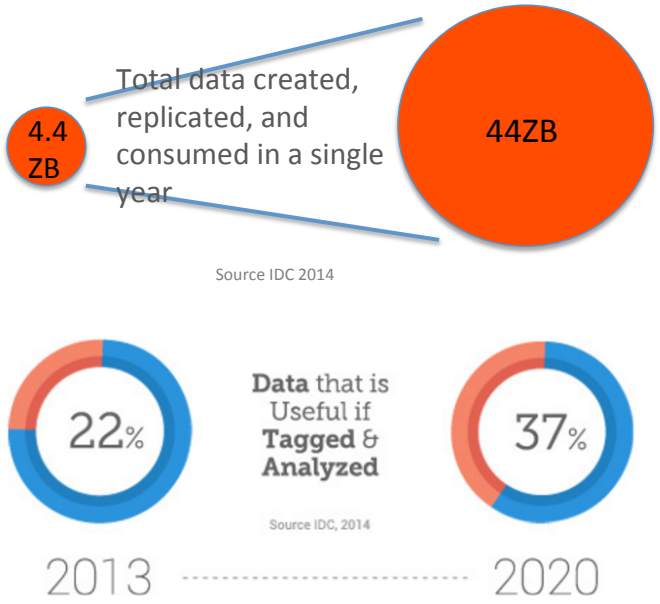
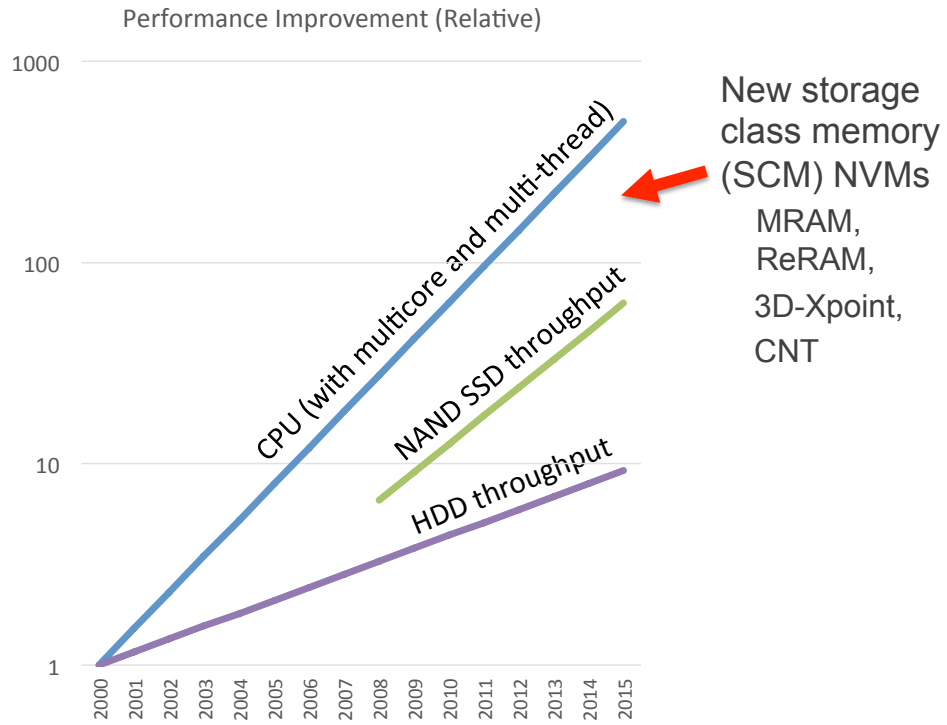
MANAGING MULTI-TIERED NON-VOLATILE MEMORY SYSTEMS FOR COST AND PERFORMANCE

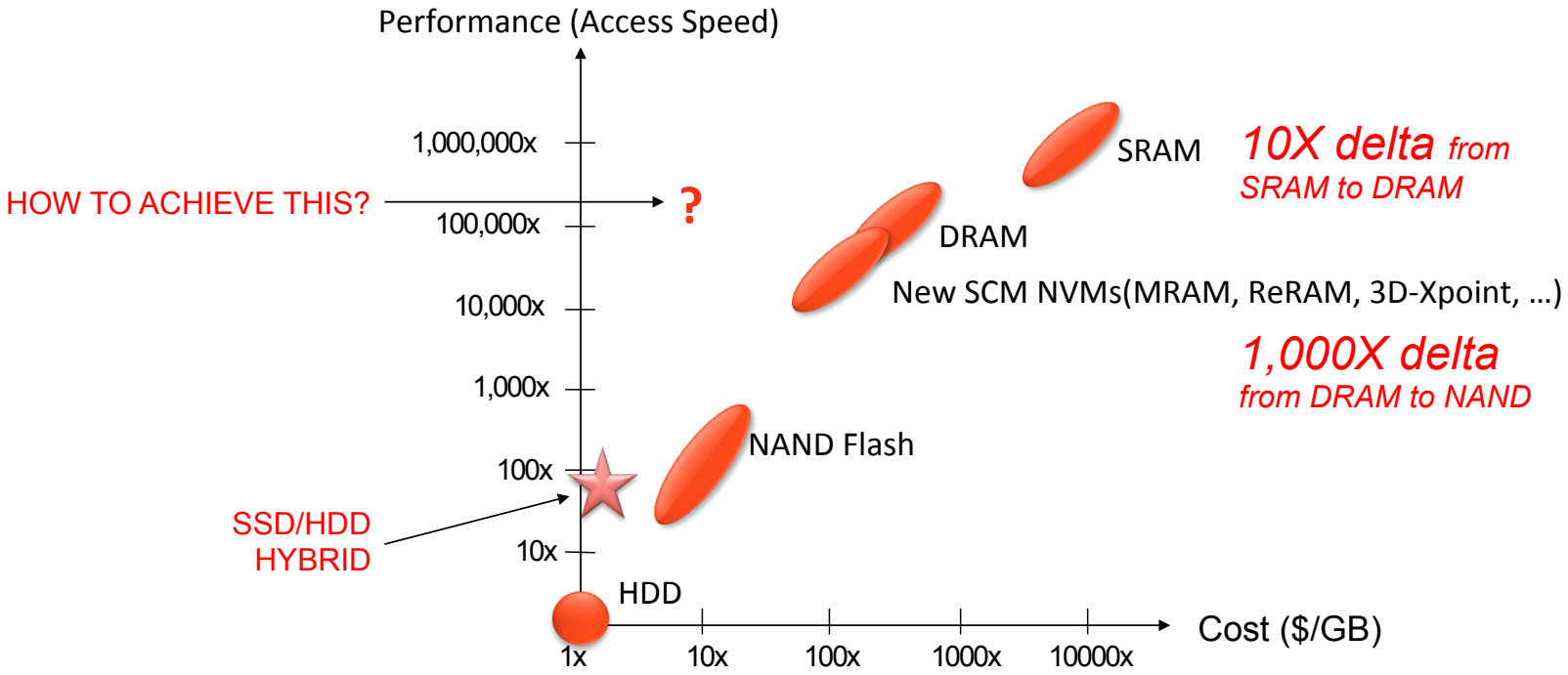
8/9/16

THE DATA CHALLENGE

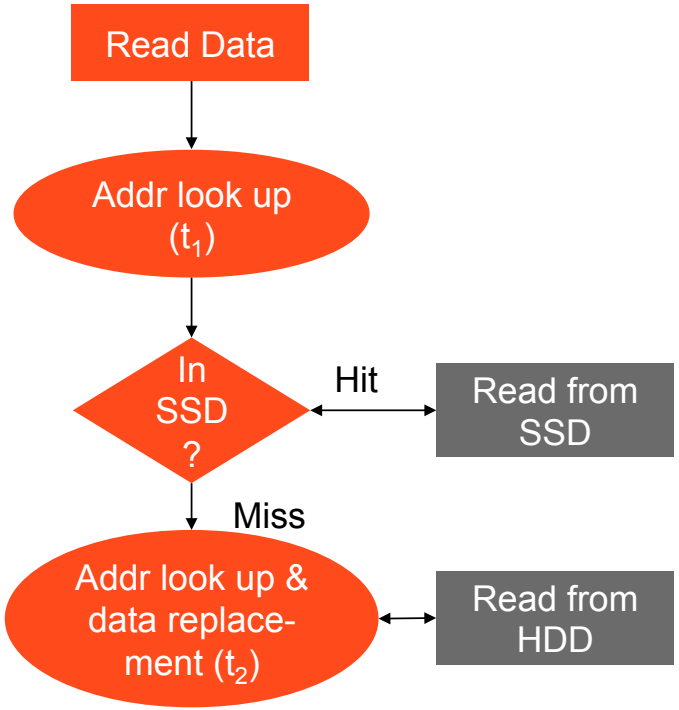




MEMORY HIERARCHIES ARE STILL NEEDED WITH NEW NVMS



BRIDGING THE GAP BETWEEN HDD AND SSD



	HDD	SSD
Access Time	10ms	100us

- Average access time =
  t_1 + Hit Rate x SSD access time
  + Miss Rate x (t_2+HDD access time)
- Design requirement – 1-2% penalty for Hybrid drive compared to SSD

TRADEOFFS BETWEEN HIT RATE, CACHE ENGINE DELAY, AND PERFORMANCE

1% PERFORMANCE (DELAY) PENALTY

t_1	1us	0.5us	0.1us
t_2	100us	50us	10us
Miss Rate Requirement	0 (Not Possible)	0.0050%	0.0091%

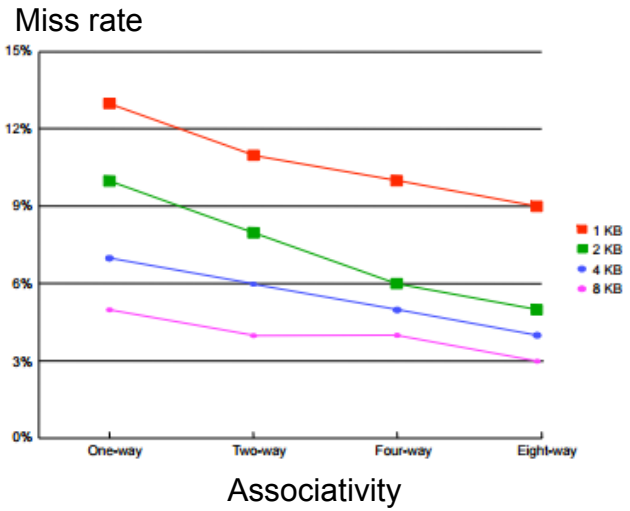
2% PERFORMANCE (DELAY) PENALTY

t_1	1us	0.5us	0.1us
t_2	100us	50us	10us
Miss Rate Requirement	0.0100%	0.0151%	0.0192%

- t_1 << SSD access time → Hardware based lookup
- t_2 << HDD access time → Software solution is acceptable
- Very high hit rate is required due to large performance gap between SSD and HDD
  - Large cache, high associativity (i.e., data can placed “anywhere” in the cache)

CURRENT MEMORY HIERARCHIES

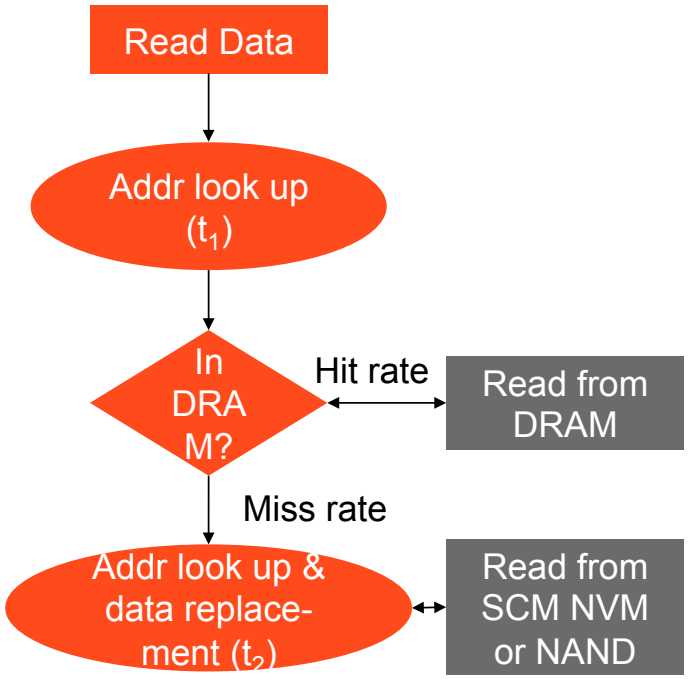
	L1 cache	L2 cache	Page (DRAM)	HDD
# of entries	100s	10,000s	millions	100s of millions
Block placement	2-4 ways associative	4-8 ways associative	Full set associative	
Lookup/Miss handling	Hardware	Hardware	Hardware /Software	Software



PERFORMANCE ← → HIT RATE, COST

Source: https://courses.cs.washington.edu/courses/cse378/09au/

USE STORAGE CLASS MEMORY NVM AS CACHE

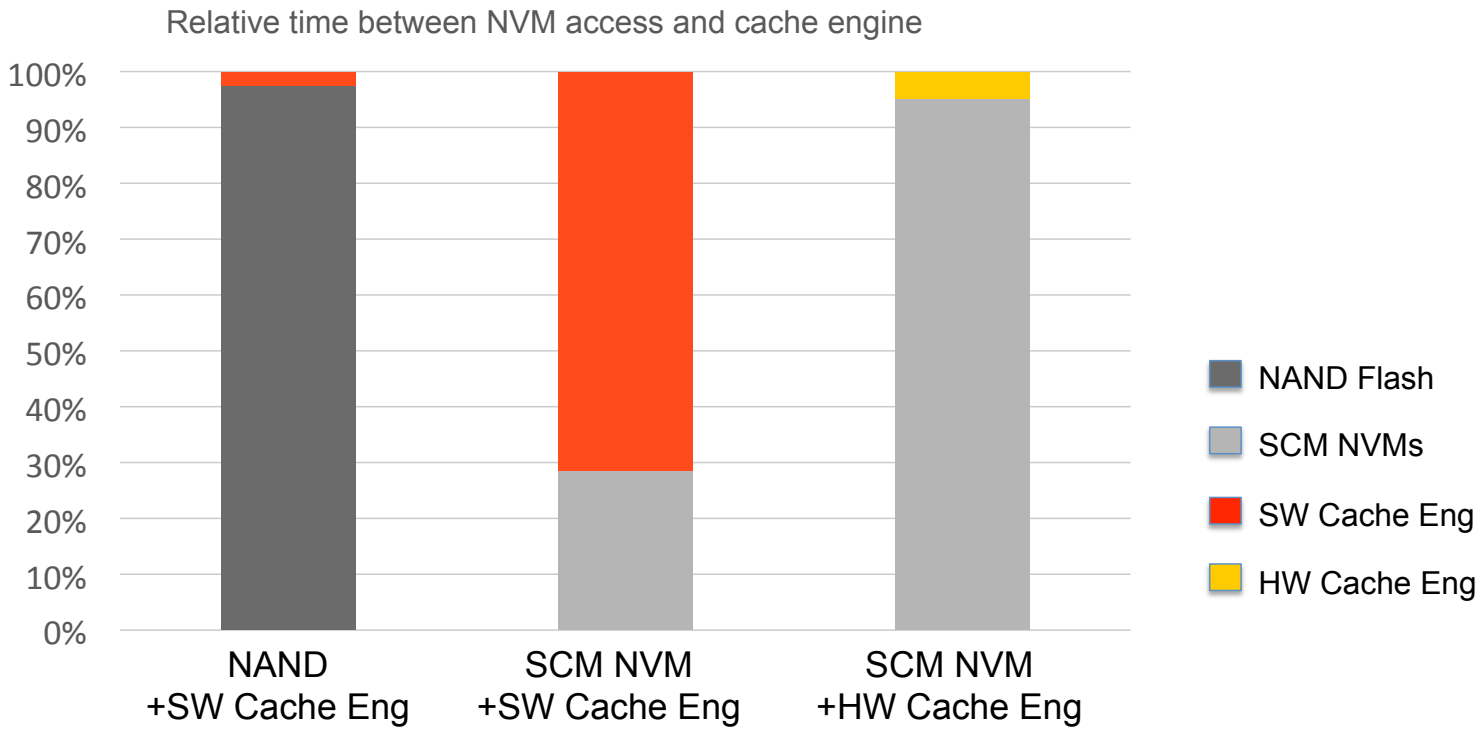


	NAND SLC	Storage Class Memory NVM	DRAM
Access Time	20us	200ns	20ns

Achieving 0.1% DRAM access delay penalty requires low miss rate and low overhead cache engine

Miss Rate	0.005%	0.001%	0.0001%	0.00005%
$t_{2\_SCM\ NVM}$ (us)	0.2	1.8	19.8	39.8
t_{2_NAND} (us)	Not possible	Not possible	Not possible	20.0

HARDWARE CACHE ENGINE IS NEEDED FOR SCM NVM



UTILIZING SCM NVMs AS A CACHE

	L1 cache	L2 cache	Page (DRAM)	SCM NVM	SSD
# of entries	100s	10,000s	100,000s - millions	Millions	100s of millions
Block placement	2-4 ways associative	4-8 ways associative	Full set associative	Full set associative	
Lookup/Miss handling	Hardware	Hardware	Hardware	Hardware	Software

DESIGN CHALLENGE
High hit rate →Large cache size, full set associative
Low overhead →Small cache size, non set associative

MARVELL FINAL LEVEL CACHE™ (FLC™) CACHE ENGINE

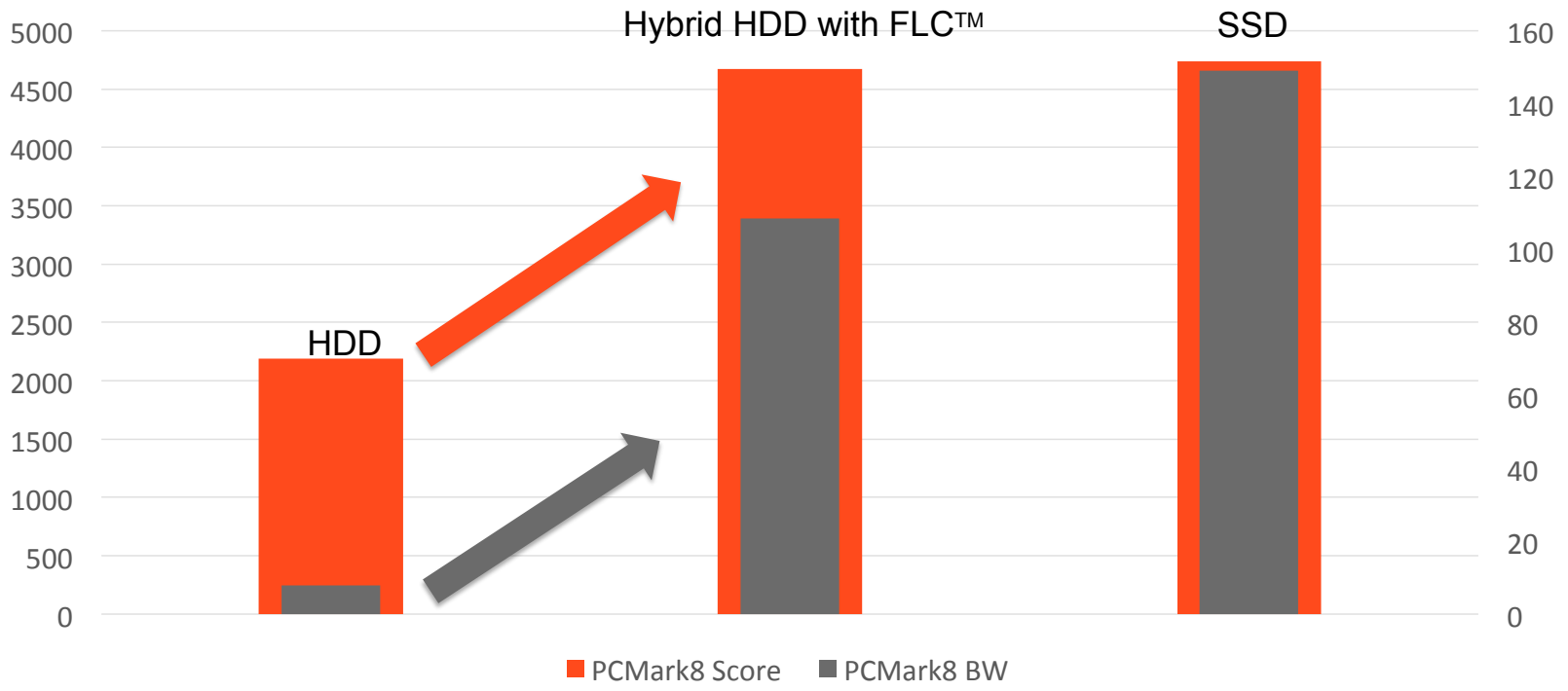
LOW MISS RATE

- Support very large cache (GB)
- Large cache line (16KB or larger)
- Full set associative

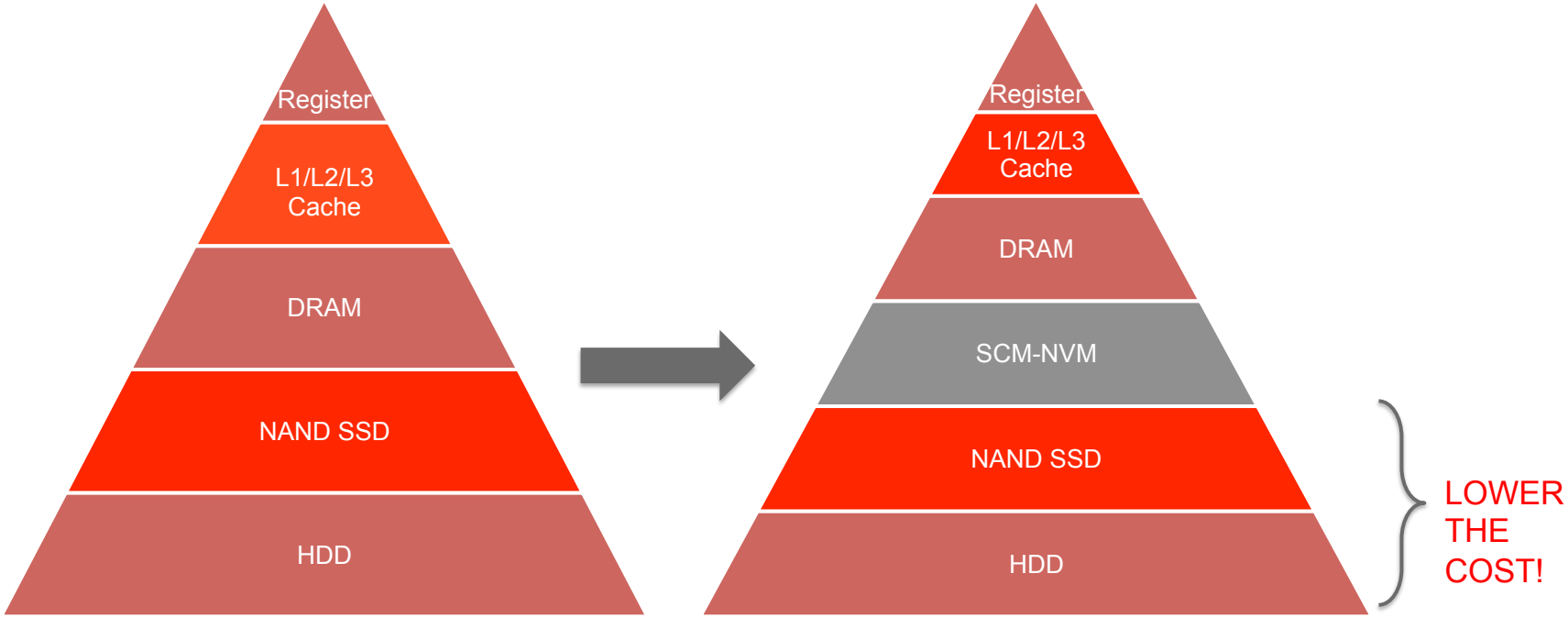
LOW OVERHEAD

- Pseudo CAM based design
- Hardware based pseudo-LRU replacement scheme
- <10 clock cycle latency

PERFORMANCE OF FLC™ (TESTED ON HYBRID HDD)



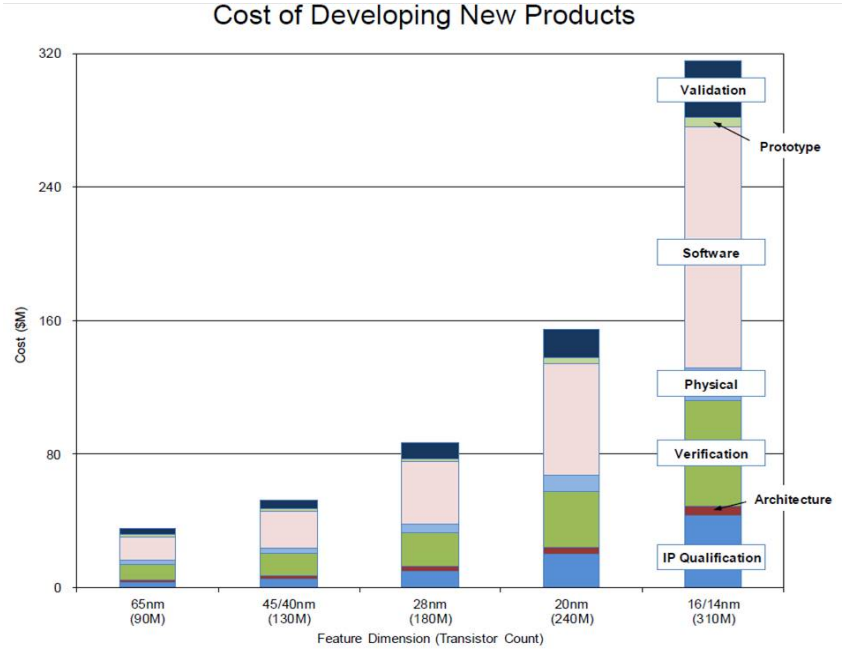
LOWER TIERS SHOULD FOCUS ON COST



ECONOMY OF CHIP DESIGN DOESN'T WORK OUT

$$\text{UNIT COST} = \frac{\text{R\&D COST + MASK COST + PACKAGE NRE}}{\text{\# OF CHIPS}} + \text{PER DIE COST} + \text{PER PACKAGE COST}$$

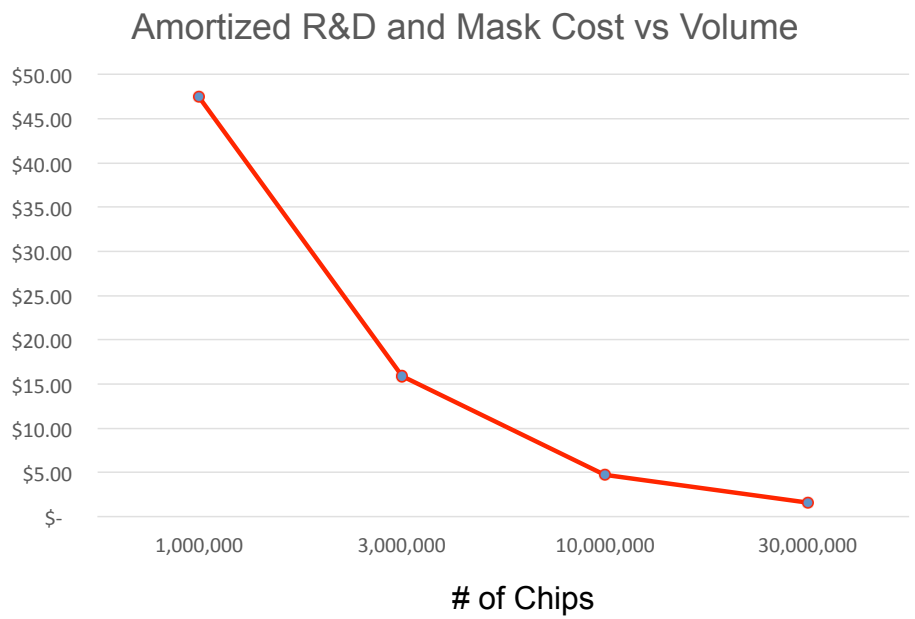
R&D cost	Increase
Mask cost	Increase
Package NRE	Increase
# of chip	Decrease
Per dies cost	Decrease
Per package cost	Flat or decrease



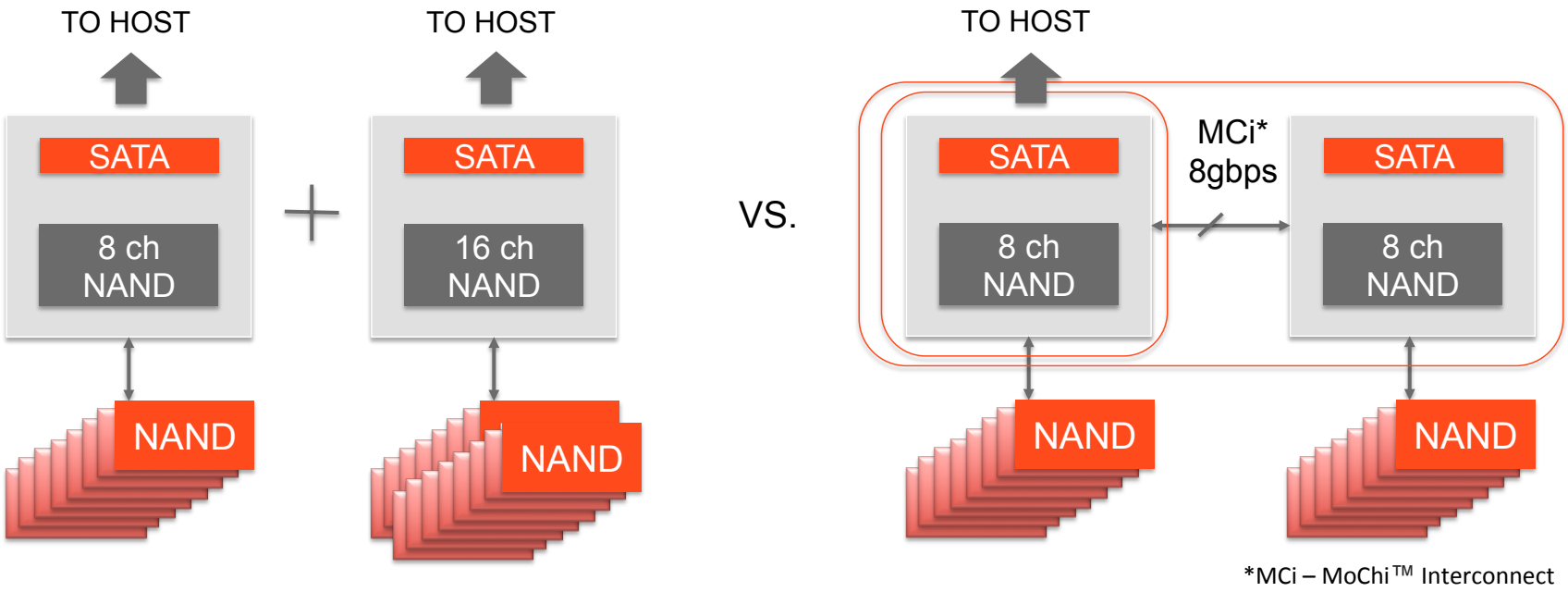
Source: http://semiengineering.com/how-much-will-that-chip-cost/

NEED HIGH VOLUME TO AMORTIZE DEVELOPMENT COST

ASSUMING A DERIVATIVE CONTROLLER - R&D COST $40M, MASK AND PACKAGING NRE $7.5M



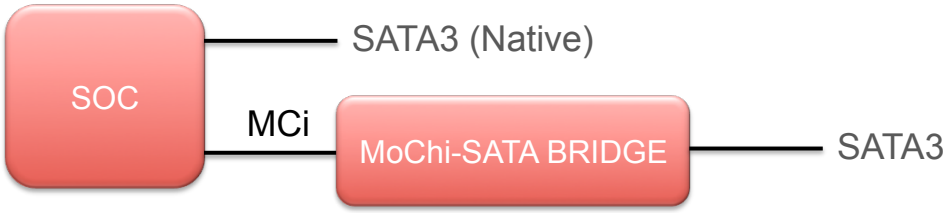
EXPANDABLE SSD CONTROLLERS



DEVELOPMENT COST: $50M + $20M

DEVELOPMENT COST: $50M

MoChi™ INTERCONNECT (MCi) IS TRANSPARENT



	READ bandwidth (MB/s)	WRITE Bandwidth (MB/s)	R/W Mixed Bandwidth (MB/s)
Native SATA	515	490	500
SATA through MCi	492	510	501

CONCLUSIONS

IT'S ALL ABOUT DATA ACCESS – GET DATA WHEN NEEDED, AT THE RIGHT COST

MULTI-TIERED DESIGN IS ESSENTIAL IN STORAGE SYSTEMS

- EVEN MORE IMPORTANT WITH NEW SCM NVMS

HARDWARE-BASED CACHE ENGINE IS NEEDED FOR HIGH PERFORMANCE TIERS

EXPANDABLE CONTROLLERS CAN HELP REDUCE DEVELOPMENT COST FOR LOW TIERS

