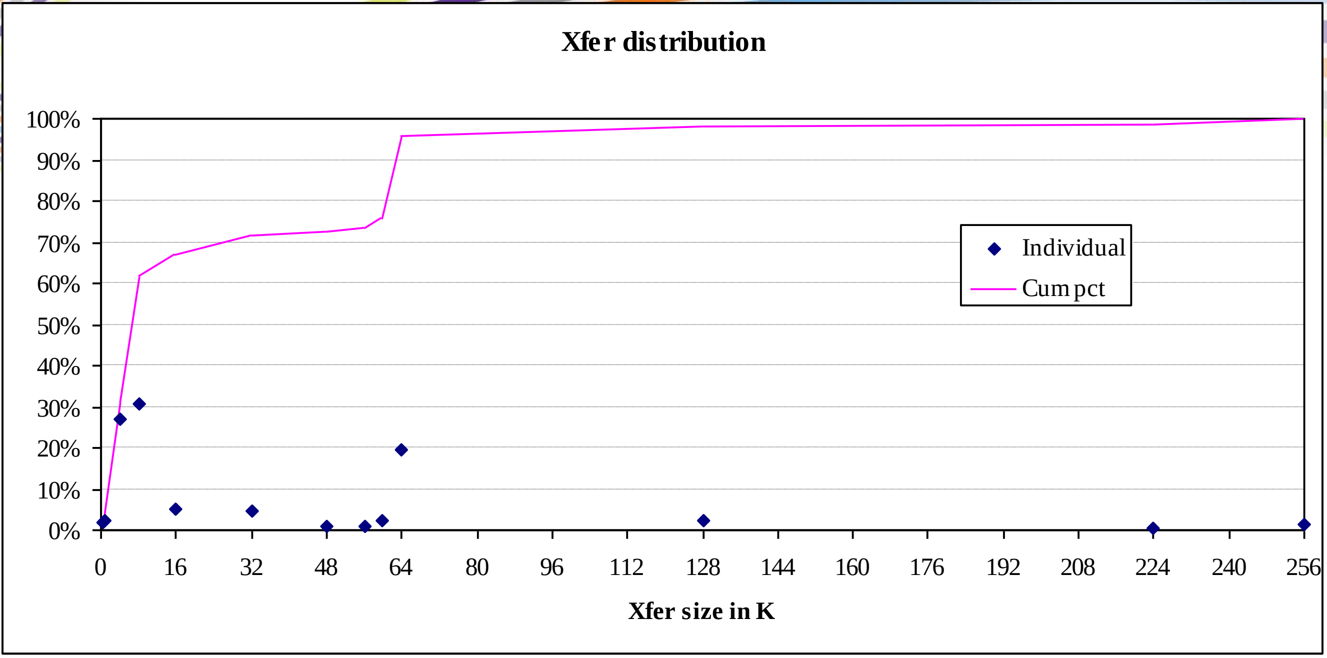
- Over all workload 66% read



Recent study of 1.1 Billion IOs 2,500+ traces Abstracted the distribution below

K	Individual	Cum pct
0.5	2%	2%
1	2%	4%
4	27%	31%
8	31%	62%
16	5%	67%
32	5%	72%
48	1%	73%
56	1%	74%
60	2%	76%
64	20%	96%
128	2%	98%
224	0%	98%
256	2%	100%

Average = ~27K



SNIA Green Storage Initiative

Vdbench script

```
compratio=2  # Compression Ratio 2:1
sd=sd1,lun=e:\junk\hotband1,size=500m  # Define Storage definition (file for testing purposes)
```

SNIA Green Storage Initiative

Vdbench script (cont)

```
wd=default,xfersize=(8k,31,4K,27,64K,20,16K,5,32K,5,128K,2,1K,2,60K,2,512,2,256K,2,48K,1,56K,1),rdpct=70
wd=wd_uniform,skew=6,sd=sd*,seekpct=rand,rdpct=50
wd=wd_hot1,sd=sd*,skew=28,seekpct=rand,range=(10,18)
wd=wd_hot2,sd=sd*,skew=14,seekpct=rand,range=(32,40)
wd=wd_hot3,sd=sd*,skew=7,seekpct=rand,range=(55,68)
wd=wd_hot4,sd=sd*,skew=5,seekpct=rand,range=(80,88)
# 5 read sequential walkers
wd=wd_99rseq1,sd=sd*,skew=5,xfersize=(8k,33,4K,29,64K,22,16K,6,32K,5,128K,3,256K,2),seekpct=1,rdpct=100
wd=wd_99rseq2,sd=sd*,skew=5,xfersize=(8k,33,4K,29,64K,22,16K,6,32K,5,128K,3,256K,2),seekpct=1,rdpct=100
wd=wd_99rseq3,sd=sd*,skew=5,xfersize=(8k,33,4K,29,64K,22,16K,6,32K,5,128K,3,256K,2),seekpct=1,rdpct=100
wd=wd_99rseq4,sd=sd*,skew=5,xfersize=(8k,33,4K,29,64K,22,16K,6,32K,5,128K,3,256K,2),seekpct=1,rdpct=100
wd=wd_99rseq5,sd=sd*,skew=5,xfersize=(8k,33,4K,29,64K,22,16K,6,32K,5,128K,3,256K,2),seekpct=1,rdpct=100
# 3 write sequential walkers
wd=wd_99wseq1,sd=sd*,skew=5,xfersize=(8k,33,4K,29,64K,22,16K,6,32K,5,128K,3,256K,2),seekpct=1,rdpct=0
wd=wd_99wseq2,sd=sd*,skew=5,xfersize=(8k,33,4K,29,64K,22,16K,6,32K,5,128K,3,256K,2),seekpct=1,rdpct=0
wd=wd_99wseq3,sd=sd*,skew=5,xfersize=(8k,33,4K,29,64K,22,16K,6,32K,5,128K,3,256K,2),seekpct=1,rdpct=0
# Actual run definition
rd=rd1_hband,wd=HOTwd*,iorate=MAX,warmup=30,elapsed=6H,interval=10,pause=30,th=XX
```

Hot Band Workloads and High End HP Storage Products

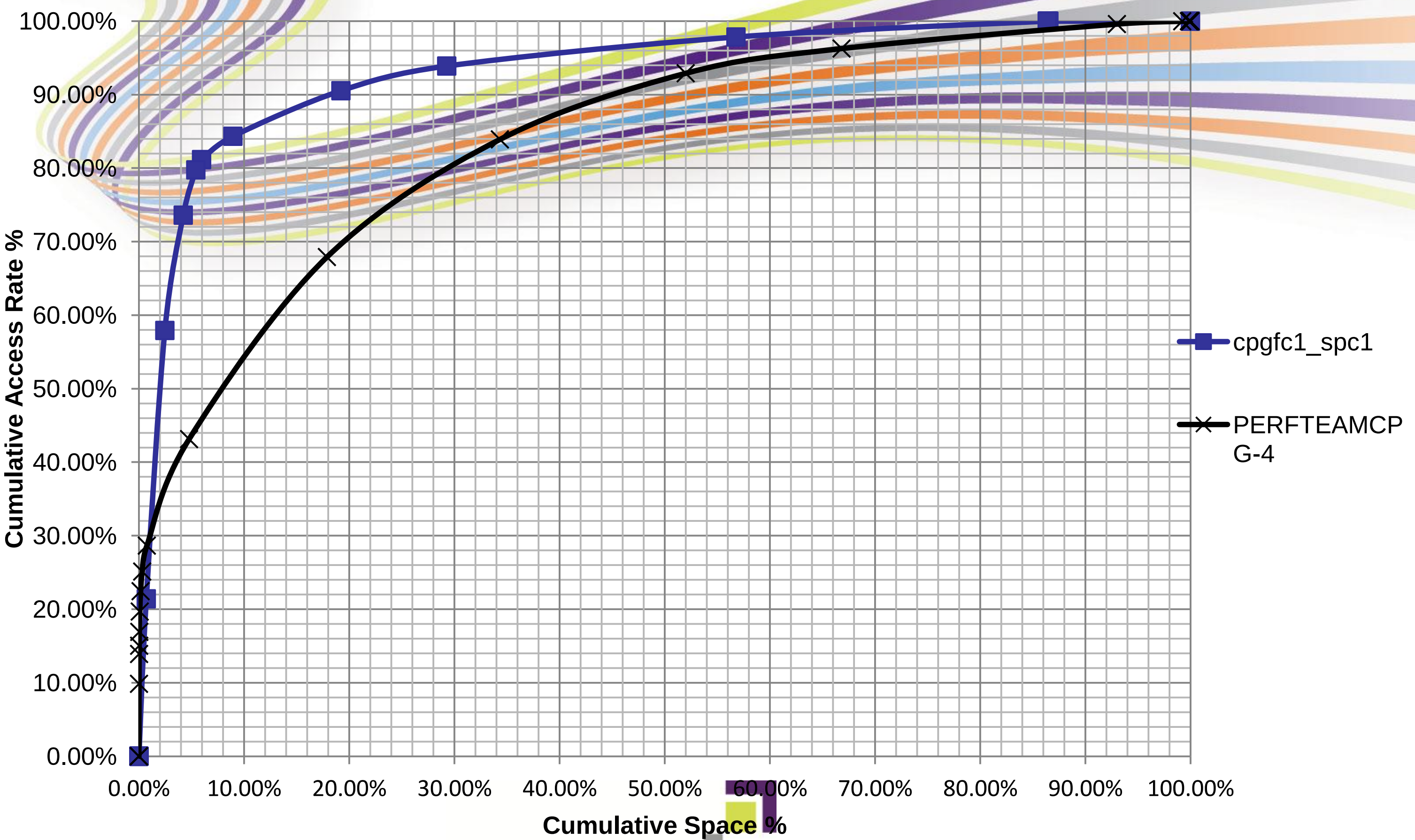
Chuck Paridon, HPES Master Storage Performance Architect

SNIA Emerald™ Training

SNIA Emerald Power Efficiency Measurement Specification, for use in EPA ENERGY STAR®

June 24-27, 2013

SNIA Green TWG Cache Friendly Performance Comparison Chart
Tiered Storage Speedup on 3PAR Storage Products



Source: HP Blue line is the SPC-1 workload IO density data. Black is Hot Band. The steeper the curve and the sharper the knee, the more Tiered Storage can enhance the IO rate. SPC speedup ~50% Hot Band ~30%
So total speedup due to both cache and tiering ~ 100 + (40 *0.3) = 112%

Hot Band IO Patterns

- The Hot Band workload is comprised of several different IO streams, some of which contain hot spots, or regions of more intense IO demand.
- This 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.

SNIA™
Green Storage Initiative

Hot Band IO Patterns

Hot Band IO Profile

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 %

Hot Bands concentrate 54% of the IO in 32% of the space

Hot Band IO Patterns

- The Hot Band workload when run on High End Storage demonstrates the power/performance advantage of two product features
  - Array Based Cache
  - Storage Tiering
- Although the initial goal was solely cache focused, there is also a benefit of implementing faster tiers in the product such as HP P9500 Smart Tiering

SNIA Green TWG Cache Friendly Performance Comparison Chart Cache and Tiering Speedup on HP P9500 Storage Product

	Hot IOPS	Hot RT	Rnd IOPS	Rnd RT	C/WS ratio	Cache Hit Ratio Hot Band vs Rnd
Small Array	4,330	32.8 ms	4,[illegible]0	[illegible]3.4 ms	<<	N/A
Large Array Cache Assist only	39,900	8.97 ms	18,410	22.[illegible] m[illegible]	[illegible]3%	60%/24%
Large Array (Tiered)	42,870	5.77 ms	N/A	N/A	~3%	60%/24%

SNIA Green Storage Initiative

SNIA Green TWG Tiered Storage w/Hot Bands Analysis and Tier Configuration Process

- The Workload Analysis Process Consists of Two Steps

1. Create a single pool large enough to hold the desired working set
2. Run the Hot Band workload on that pool and use the analysis tools to produce a report guiding the composition of the tier(s)

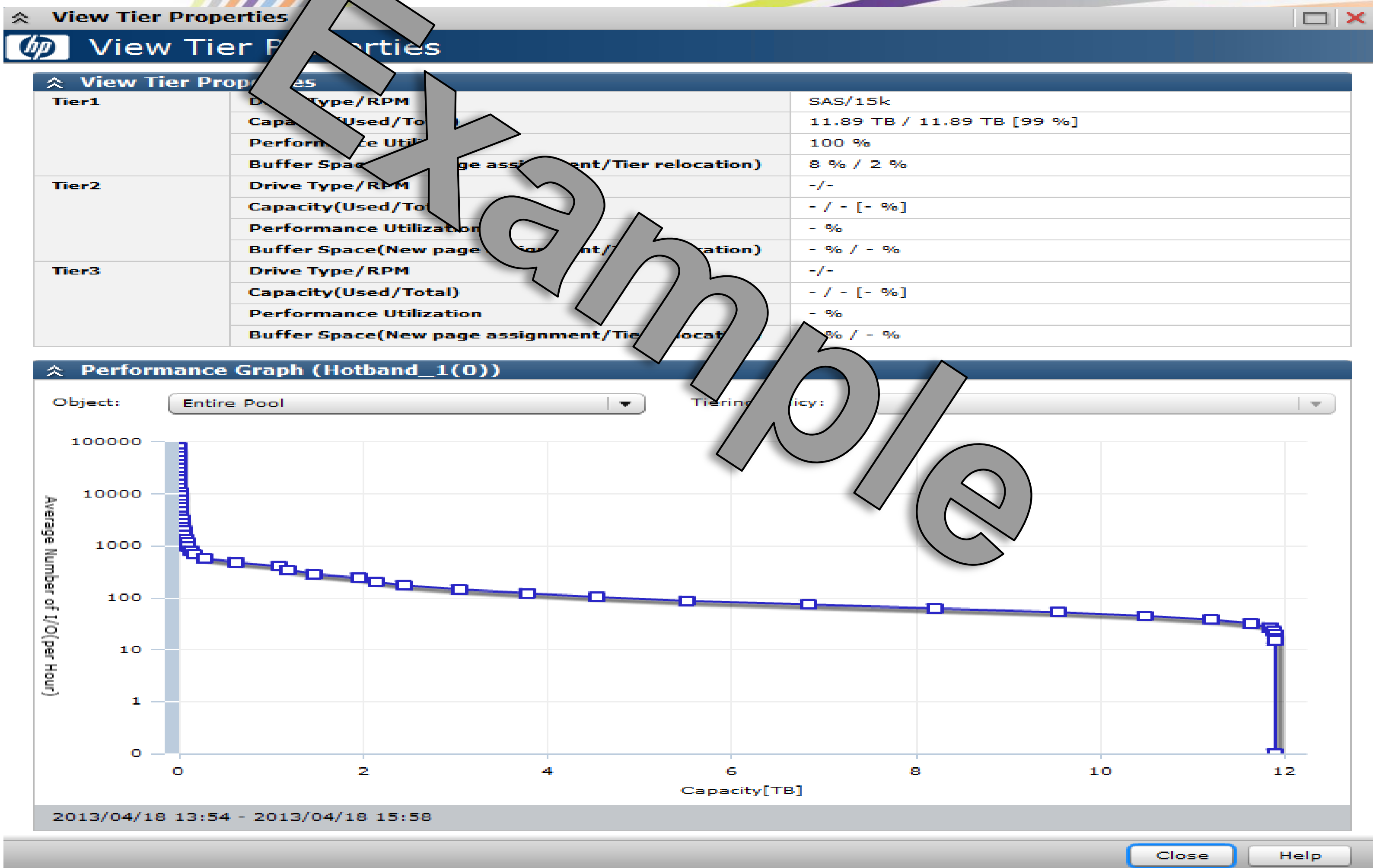
- The tier construction process involves 2 decision points.

1. Which technology to deploy
2. The capacities of the tier(s)

- The following slides illustrate an example of these activities

SNIA™ Green Storage Initiative

SNIA Green TWG Tiered Storage w/Hot Bands
Initial Tier Property Analysis



SNIA Green TWG Tiered Storage w/Hot Bands Final Tier Property Analysis

View Tier Properties

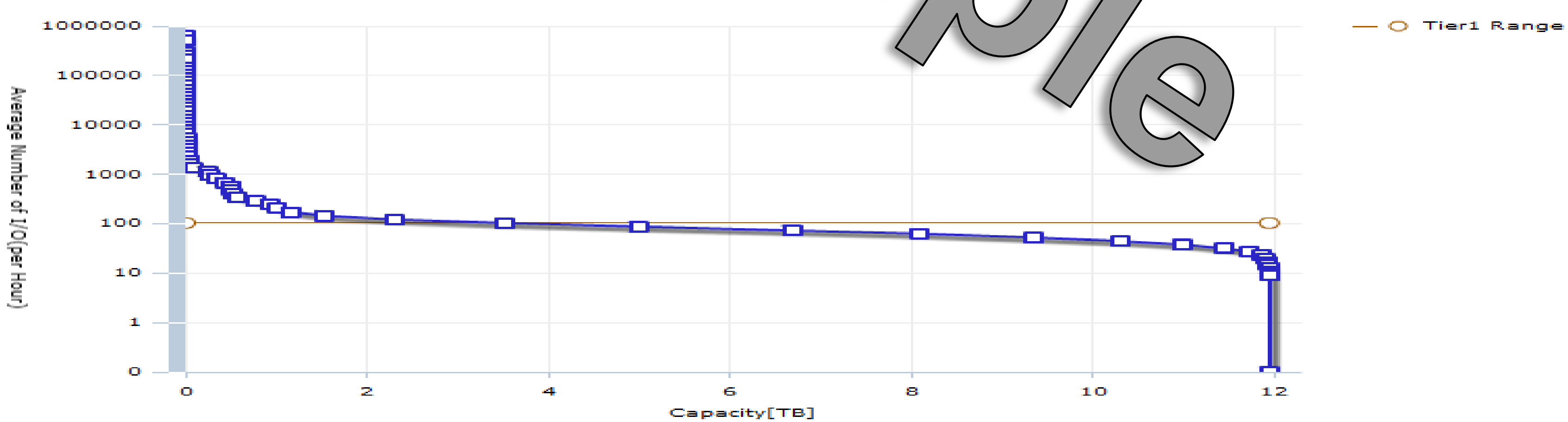
View Tier Properties

View Tier Properties

Tier1	Drive Type/RPM	SSD/-
	Capacity(Used/Total)	2.50 TB / 2.50 TB [100 %]
	Performance Utilization	74 %
	Buffer Space(New page assignment/Tier relocation)	0 % / 2 %
Tier2	Drive Type/RPM	SAS/15k
	Capacity(Used/Total)	9.43 TB / 10.06 TB [93 %]
	Performance Utilization	59 %
	Buffer Space(New page assignment/Tier relocation)	8 % / 2 %
Tier3	Drive Type/RPM	-/-
	Capacity(Used/Total)	- / - [- %]
	Performance Utilization	- %
	Buffer Space(New page assignment/Tier relocation)	- % / - %

Performance Graph (hot_band(0))

Object: Entire Pool Tiering Policy:

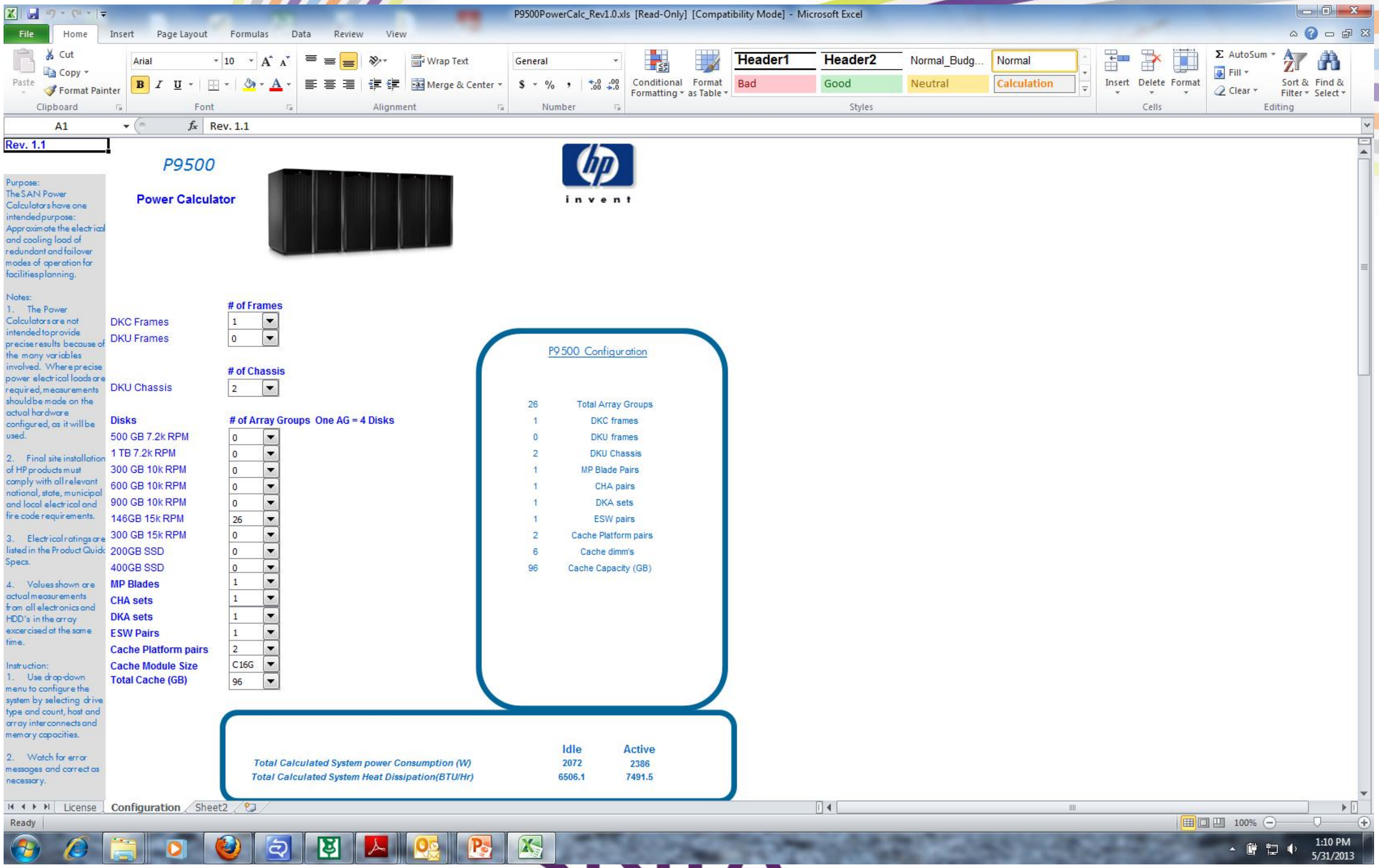


2013/04/26 09:54 - 2013/04/26 11:56

Close Help

Example

The HP P9500 Power Calculator



SNIA™
Green Storage Initiative

SNIA Green TWG Tiered Storage w/Hot Bands Primary Metric Comparison

Configuration	Tier Type	Power Consumption (Modeled)	IOPS	IOPS/Watt
Large Array (Initial)	15k RPM	7,491 Watts	18,[illegible]0	2.457
Large Array (Cache Assist)	15k RPM	7,491 Watts	39,[illegible]00	5.326
Large Array (Tiered)	15k + SSD	7,283 Watts	42870	5.886

SNIA™ Green Storage Initiative

SNIA Green TWG Hot Band Workload Conclusions and Observations

- The Hot Band workload is amenable to performance optimization by both storage subsystem cache and the proper deployment of tiered storage.
- As a result of high cache hit rates (~60%) the overall performance contribution of tiered storage is limited.
- In addition to the increase in IO rate (132%), there is also a corresponding decrease in power consumption from the substitution of SSDs in the configuration
- The net effect of these two parameter changes is a 140% improvement in the SNIA primary active metric (IOPS/Watt) of

SNIA™ Green Storage Initiative