The Storage Chasm:
Implications for the Future of HDD and Solid State Storage

An examination of the storage hierarchy from an economic point of view

Steven R. Hetzler

IBM Fellow, Manager Storage Architecture Research

hetzler@almaden.ibm.com

Outline

- ## Chasm Analysis
  - Details of the Storage Hierarchy

- ## Tales from the Supply Side
  - The capital cost of producing HDD and solid state storage

- ## Conclusions and Outlook
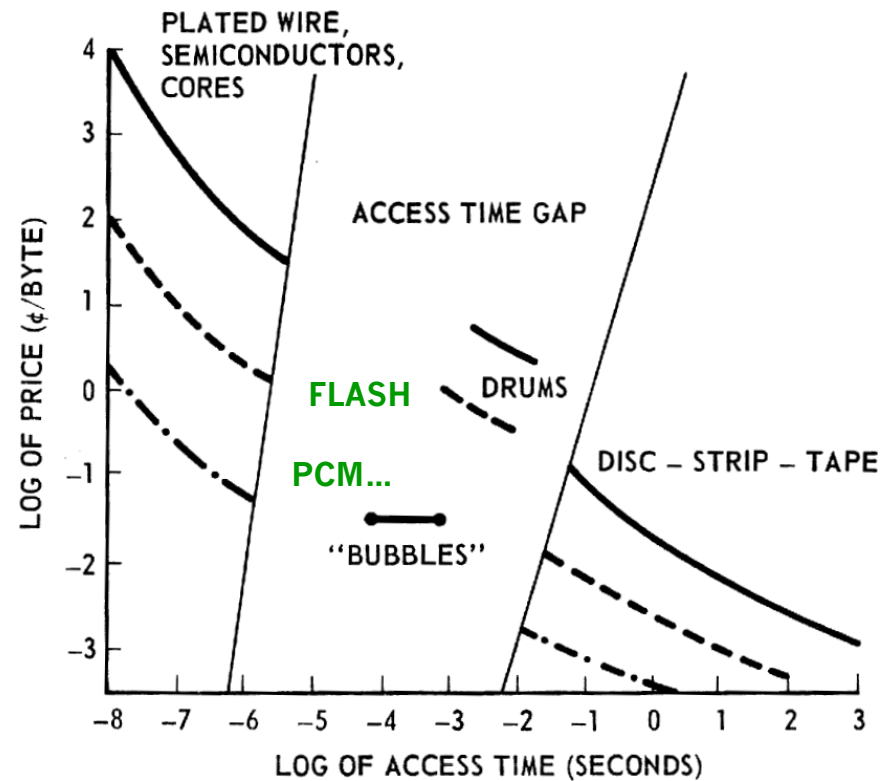 - What does it all mean

Acknowledgements:
  - Robert Fontana
  - Daniel Smith

Details of the Storage Hierarchy

CHASM ANALYSIS (CHASMOLOGY)

The Storage Gap

- **A large gap between main memory and storage has been noted**
  - In cost and performance
  - Looks exploitable by new technologies


(Bobeck, BSTJ 46, 1901[1967])

Take Away

Access time plot shows a large gap

Gapology

- **If the "gap" has been known for 40 years, why hasn't it been successfully exploited?**
  - Surely something would have filled it
  - Yet all candidates have failed in the "gap"
    - Drums, bubbles, MRAM, MO optical…

- **The issue is with the axes**
  - Market compares on $/GB and $/Performance
    - Gap chart is a technologists view
  - Let's examine this construct
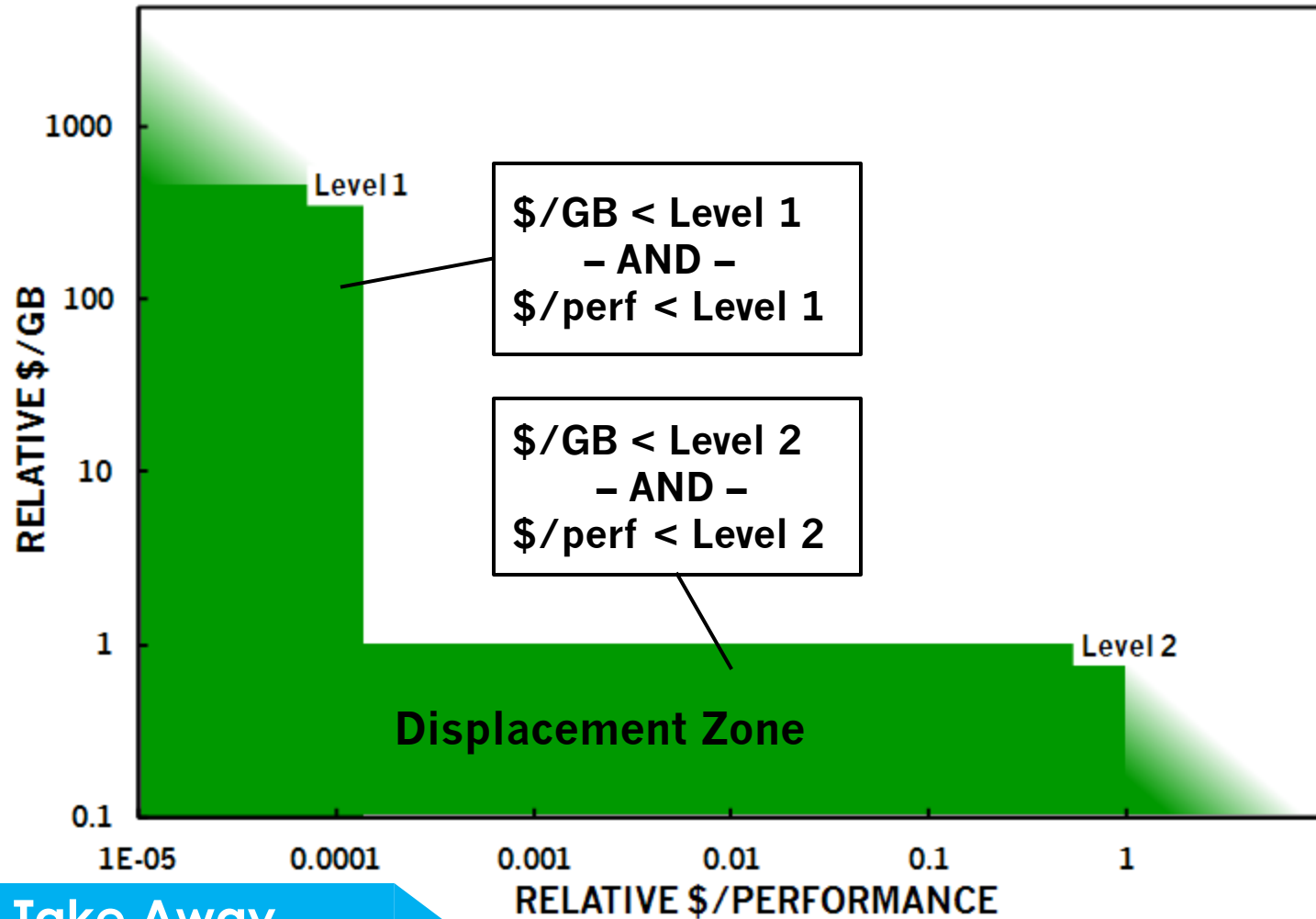
The Storage Hierarchy Axes

The Displacement Zone

$/GB < Level 1
– AND –
$/perf < Level 1

$/GB < Level 2
– AND –
$/perf < Level 2
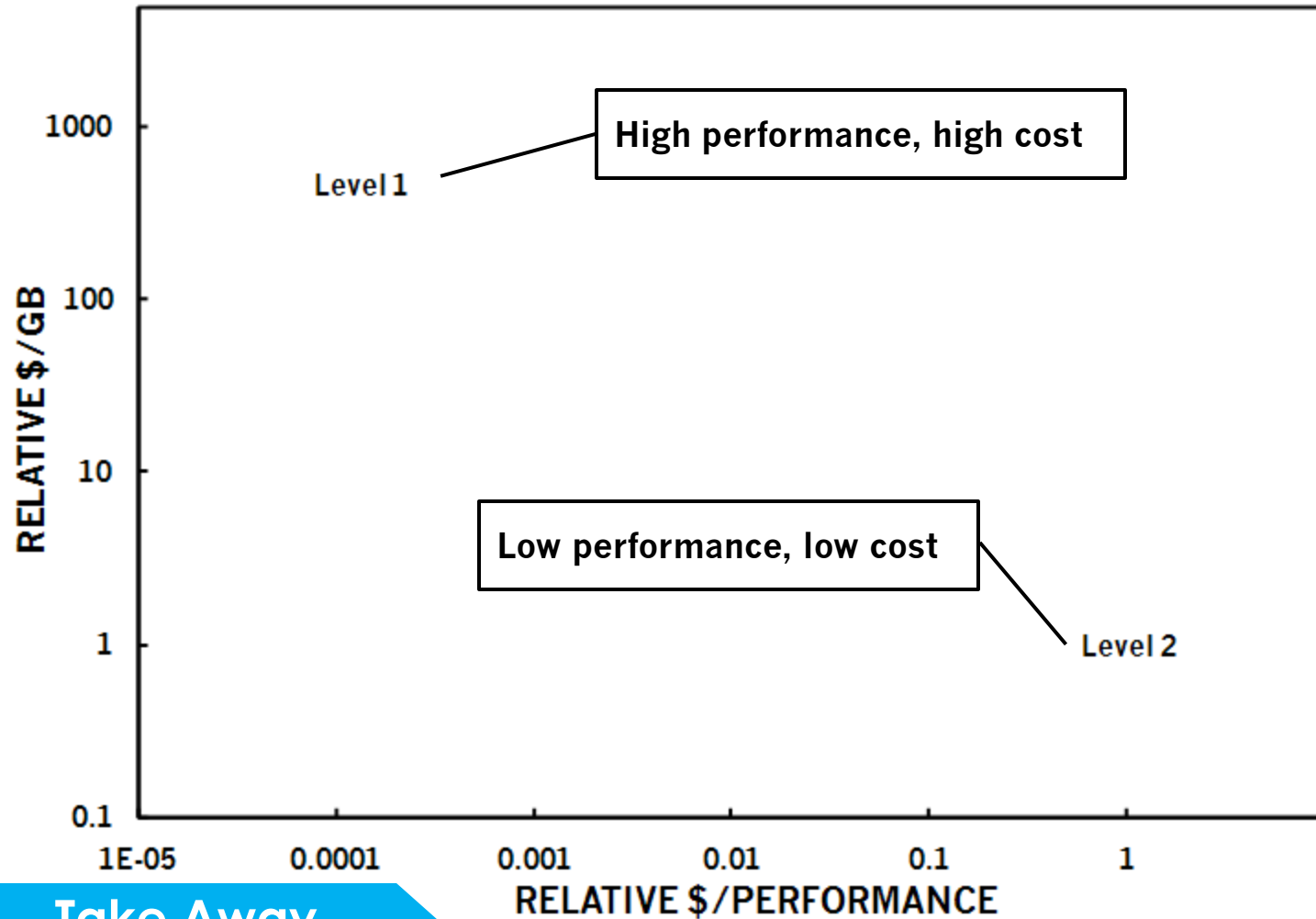
Displacement Zone

Take Away

Displacement possible when better in both dimensions

The Dead Zone

Dead Zone

$/GB > Level 1
– AND –
$/perf > Level 1

$/GB > Level 2
– AND –
$/perf > Level 2

Level 1

Level 2

RELATIVE $/GB

RELATIVE $/PERFORMANCE
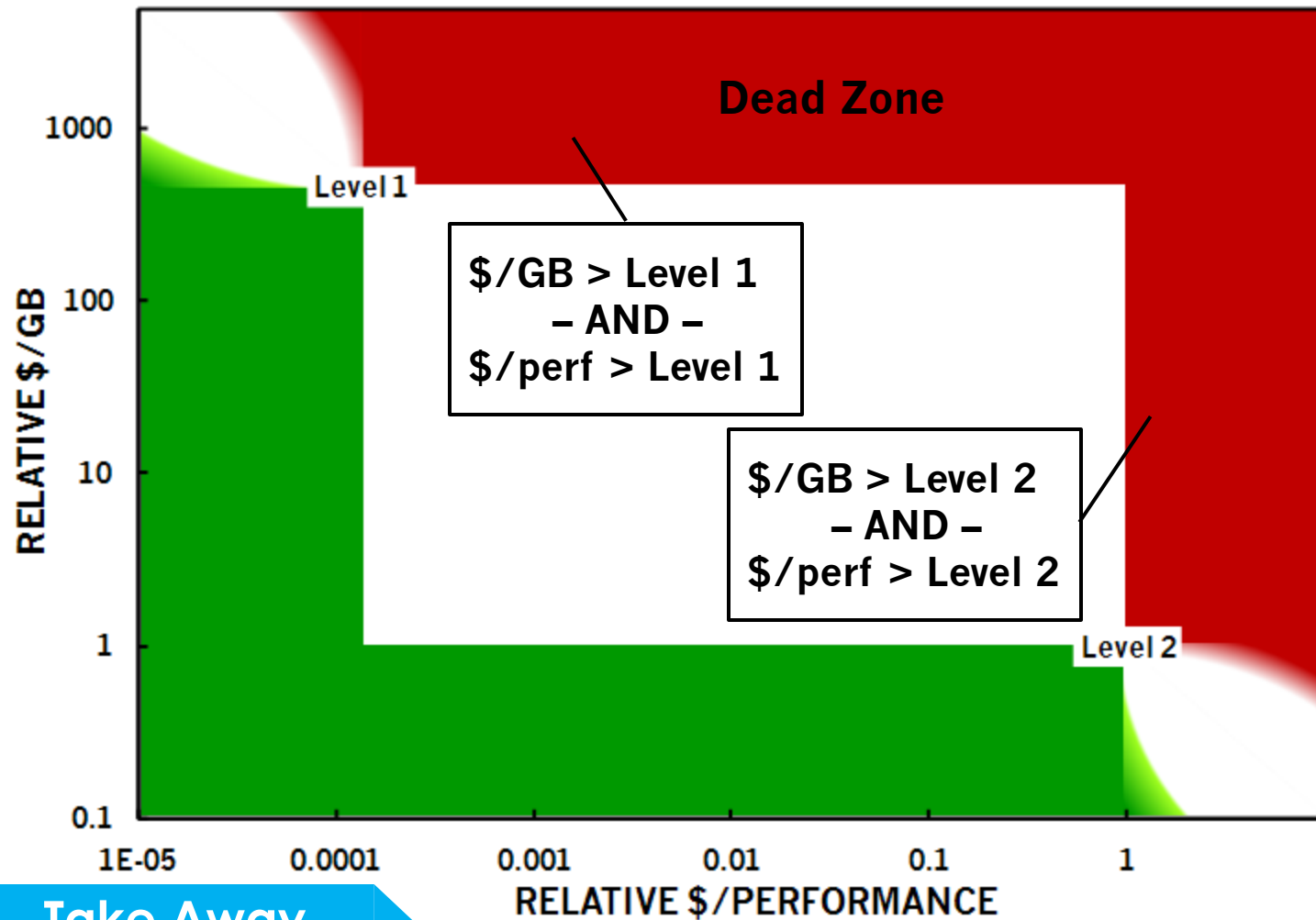
Take Away

No market when worse in both dimensions

The Gap

$/GB > Level 2
– AND –
$/perf > Level 1

$/GB < Level 1
– AND –
$/perf < Level 2
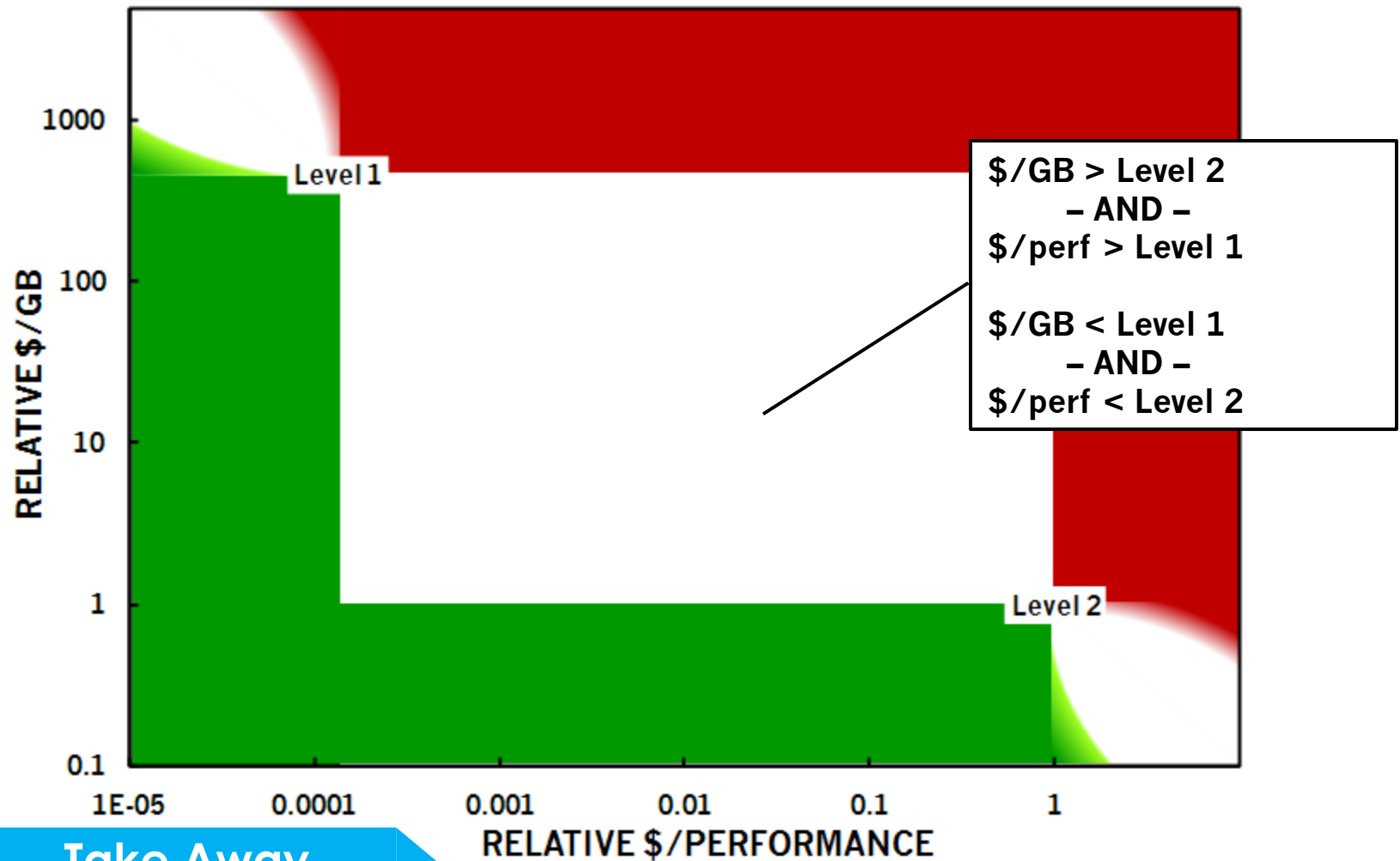
Take Away

Some say, the middle is an exploitable gap…

Increasing Potential

Market share negligible

Market share depends on perceived value

Level 1

Level 2

RELATIVE $/GB: 1000, 100, 10, 1, 0.1

RELATIVE $/PERFORMANCE: 1E-05, 0.0001, 0.001, 0.01, 0.1, 1
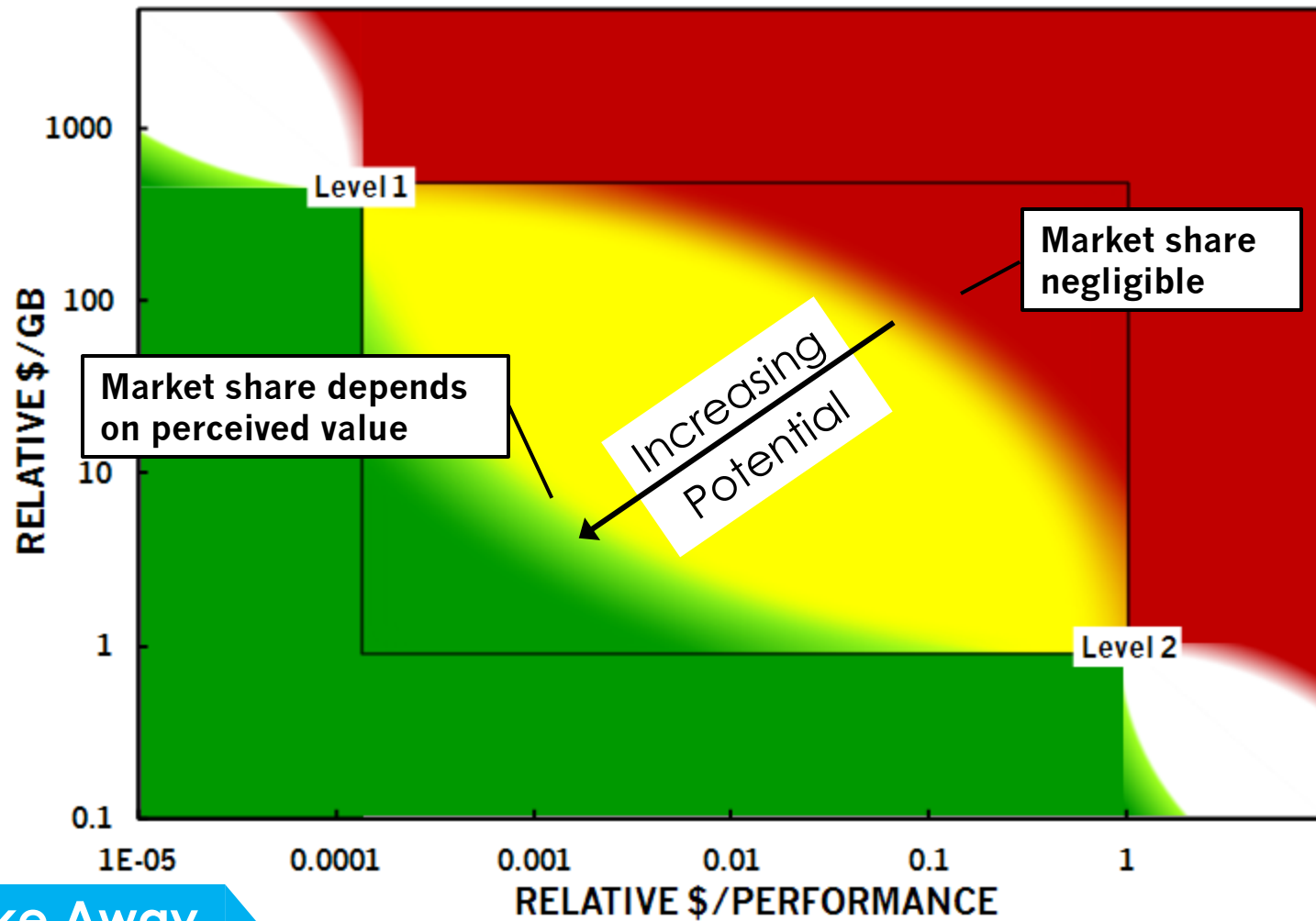
Take Away

The gap has an elasticity surface

The Cache Zone

Level 1 can be used to cache Level 2

Level 1

1000

100

Cache Zone

10

1

Level 2

0.1

RELATIVE $/GB

1E-05   0.0001   0.001   0.01   0.1   1
RELATIVE $/PERFORMANCE
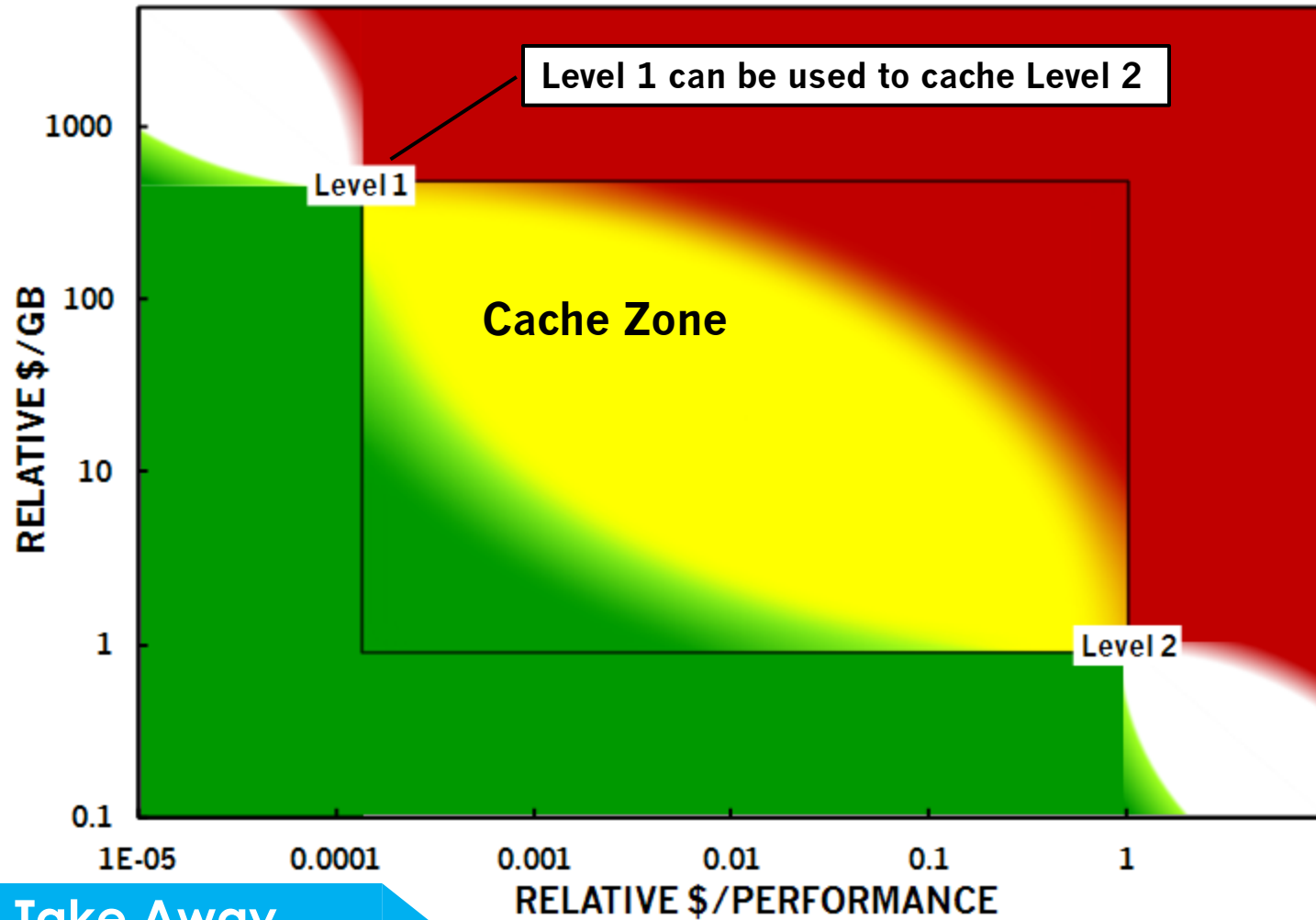
Take Away

Level 1 cache for level 2 makes displacement hard

The Cache Zone

Cost of system here goes up:

CZ $/GB > L2
$\therefore$ $ at fixed GB ↑

CZ $/perf > L1,
$\therefore$ $ at fixed perf ↑

Value of system must increase
enough to compensate
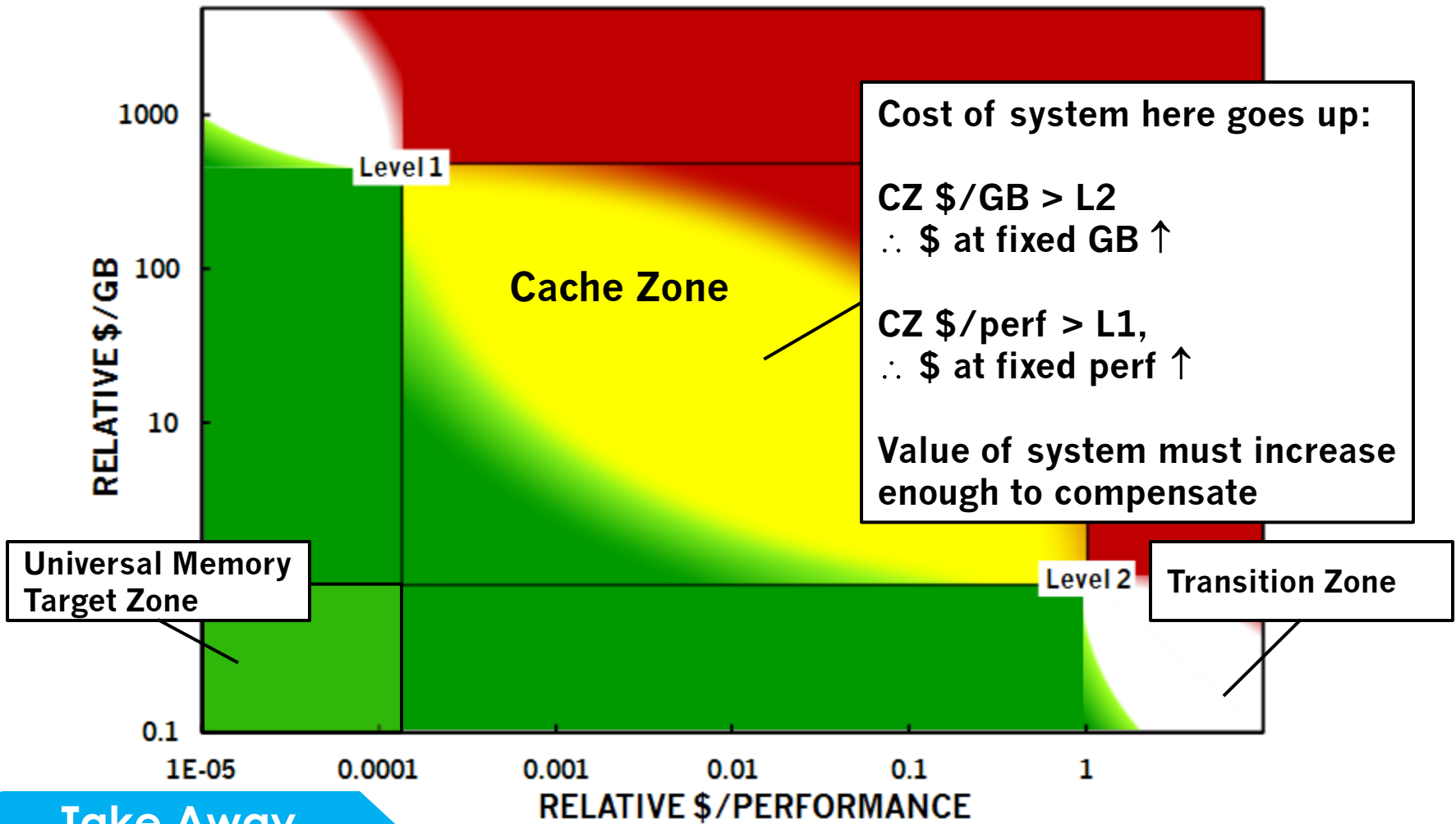
Take Away

Cache Zone storage increases system cost

The Storage Chasm

- **The "gap" isn't a gap at all**
  - It's really the Cache Zone
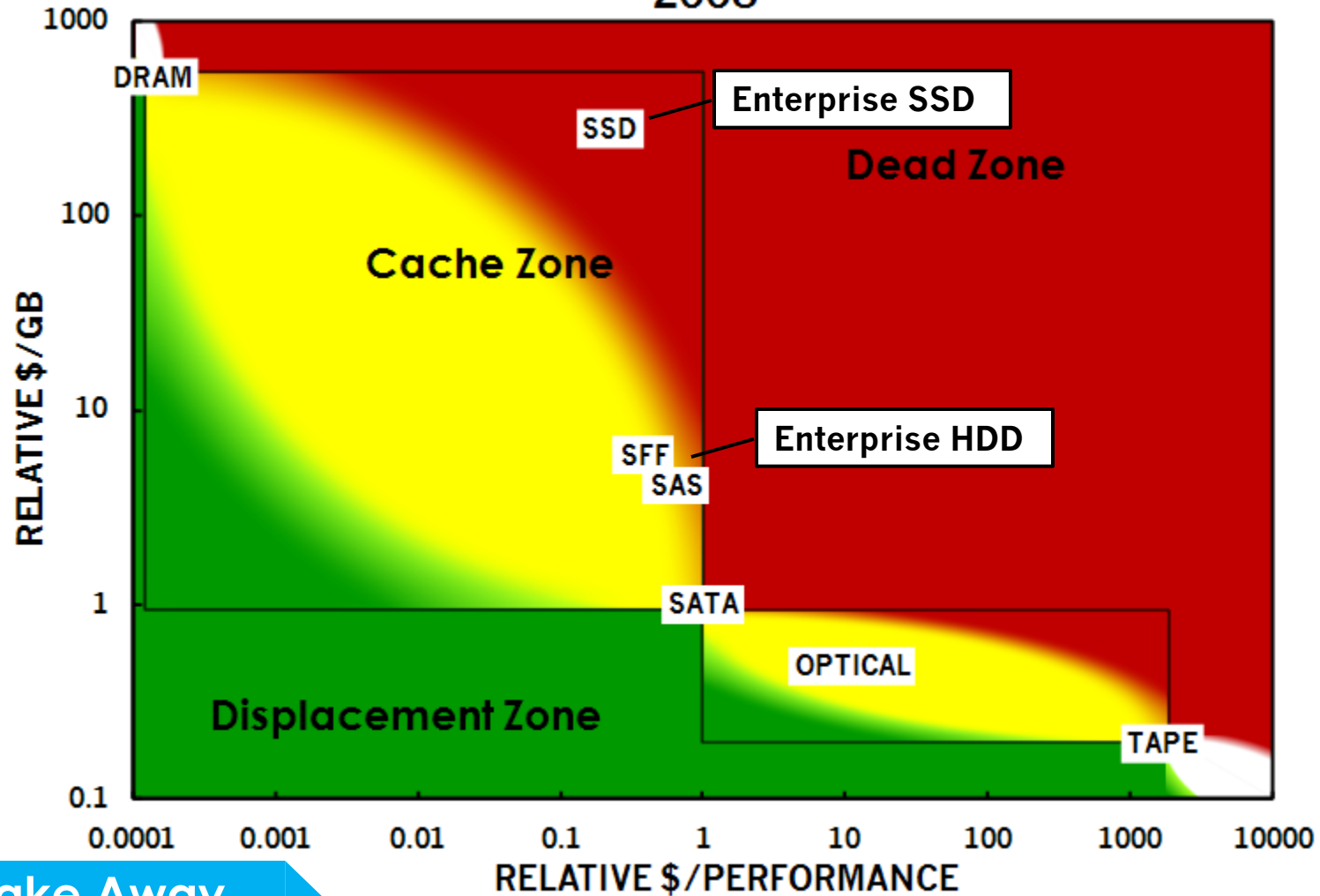 - Level 1 caching Level 2 limits opportunity in Cache Zone
  - Cache Zone storage increases systems costs
  - Closer levels are, harder life is in Cache Zone

- **This is the Storage Chasm**
  - The Chasm is a fundamental property of the hierarchy
  - Storage technologies do not thrive here

  - *That's why it has remained empty for 40 years*

Take Away

The Cache Zone is the Storage Chasm!

The IT Storage Hierarchy Today

2008

Chart showing RELATIVE $/GB (y-axis, 0.1 to 1000) vs RELATIVE $/PERFORMANCE (x-axis, 0.0001 to 10000) with zones: Dead Zone, Cache Zone, Displacement Zone. Storage types labeled: DRAM, SSD (Enterprise SSD), SFF, SAS (Enterprise HDD), SATA, OPTICAL, TAPE.

Take Away

Two Chasms: DRAM – SATA, SATA – Tape

Case Study #1: The Hybrid HDD

- ## Soundbite
 - Will increase system performance due to non-volatile cache

- ## Chasm Analysis
  - 2007 ~ $70 disk + $7 flash (256MB)
  - Cost now 110% of base
  - System already has a DRAM cache
    - On the faster side of the bus
      (oh - and it's NV due to laptop battery!)
    - Realizable performance gains likely to be small as result
  - Hybrid HDD is therefore in cache zone (higher margins aside..)
    - Increases cost of the HDD (and $/GB)
    - System performance increased more by spending $ on DRAM or CPU,…

- ## Field results
  - Hybrid HDD missed the market (value proposition was negligible)
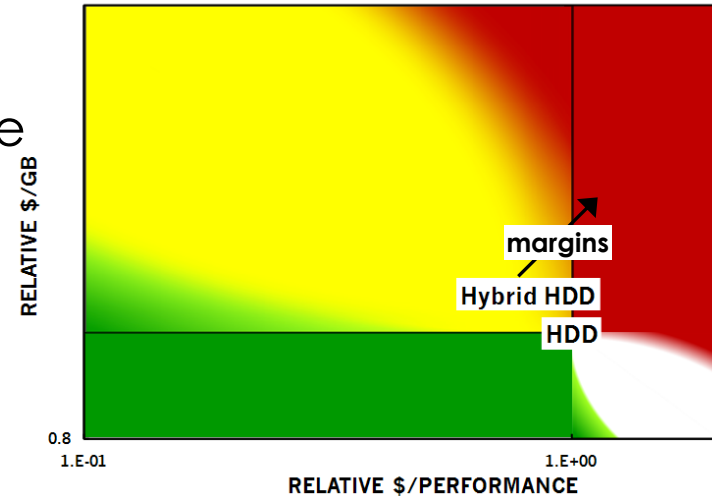
Take Away

Customer value highly sensitive to cost in this space.

Case Study #2: Holographic Storage

- ## Soundbite
  - High density by recording in 3 dimensions

- ## Chasm Analysis
  - Starts in the SATA Dead Zone
    - In-Phase - $180 for 600GB 2009*,
      250ms access time, 3.5s media change**
    - $0.3/GB vs. SATA $0.12/GB
 - At best, might move to SATA-Tape Cache Zone in 2012
    - Requires aggressive density timeline to be met
    - (InPhase just delayed shipment until late 2009)*

- ## Expected Field Results
  - Negligible market share
  - Doesn't have a value proposition for the total market
    - Squeezed by SATA, BlueRay and Tape

* "InPhase delays holographic storage to late 2009", The Register, 11/3/2008. 2x capacity credit for compression.
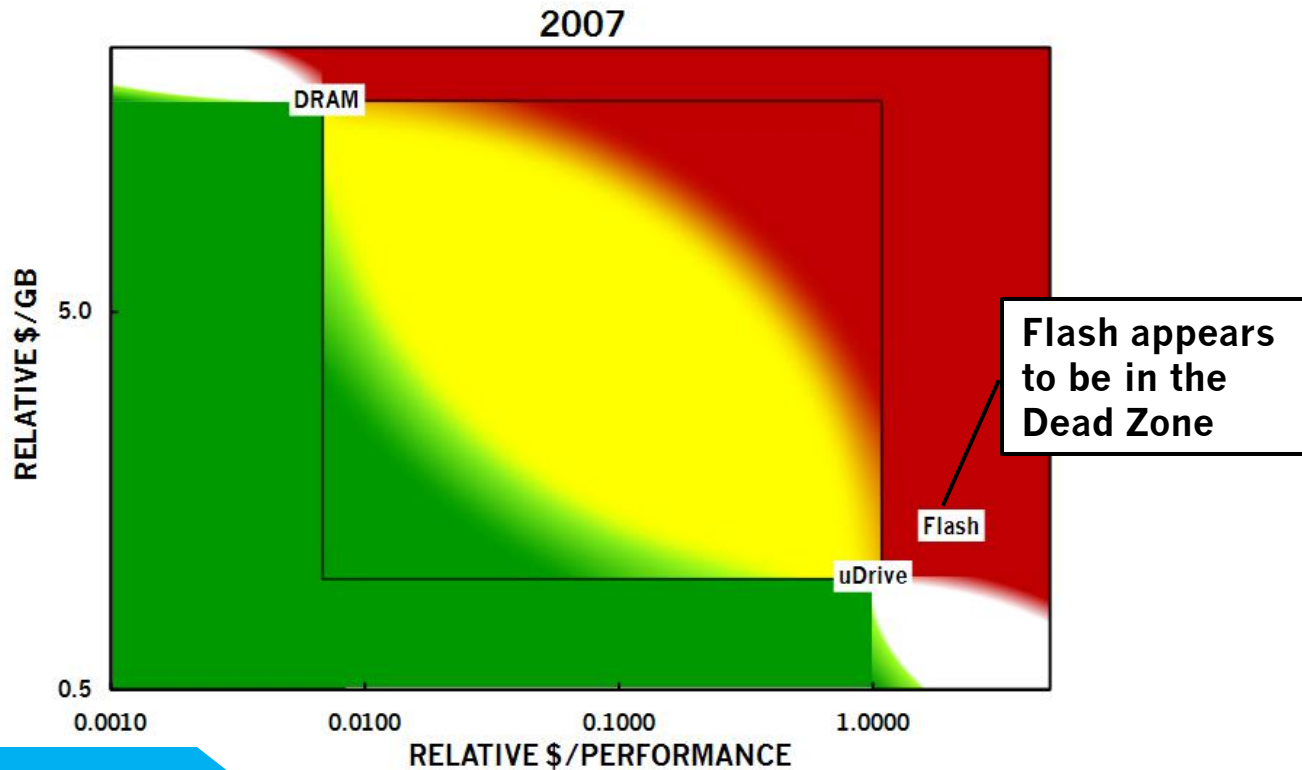** Tapestry Product Brochure, InPhase Corp.
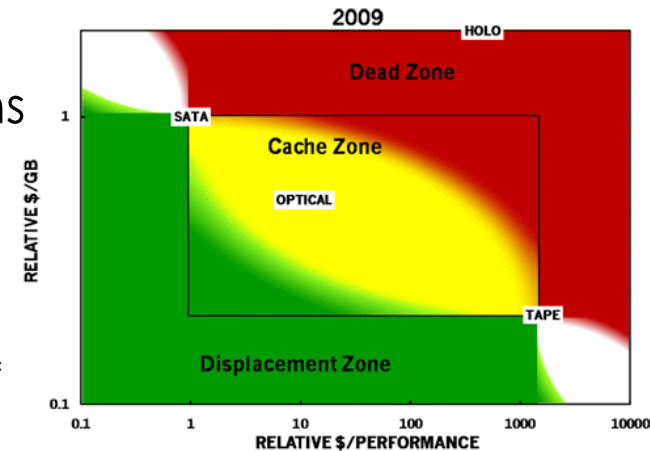
Take Away

Holographic will be victim of the Chasm.

Case Study #3: Microdrives

- ## **Soundbite**
  - Better $/GB than flash, better streaming – consumer market
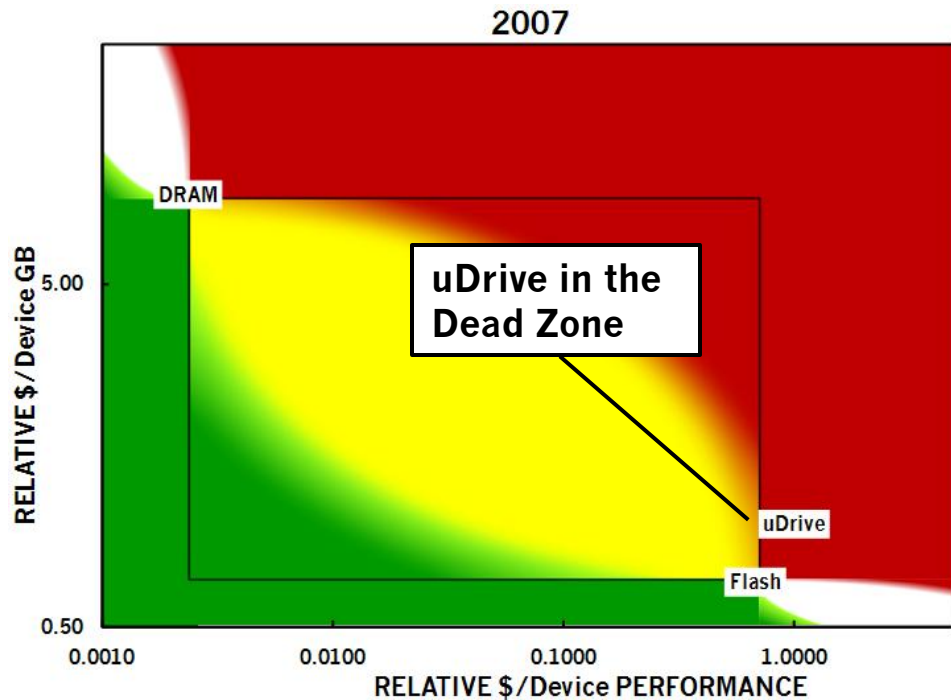   - Sounds like displacement zone in the consumer space

2007

Flash appears to be in the Dead Zone

Take Away

This is not what happened, because what matters…

Case Study #3: The System Design View

- ## Chasm Analysis
  - Need to look at system designs in this space
    - MP3 average ~4GB storage (2007)
    - Minimal streaming performance required (USB)

2007

uDrive in the Dead Zone

DRAM

uDrive

Flash

RELATIVE $/Device GB

RELATIVE $/Device PERFORMANCE

Take Away

What the system needs: 4GB & good enough perf.

Case Study #3: Microdrives Result

- **Chasm Analysis**
  - Microdrive was really in the Flash-DRAM Dead Zone
    - Perhaps not at first, but time dependency moved it there
      - Driven by capacity overshoot

- **Field Results**
  - Essentially no more microdrives
  - Displaced by flash
    - Not because "non-mechanical" is better
    - But because flash is good enough,
      hence **cheaper** for solution

Take Away

Low entry cost made flash-based solutions cheaper

Flash Value Proposition

- **Hypothesis for microdrive replacement**
  - Customer will pay higher $/GB for flash value add:
    - Solid state (never fails…), shock, power, performance

- **However, if that were true, then why:**

Flash		**HDD**	
	8 GB iPod Nano $150		120 GB iPod Classic $250
	4 GB Sony HDR-CX12 $900		120 GB Sony HDR-SR12 $1200

  - Consumer pays *less* for 4GB of flash than for 4GB of HDD
  - Flash $/GB premium at solution is not justified
    - Otherwise we would have 120GB flash devices in consumer

Take Away

Measure is cost of the solution – no $/GB premium

Case Study #4: Solid State Disk

- ## Soundbite
  - Flash is better than HDD because it is solid state
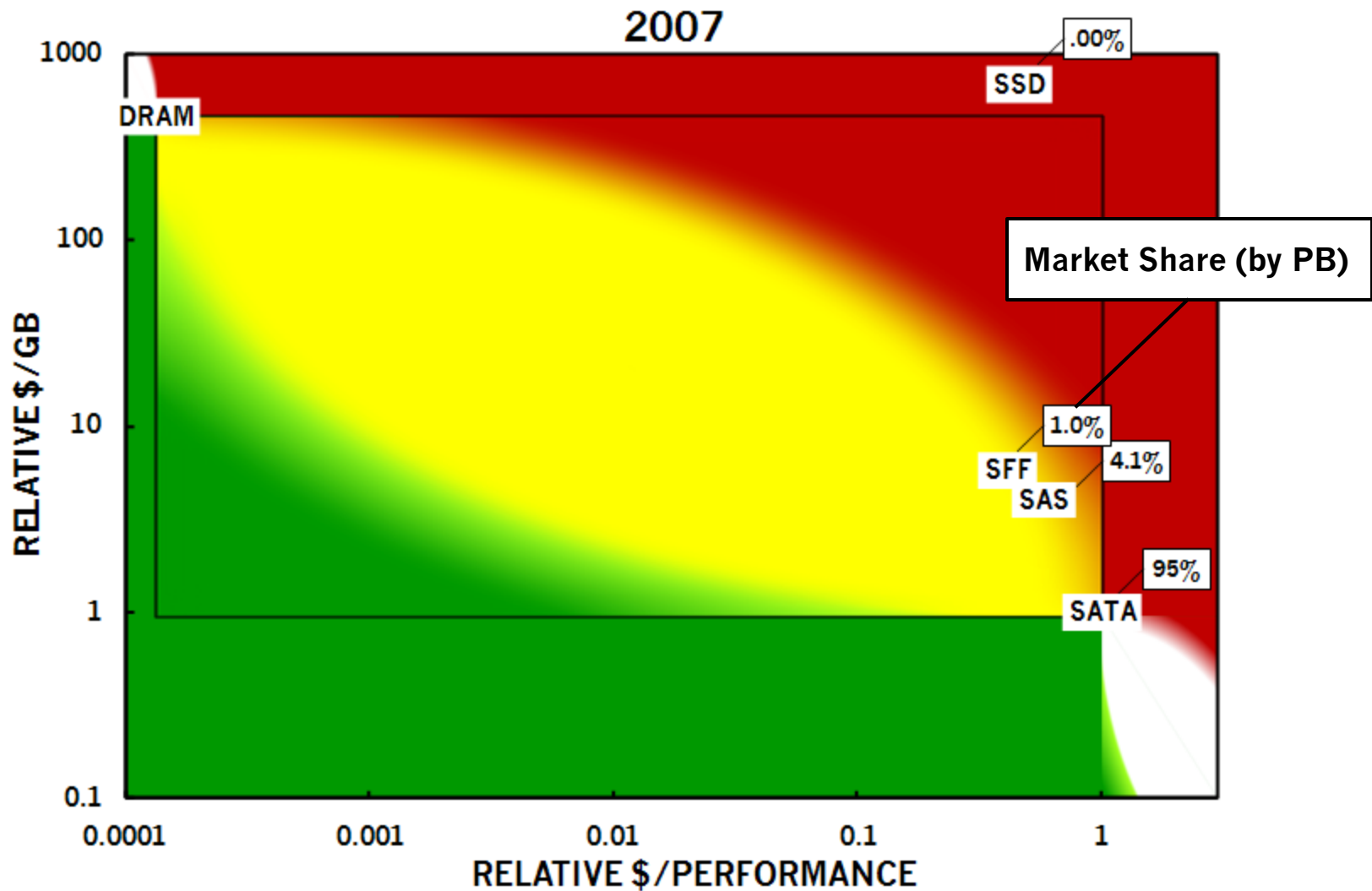
- ## The following use IDC HDD and SSD data*

- ## Chasm Analysis
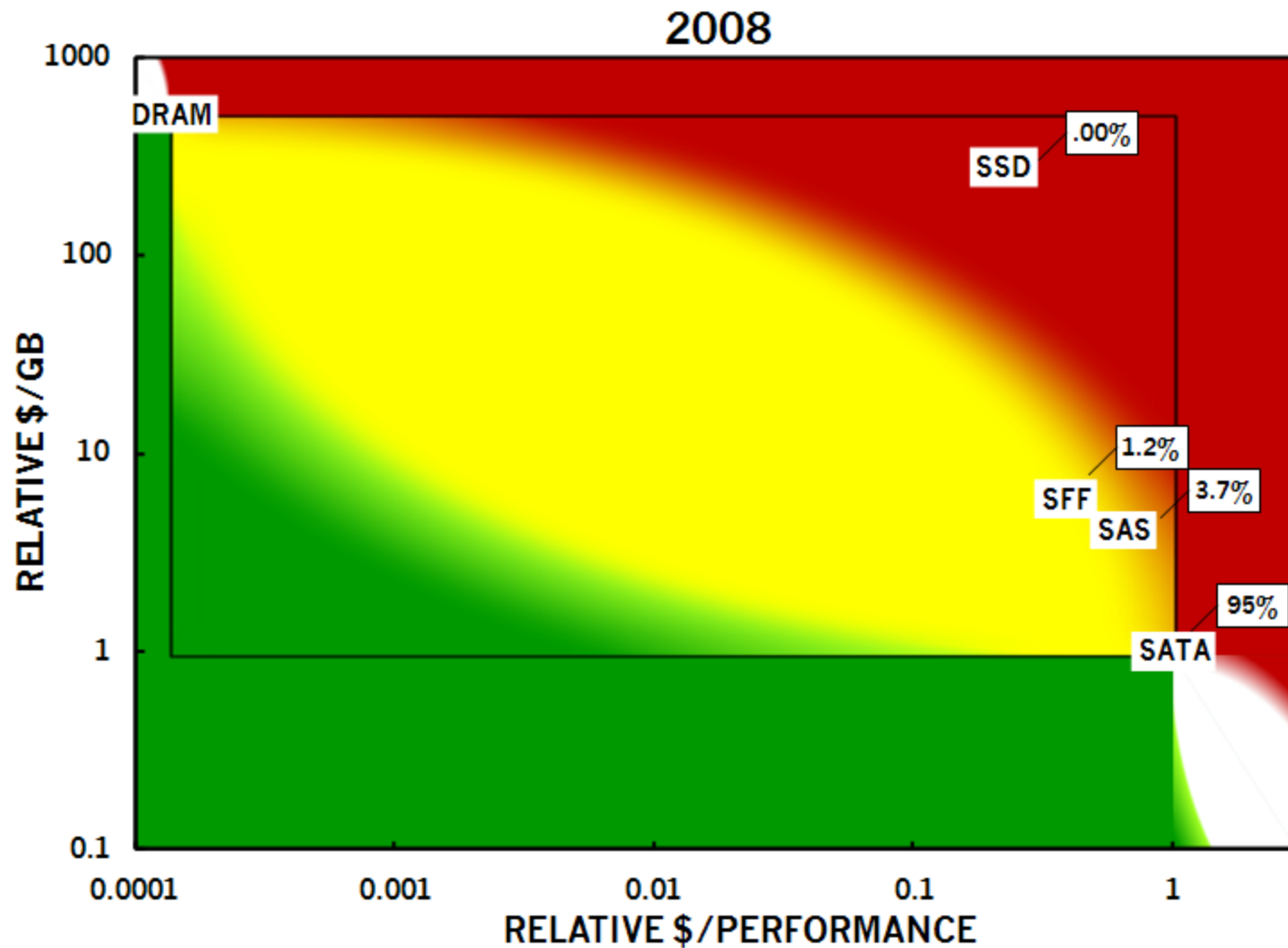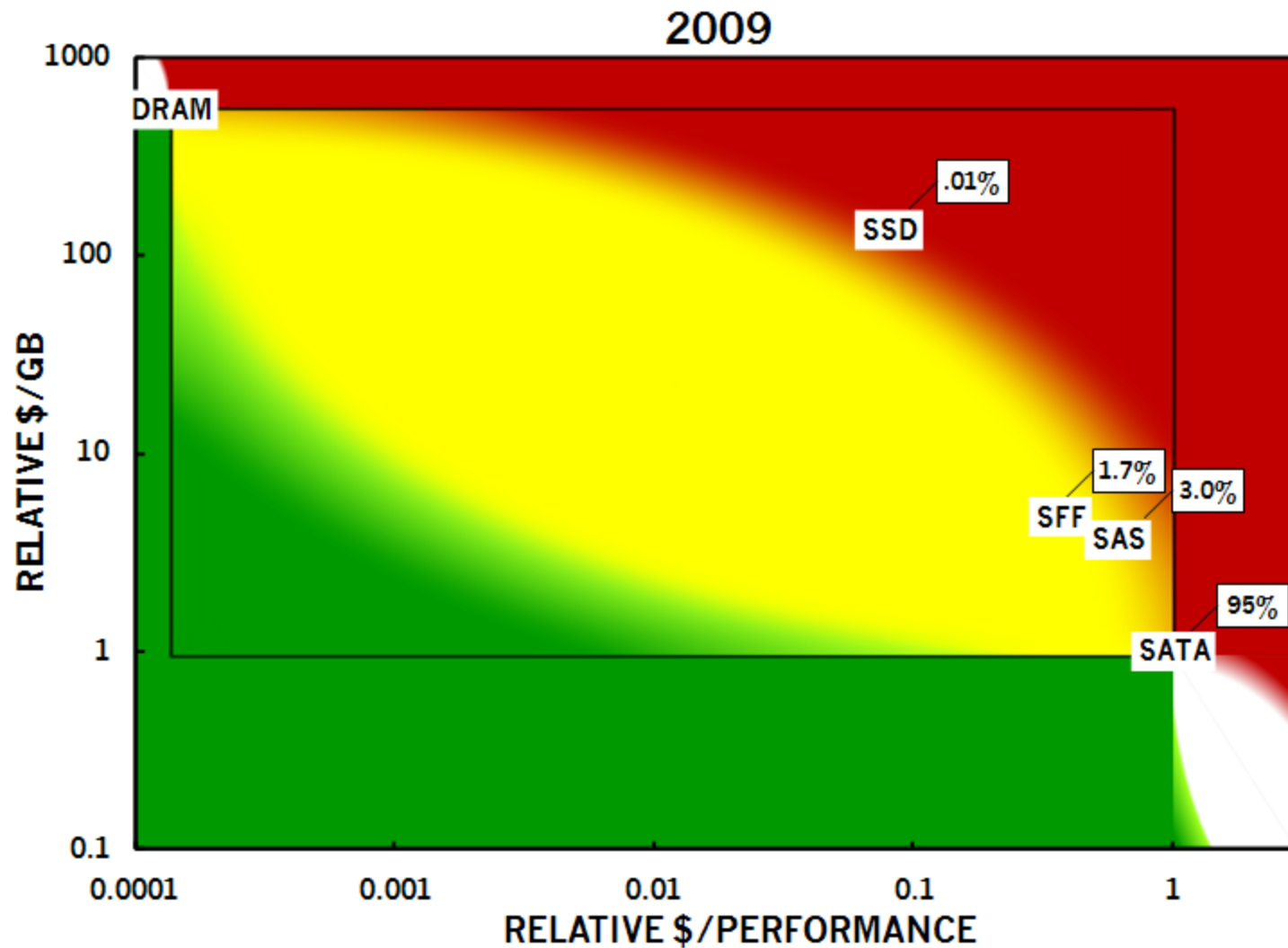 - We can use enterprise disk to measure elasticity in the Chasm
  - Market share by PB

*Sources:
IDC WW HDD 2008-2012 forecast
IDC WW SSD 2008-2012 forecast

2007

Market Share (by PB)
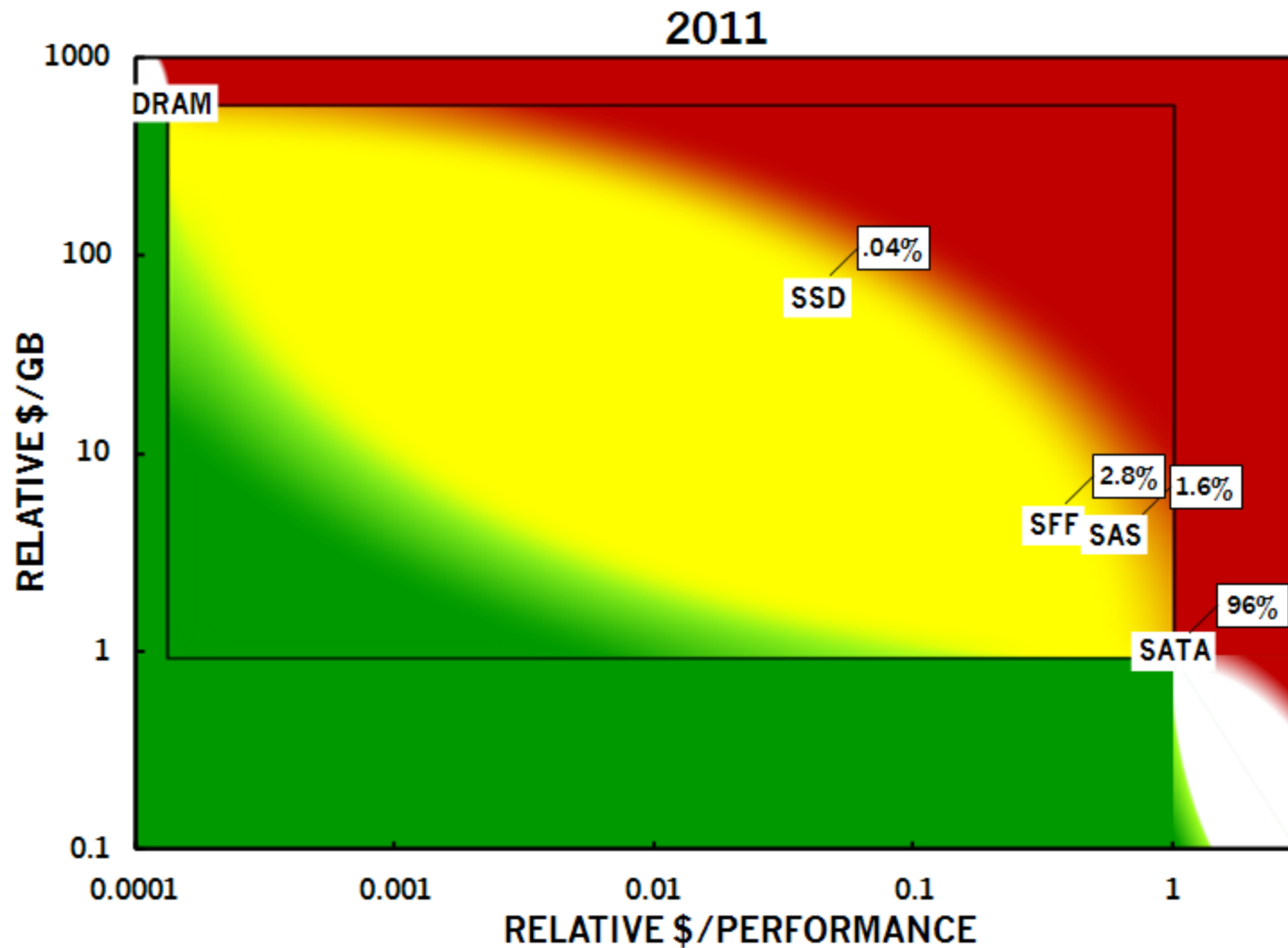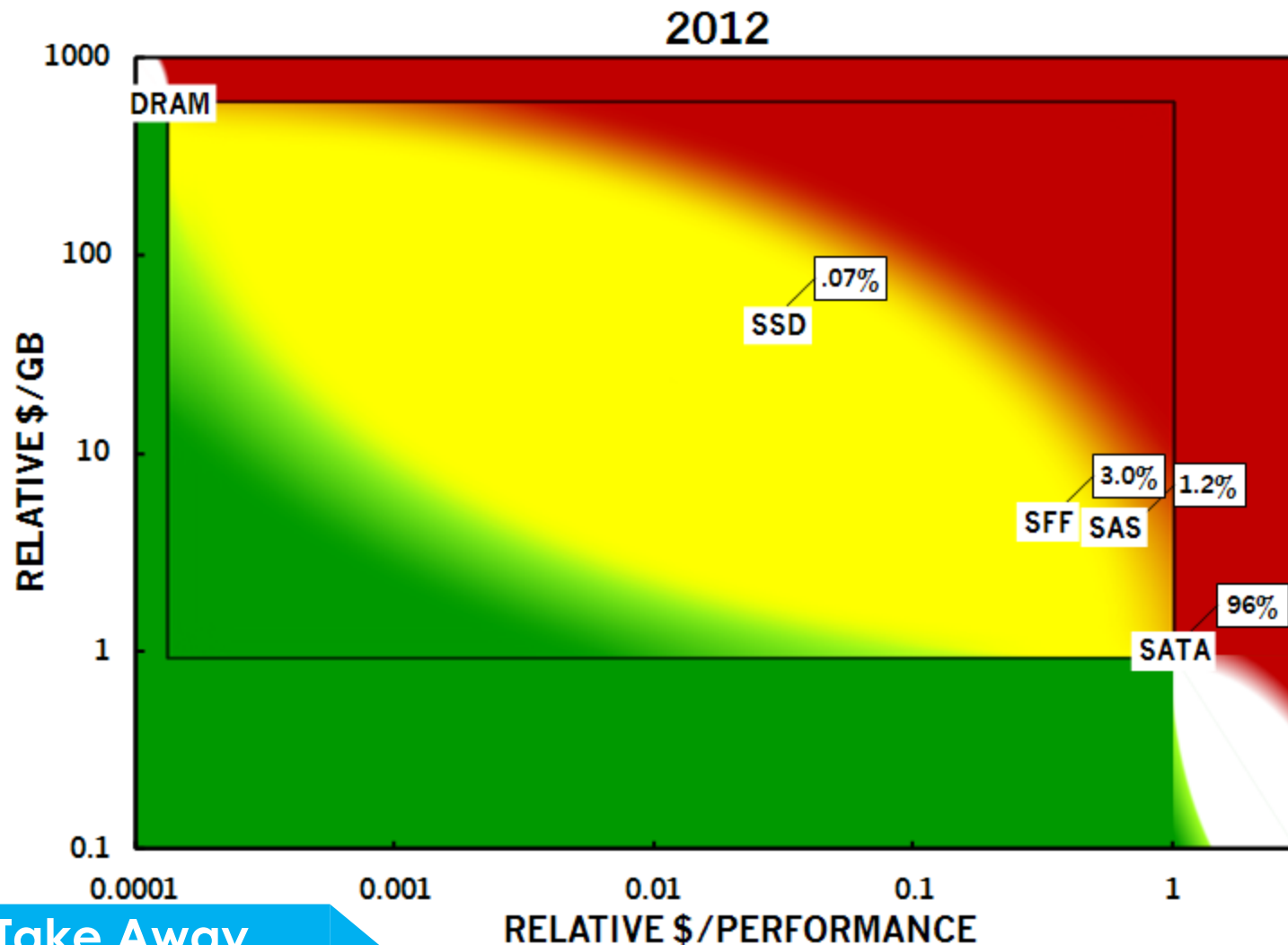
2008

RELATIVE $/GB

DRAM .00% SSD 1.2% SFF 3.7% SAS 95% SATA

RELATIVE $/PERFORMANCE

2009

Steven R. Hetzler, © Copyright IBM Corp. 2008,2009

2010

2011

2012
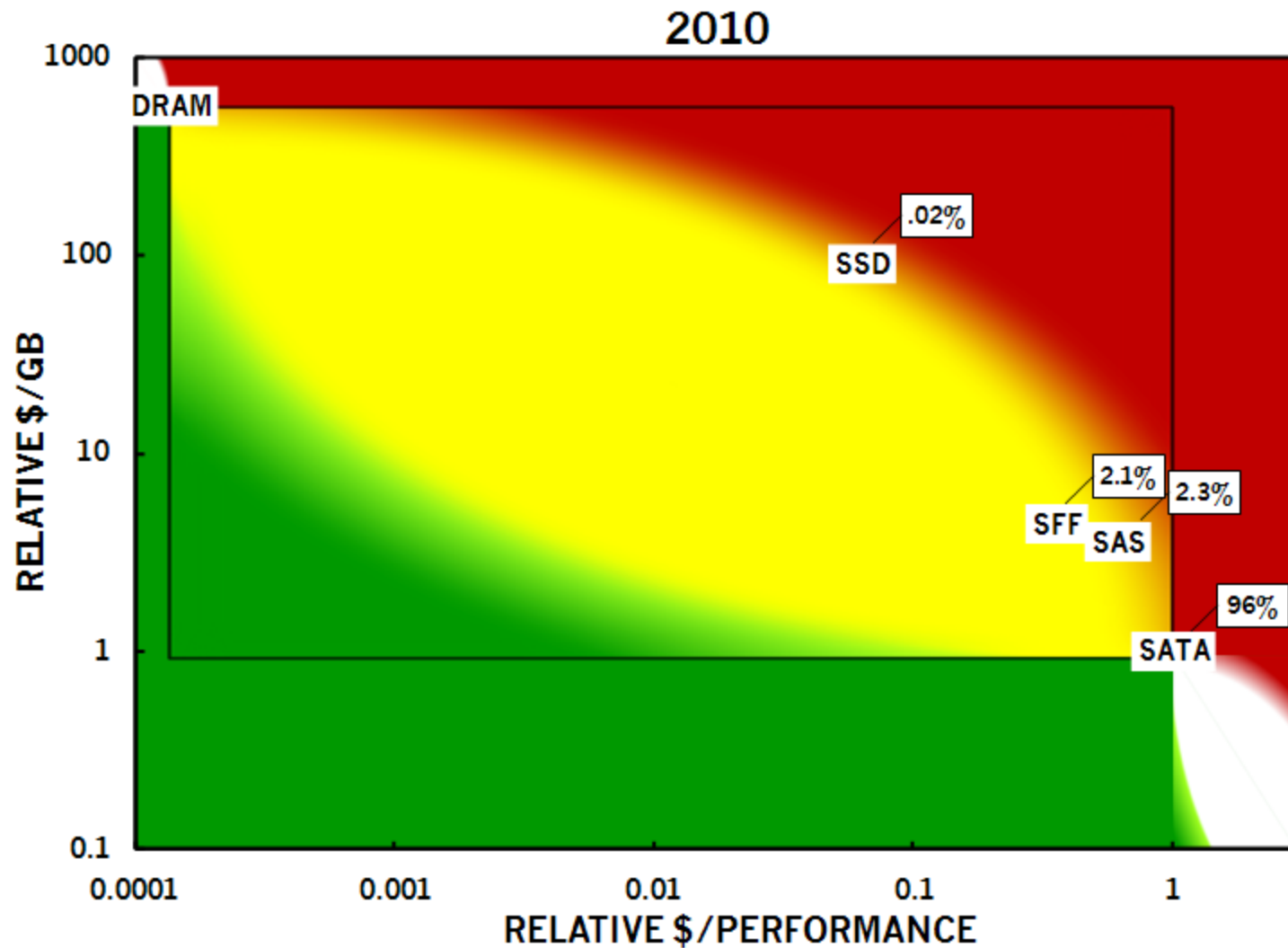
Take Away

It's hard to earn a living in the Storage Chasm

Observations

- **Storage Chasm is real**
  - A few orders of magnitude seems to be the historically preferred separation in the storage hierarchy

- **SSD is in the Chasm for many years**
  - Still negligible market share in 2012
    - Nowhere near displacement
  - Market value for SSD performance is limited
    - Enterprise HDD shows elasticity for performance at higher $/GB – < 5% share

- **Solid State strength**
  - Smaller unit granularity
    - Lower cost of first byte
  - Winning applications will play to this strength

Take Away

Solid state storage wins when the solution is cheaper

The economics of producing solid state storage

TALES FROM THE SUPPLY SIDE

Why Economics Matters

Why Economics Matters

NAND Flash and HDD Fabrication YE2008

MLC NAND Flash		HDD	
Wafer diam.	300mm (12'')	Head Wafer	6''
Node	45nm	Head Node	60nm
Die capacity*	2GB	Disk diam.	3.5''
Dies/wafer	425	Heads/wafer	30,000
Wafer Capacity	850 GB	Disk Capacity**	375 GB
Daily output	1,250 wafers	Disks/day	100,000
PB/line/day	1.1PB/day	PB/line/day	38PB/day
PB/line/year	**390PB/year**	PB/year	**14,000PB/year**
Fab cost	**$3.4B*****	Fab cost (est)	**$1B**

Only ~
½ SATA HDD!

* Assumes highest density 2-bit MLC flash. Wafer GB for SLC flash will be lower.
** Assumes highest density SATA disk .
***Source: IBS

Storage Market Compare 2008

- ## WW storage
  - 125kPB HDD shipped*
    - Growing at 40%/year
  - 8.9B sq" Si WW capacity** - that's 2.2 square miles
    - 82M equivalent 300mm wafers
      - Only 2% is at 45nm node
  - If all world's Si capacity moved to MLC flash today:
    16kPB = 12% of HDD PB!
      - That's assuming no tooling upgrades (constant node)

- ## WW revenue
  - $35B HDD*
  - $280B Silicon***
  - Will world pay more for storage just because it's "faster"?
    - Chasm Analysis shows this is not probable
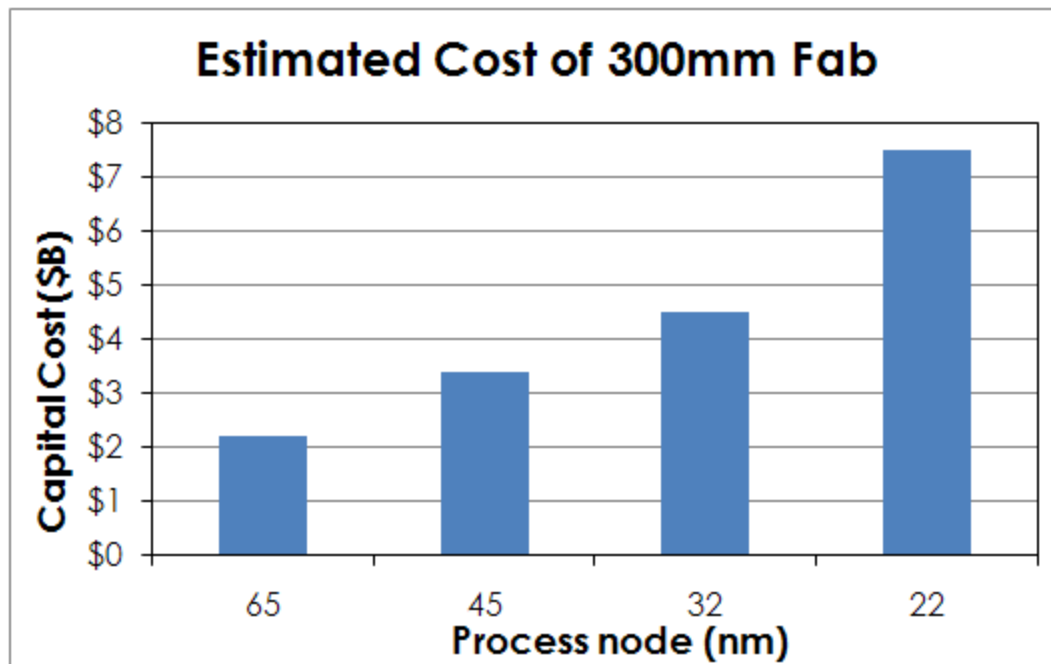
* IDC WW HDD 2008-2012 forecast
**Gartner, Jan 2008. Fab capacity about 88% utilized, not including empty fabs.
***Gartner, Nov 2008.

300mm Fab Cost

- ## **Capital costs increase as node size shrinks**
  - Today's 45nm fab $3.4B*
  - 22nm fab increases to $7.5B
    - Expected to start around 2012**

Estimated Cost of 300mm Fab

* IBS
** Gartner, Jan 2008

HDD Displacement

- ## It will take significant investment to dent HDD
  - 2008 est. $40B semico capital spending for $280B revenue*

* Gartner, Nov 2008

Fabs and Costs of HDD Displacement

2008	HDD Rev ($B)	Equiv. Flash Capital ($B)	% Semico Rev
1.8" mobile	$1	$17	0.4%
Enterprise	$5	$53	1.8%
2.5" mobile	$12	$287	4.3%
All HDD	**$35**	**$1,080**	**12.5%**
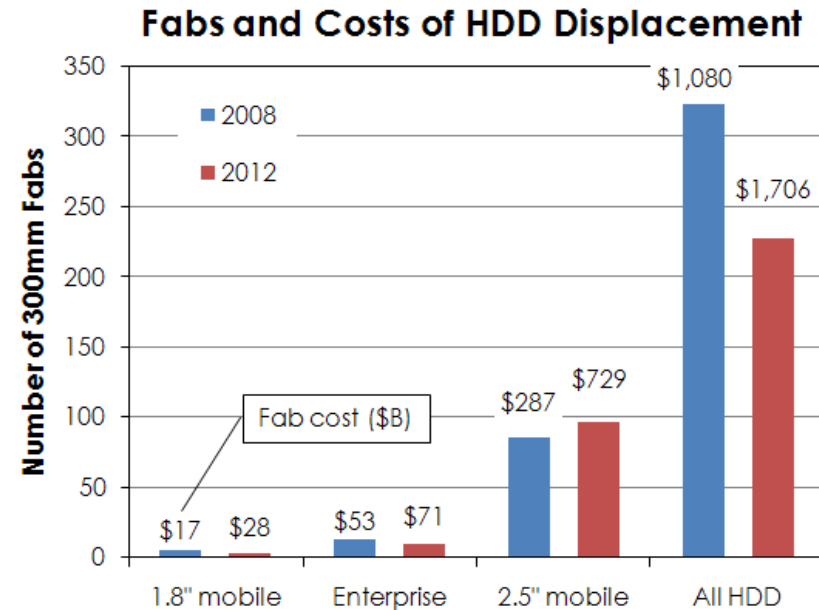
Take Away

Be careful what you wish for…

It's A Small NAND World After All…

- ## Some numbers from the flash market

2008	Units (M)	PB	% HDD PB
HDD	580	125,400	100%
iPods*	55	357	0.28%
iPhones*	14	100	0.08%
USB Key**	179	598	0.48%
Flash Card**	752	1,548	1.23%
SSD***	3	48	0.04%
Other	146	156	0.12%
Total	1,150	2,782	2.2 %

– Capital cost indicates that total PB must be small

* Apple
** Gartner, Oct 2008
***IDC, Aug 2008

A 1% NAND World

- ## For NAND to get 1% of HDD market (by PB)
  - 2008: 125kPB HDD total (2012: 400kPB)
    - 1% = 1.25kPB (>40% increase in NAND supply)
    - Using 2GB MLC dies (45nm) – 3.2 fabs

  - $11B @ $3.4B/fab (for SLC, 2-4x more)
    - That's just the fab cost, doesn't include operating costs

  - NAND economics will change for 1% HDD market share

- ## Capital cost to replace HDD <u>increasing</u> 19%/Y
  - HDD kPB 40%, NAND density 40%, fab costs 19%

Take Away

At these volumes, NAND no longer in oversupply

Observations

- **Flash is not likely to displace HDD in IT storage space**
  - In terms of capital, cost and reliability
  - Will be the case for most Si storage technologies
  - Better financial return in other markets
    - Unless you believe IT is underpaying for storage…

- **Why will it be so hard for Si storage?**
  - **Flash patterns one unit (cell), gets 2 bits**
  - **HDD patterns one unit (head), gets 10^{12} bits**

T-shirt of the Day:

I learned why they call it "Mechanical Advantage"

And prognostications

CONCLUSION AND OUTLOOK

Summary 1

- **There is a Chasm in the storage hierarchy**
  - Not a "performance gap"
  - Not filled in over 40 years – may not be possible
  - Fundamental – Chasm technologies *increase* system costs

- **Elasticity for $/performance is steep**
  - Very difficult to obtain significant penetration
  - Enterprise HDD at 5% market share shows just how steep
  - Flash dominates where it makes <u>solutions</u> cheaper

- **Chasm Analysis useful for any proposed storage technology**
  - Shows how difficult displacement will be
  - Prospects for universal memory appear slim

Summary 2

- **Capital cost of Si storage can't be ignored**
  - Si is substantially more expensive to produce than disk
    - It's not just areal density that matters
  - Not enough fab capacity on the planet to replace HDD
    - Not enough money to build it either (~$2T in 2012!)

- **All solid state storage should be tested against capital cost**
  - Significant HDD penetration at odds with capital cost
  - Possible to displace flash in consumer market
    - Perform DRAM-flash-"new memory" Chasm Analysis
      - (see above microdrive case study)

Signposts for Avoiding the Chasm

- ## HDD Capacity Overshoot
  - If HDD capacity growth overshoots, greater SSD opportunity
    - Perhaps in mobile…
    - But as mobile replaces desktop, storage needs likely to grow
- ## HDD $/GB slows
  - If it drops below 40% CGR
    - Over extended period opens up opportunities
      - If disk density growth stopped today…
        » SSD gains still limited by cost disadvantage
  - Move to patterned media could hurt HDD economics
    - Reduces the mechanical advantage
    - Would pattern 2.4B surfaces in 2012 (3.3 square miles!)
      - Might need to check the capital cost here…

Outlook

- **Enterprise HDD erosion continues**
  - Moves further into SATA Chasm as $/GB falls behind
  - PB share on enterprise disk will drop below 4% in 2012
    - Areal density growth faster on SATA

- **SSD unlikely to get significant market penetration**
  - Stays in the Cache Zone relative to SATA HDD
  - Best opportunities enterprise and 1.8" HDD
  - But remain too expensive to get near 1% share until after 2012
    - Capital costs prevent rapid adoption
    - Rapid adoption will change NAND market to undersupply