New Directions in Memory Architecture

June 12, 2014

Bob Brennan,
Senior Vice President
Memory Solutions Lab
Bob.Brennan@Samsung.com

Ecosystem MSL Research Development

Memory Solutions Lab

SAMSUNG

Legal Disclaimer

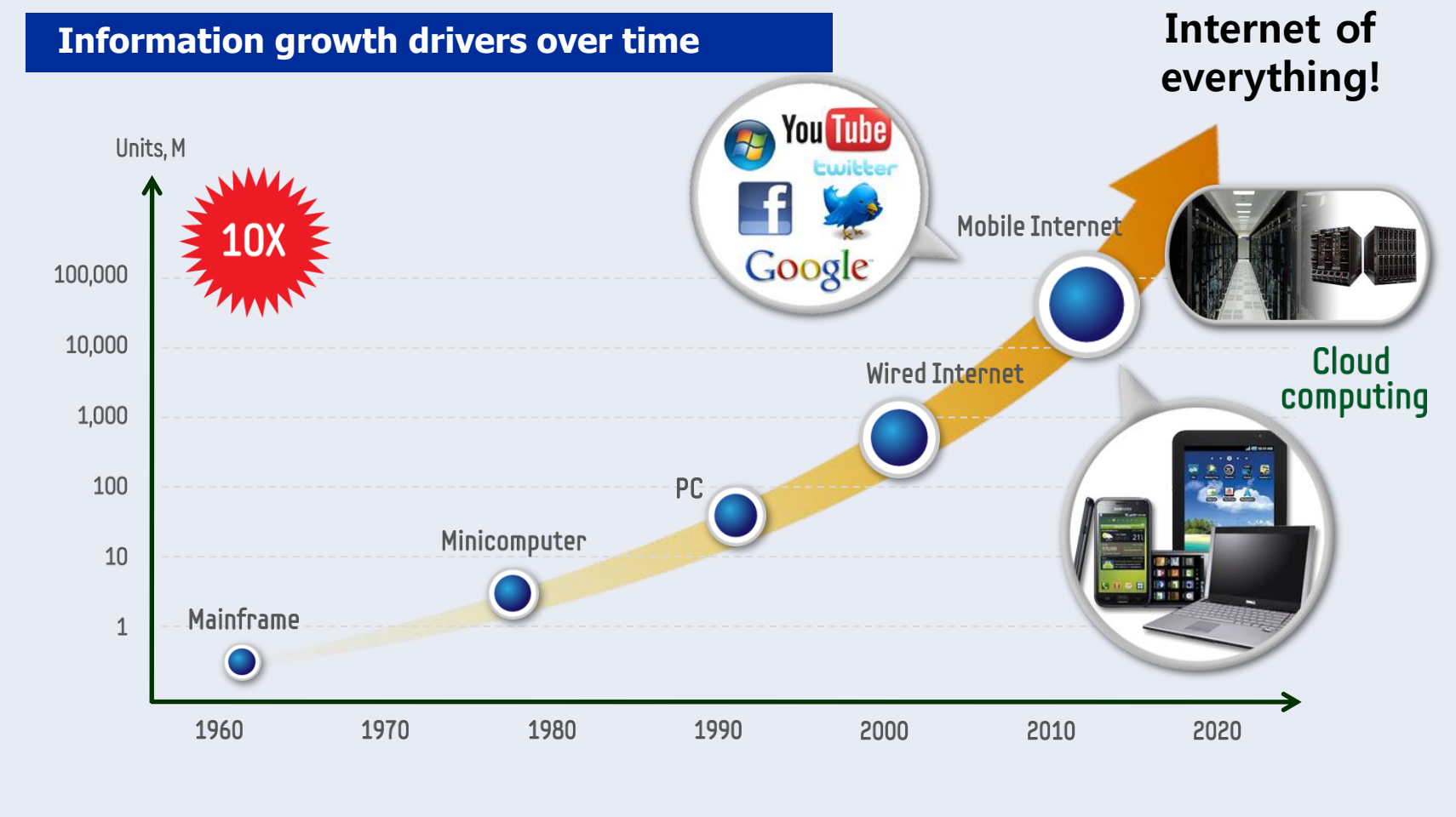
This presentation is intended to provide information concerning memory industry trends. We do our best to make sure that information presented is accurate and fully up-to-date. However, the presentation may be subject to technical inaccuracies, information that is not up-to-date or typographical errors. As a consequence, Samsung does not in any way guarantee the accuracy or completeness of information provided on this presentation. Samsung reserves the right to make improvements, corrections and/or changes to this presentation at any time.

The information in this presentation or accompanying oral statements may include forward-looking statements. These forward-looking statements include all matters that are not historical facts, statements regarding the Samsung Electronics' intentions, beliefs or current expectations concerning, among other things, market prospects, growth, strategies, and the industry in which Samsung operates. By their nature, forward-looking statements involve risks and uncertainties, because they relate to events and depend on circumstances that may or may not occur in the future. Samsung cautions you that forward looking statements are not guarantees of future performance and that the actual developments of Samsung, the market, or industry in which Samsung operates may differ materially from those made or suggested by the forward-looking statements contained in this presentation or in the accompanying oral statements. In addition, even if the information contained herein or the oral statements are shown to be accurate, those developments may not be indicative developments in future periods.

Agenda

- **Environment – BW & Capacity growth**
- **DRAM – BW & Capacity -> Tiering**
- **Flash –Scales, Gets Intelligent, Tiers**
- **New "Persistent Performance"**

Environment: Mobile & Cloud



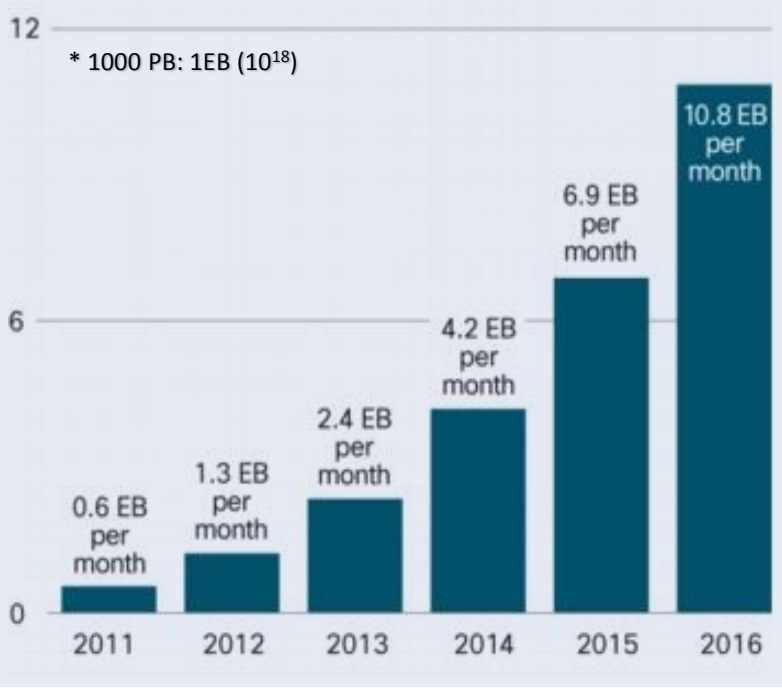
2012: Mobile connected devices exceeded the world's population

Environment: Datacenter Infrastructure

More applications for data

Billions of Devices!

Data traffic: 78% CAGR



Source: Cisco Visual Networking Index

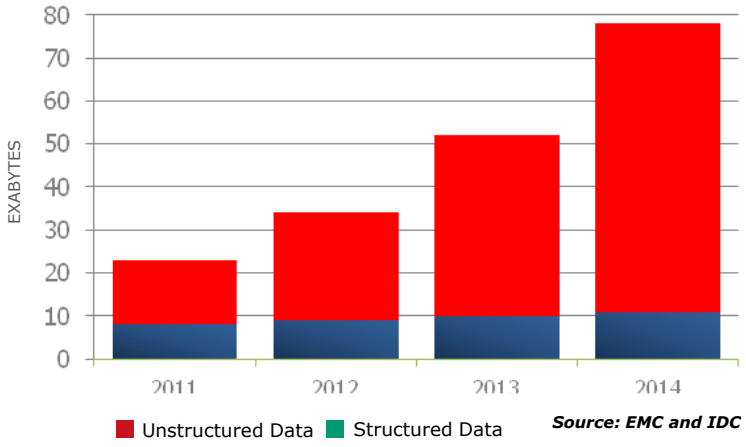
More video is uploaded to YouTube in one month than the 3 major US networks created in 60 years

What about Exabytes?

5 EB: Total data created between the dawn of civilization and 2003

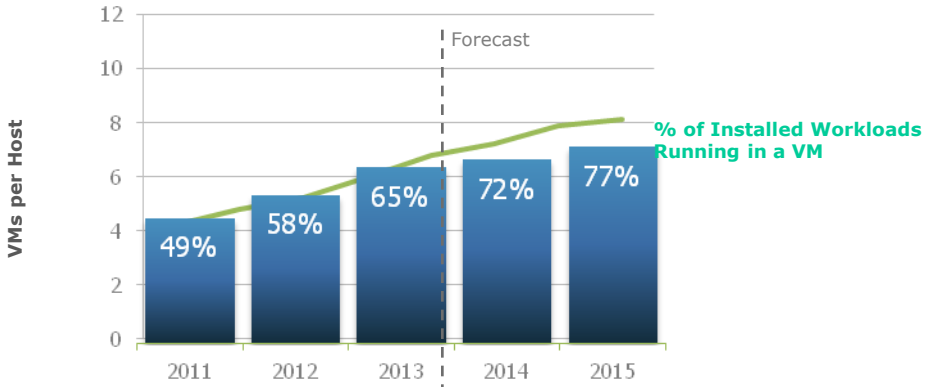
Environment: Escalating Demand for DRAM and Storage

In-Memory Analytics for Big Data



Source: EMC and IDC

Growing x86 Server Virtualization Density



Source: Gartner and IDC

Escalating Memory-Intensive Workloads

HPC

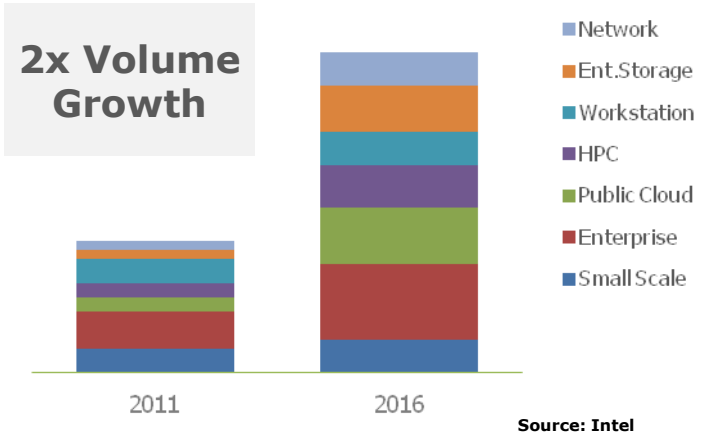
Financial

Graphics

Gaming

Big Data

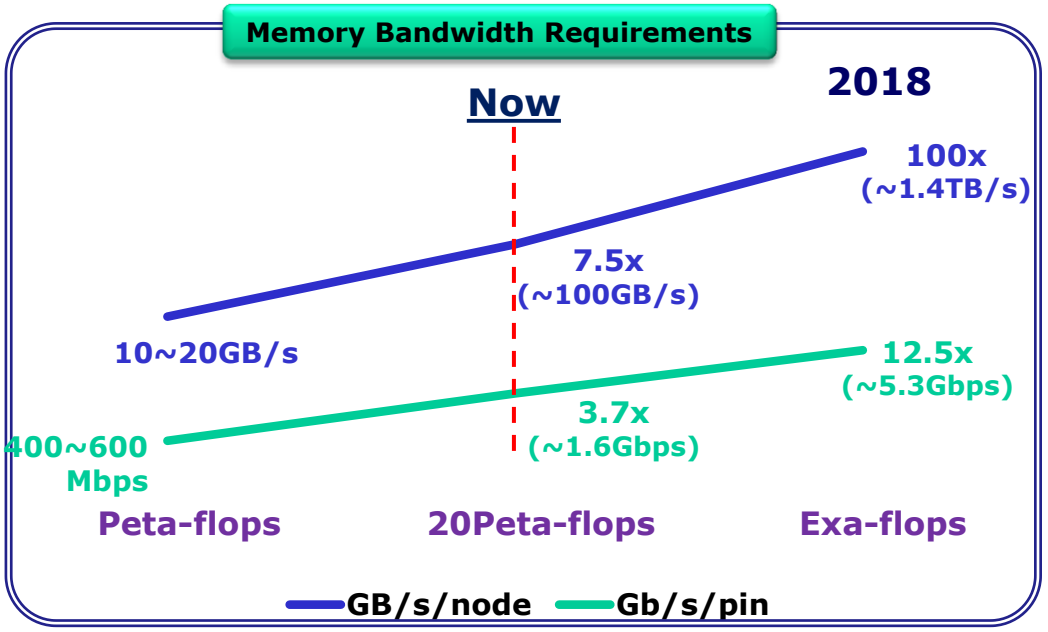
Data Center Processor Growth



Source: Intel

Environment – Bandwidth Demand

	FHD (1920x1080)	UD (3840x2160)
Display		
Camera	13MP	20+MP
Video	1080p	4K
N-screen	F-HD	UHD



[Source: "Memory systems for PetaFlop to ExaFlop class machines" by IBM, 2007 & 2010]

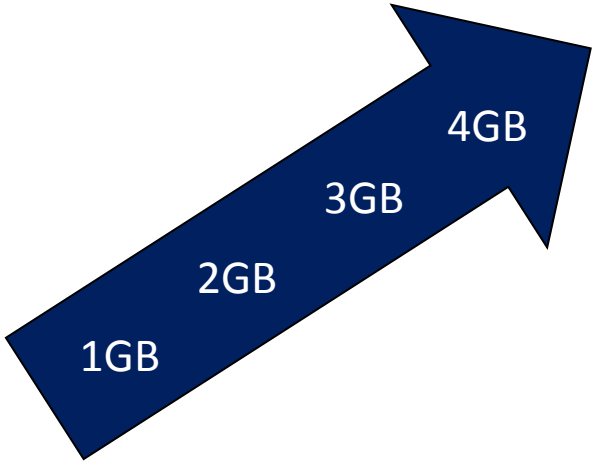
Mobile:
Display/GFX/Camera

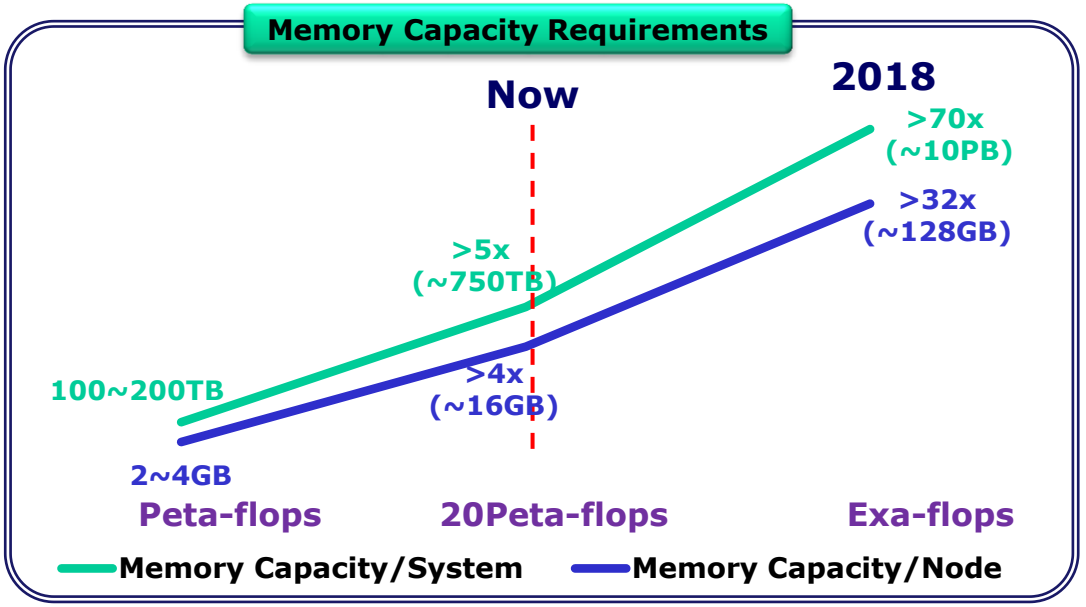
Exponential Bandwidth Demand

Server:
Core Scaling

Linear to Exponential Bandwidth Demand

Environment – Capacity Demand





[Source: "Memory systems for PetaFlop to ExaFlop class machines" by IBM, 2007 & 2010]

Mobile:
Display/GFX/Camera

~Linear Capacity Demand

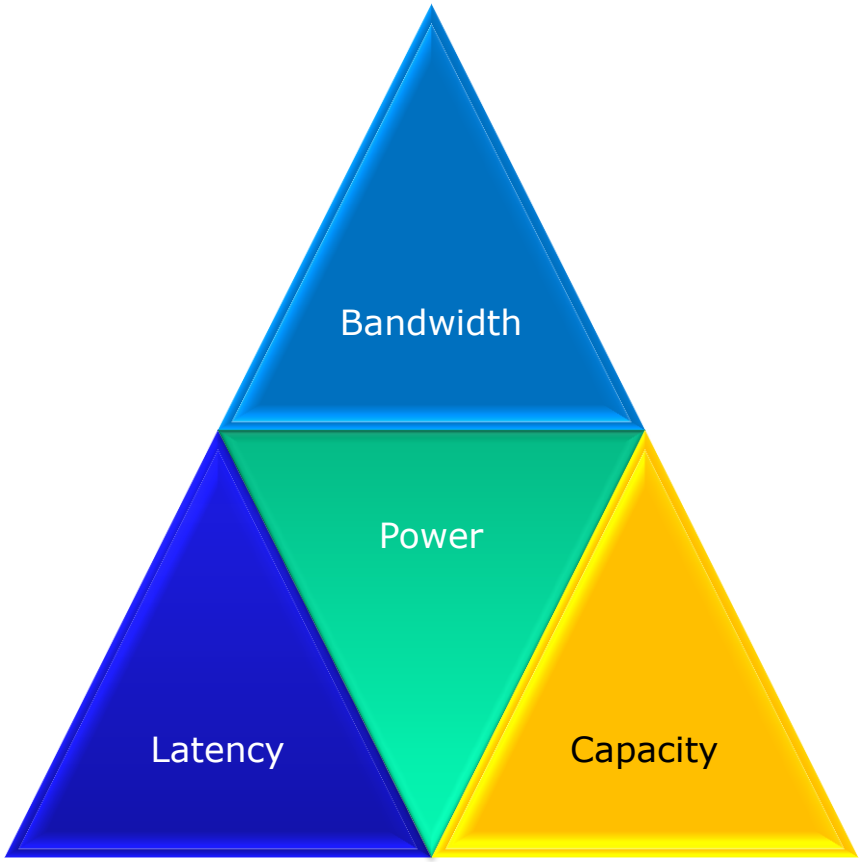
Server:
Core Scaling

Linear - Exponential Capacity Demand

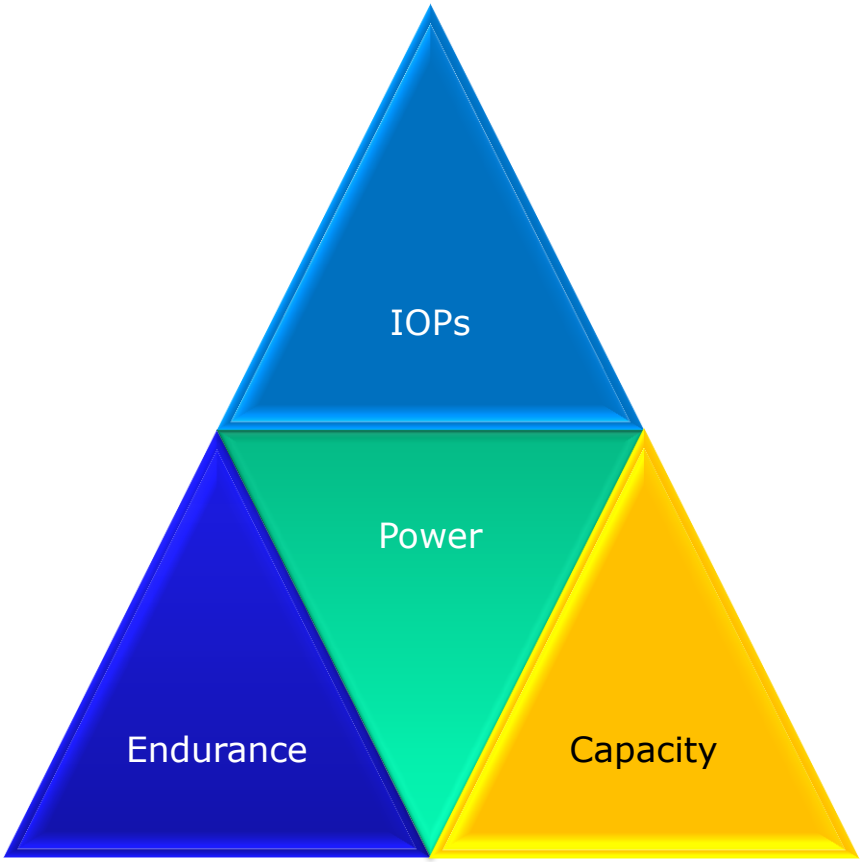
Agenda

- Environment – BW & Capacity growth
- **DRAM – BW & Capacity -> Tiering**
- Flash –Scales, Gets Intelligent, Tiers
- New “Persistent Performance”

The “Trade-off Triangles”

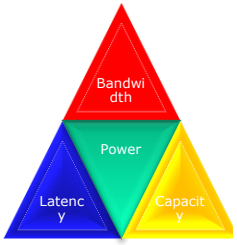


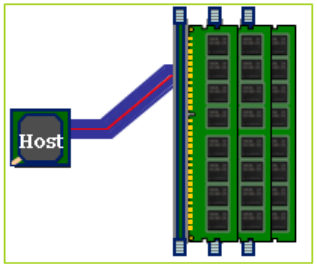
DRAM



Non-Volatile

DRAM: Bandwidth Scaling





[Mbps]

4200

3600

3200

2400/2667

2133

1866

1600

1333

2011

2015

2018

[Year]

DDR4

DDR Wall?

DDR5 (?) & New I/F (?)

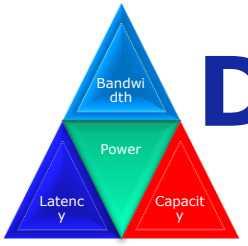
Optical (?)

Subject to cost/energy efficiency, scaling, …

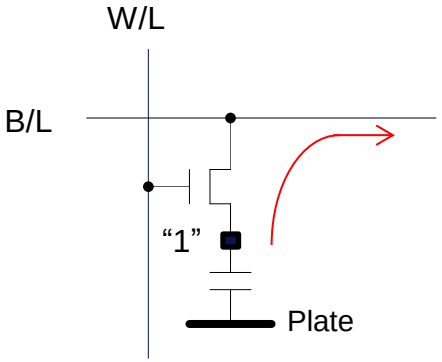
Multi-Drop Bus Challenge:

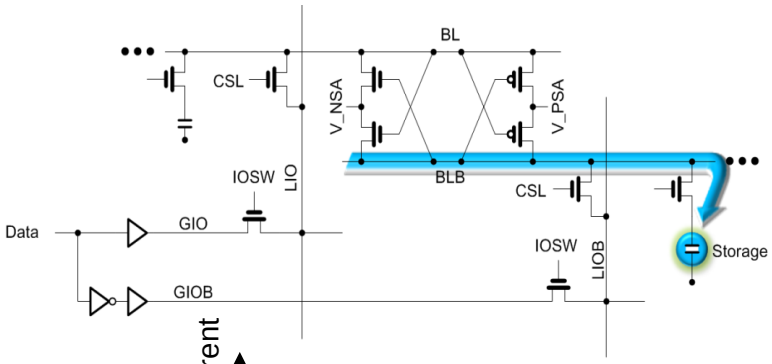
Higher BW, Lower VDD

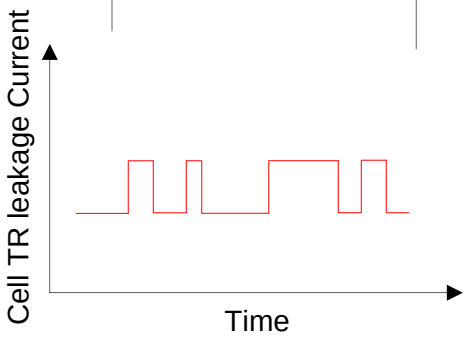
New Solution Needed



DRAM: Scaling Challenges







❖ Refresh

- Difficult to build high-aspect ratio cell capacitors decreasing cell capacitance
- Leakage current of cell access transistors increasing

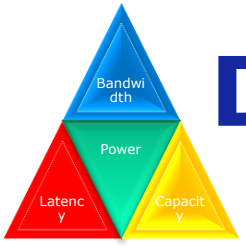
❖ tWR

- Contact resistance between the cell capacitor and access transistor increasing
- On-current of the cell access transistor decreasing
- Bit-line resistance increasing

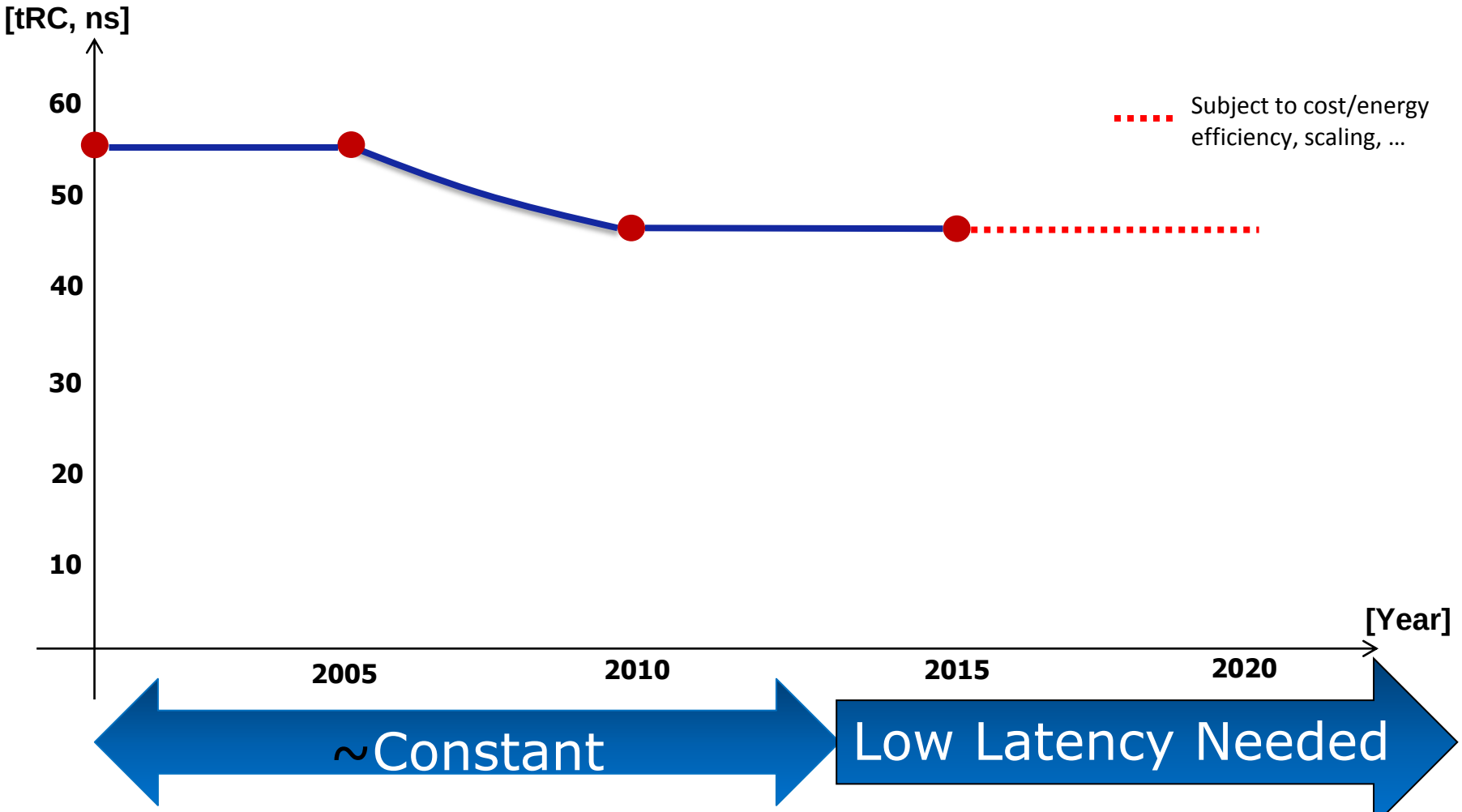
❖ VRT

- As cell capacitance shrinks, more frequent

Innovation needed to scale to 10nm & beyond

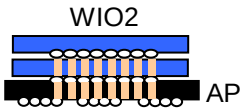
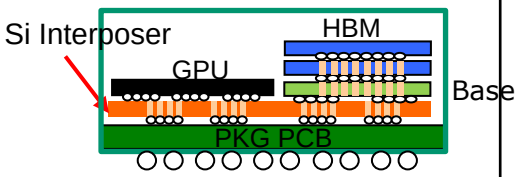


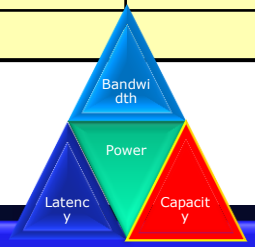
DRAM: Latency Challenge



Disruptive Solution Needed

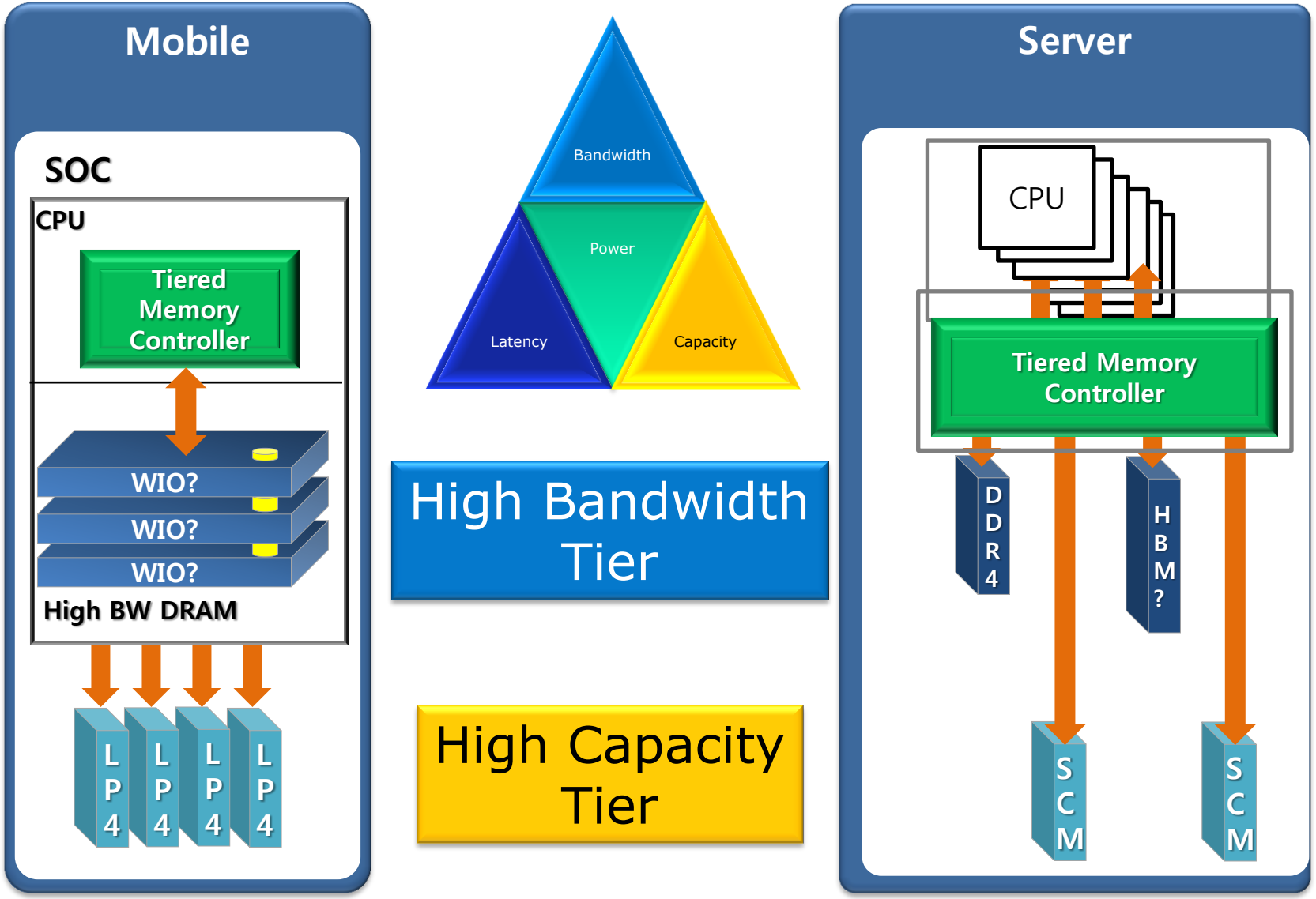
DRAM: "Go Wide" for Bandwidth

ITEM		Mobile WIO2	HBM (High B/W Memory)
		DRAM 	Base die + DRAM 
Bottom die		N/A	Buffering & Signal re-routing
BW (GB/s)		25.6~51.2	128~256
Pin	Speed	0.4~0.8 Gbps	1~2 Gbps
	# I/O	512	1,024
#Bump	Logic	1~2K	6K~8K
	DRAM	1~2K	~3K
Cube (GB)		1 / 2	1 / 2 / 4
# TSV stack		1 / 2 / 4	1 / 2 / 4
DRAM density		8Gb	8Gb
Application	GFX card	○	○
	ULT	○	-
	HPC	-	○
	Server	-	○**(Cache)**
	Mobile	○	-



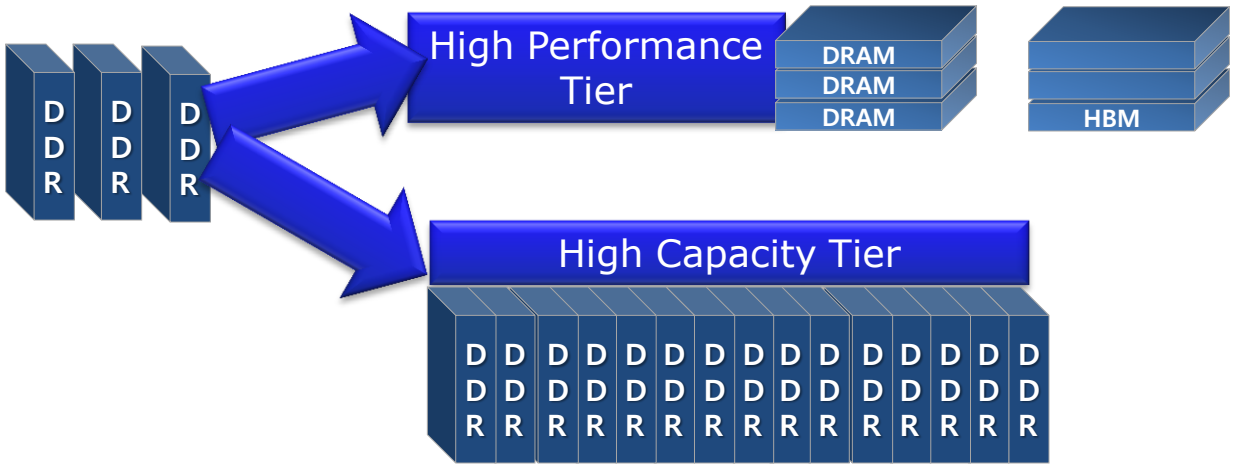
Good BW & Latency – Still Need Capacity

DRAM: Hybrid Memory Systems



Tiered Capacity, Tiered Latency, TL-DRAM?

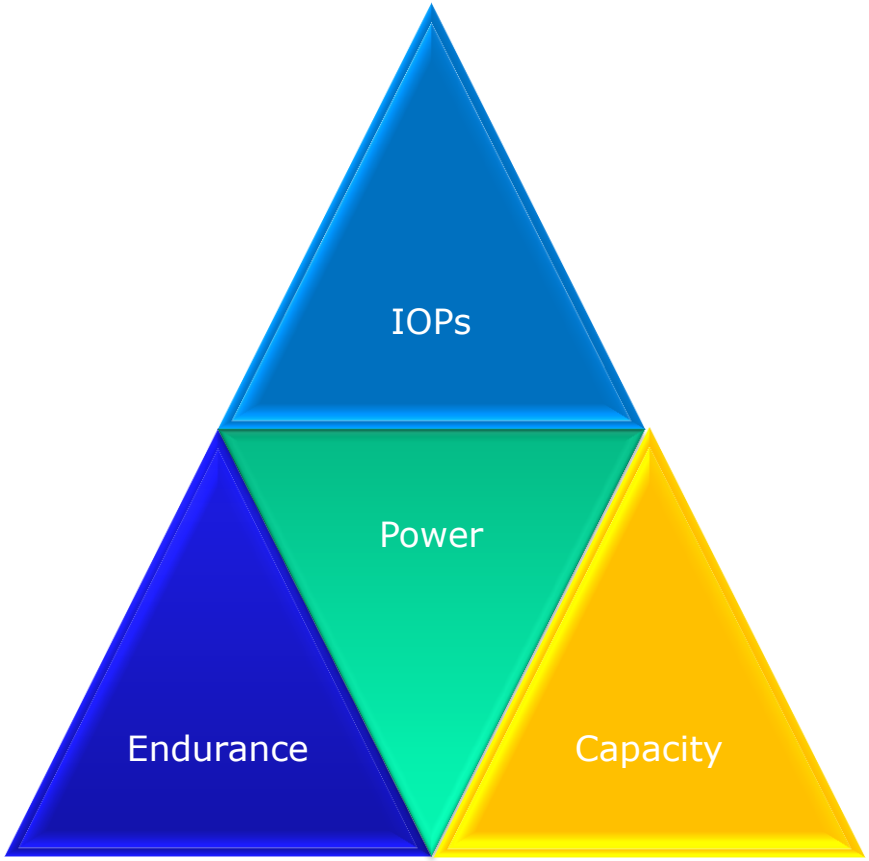
1st Step: System Tiering DRAM

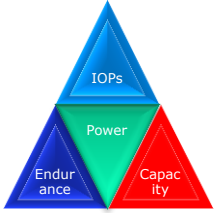


Agenda

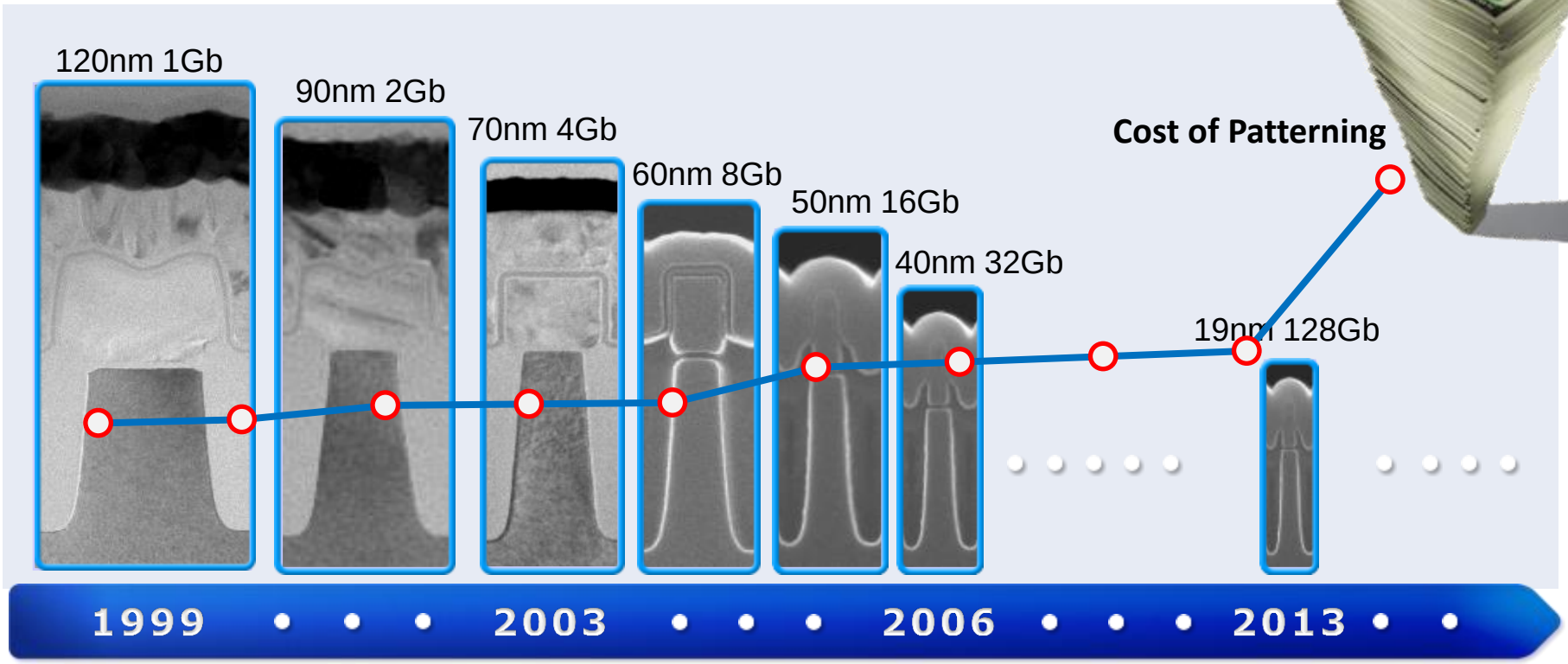
- Environment – BW & Capacity growth
- DRAM – BW & Capacity -> Tiering
- **Flash – Scales, Becomes Intelligent, Tiers**
- New "Persistent Performance"

Flash



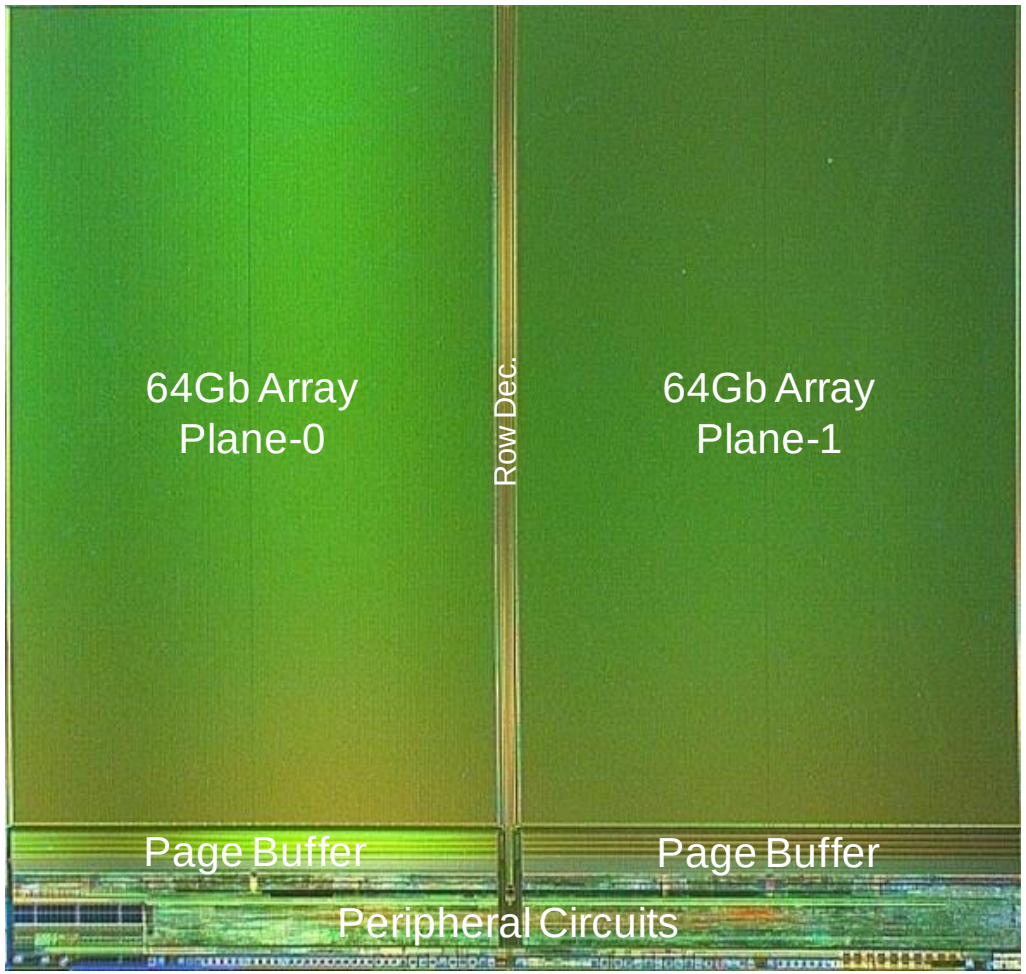


Flash: Capacity Scaling



Scaling Becomes Difficult – Need a New Solution

Breakthrough: 128Gb V-NAND



- **Vertical-NAND Technology**

- **Chip Size**
  : 133mm² → 0.96Gb/mm²

- **24-WL Stacked Layers**

- **64Gb Array × 2-Plane**

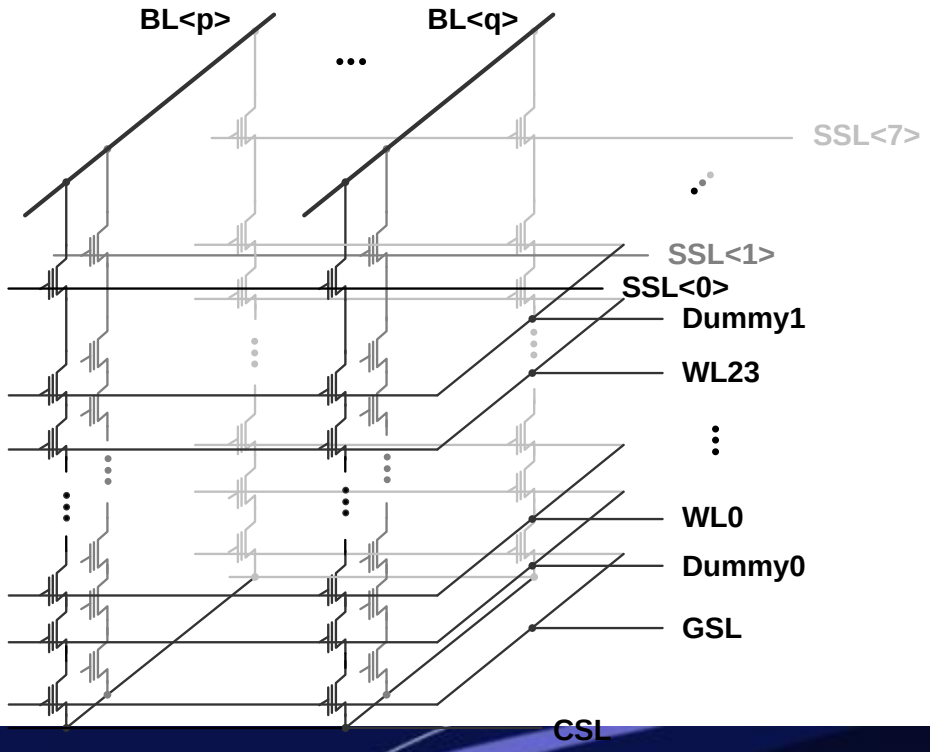
- **One-sided Page Buffer**
  : (8KB x 2) Page Size

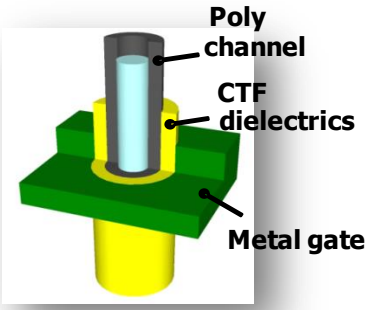
- **Asynchronous DDR Interface**
  : Wave-pipeline datapath
  : 667Mbps at Mono Die
  : 533Mbps at 8-stacked Dies

World's 1st 3D V-NAND Mass Production Flash

V-NAND Array Structure

- ✓ Advanced V-NAND Technology with Damascened Metal Gate
- ✓ Cell : All-around Gate Structure + Charge Trap Flash
- ✓ String : 24-WL + 2-DWL + 2-Select WL
- ✓ Block : 8 Strings with Shared BL (8KB)



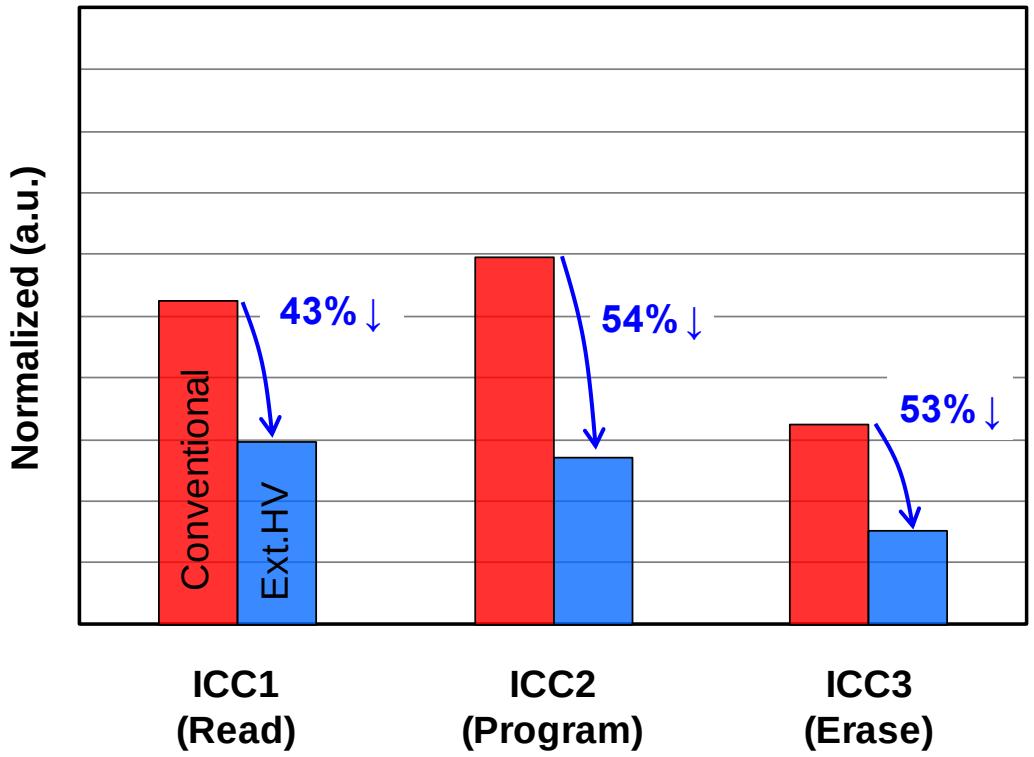


V-NAND Features

Bits per Cell	**2**
Density	**128Gb**
Technology	**Three Dimensional** **Vertical NAND, 3-metals**
Organization	**8KB × 384 pages** **× 5464 blocks × 8**
Program Performance	**50MB/s for Embedded App.,** **36MB/s for Enterprise SSD**
Data Interface Speed	**667Mbps@Mono, 533Mbps@8-stack**
Power Supply	**Vcc=3.3V / Vccq=1.8V**

Measured Active Power Improves

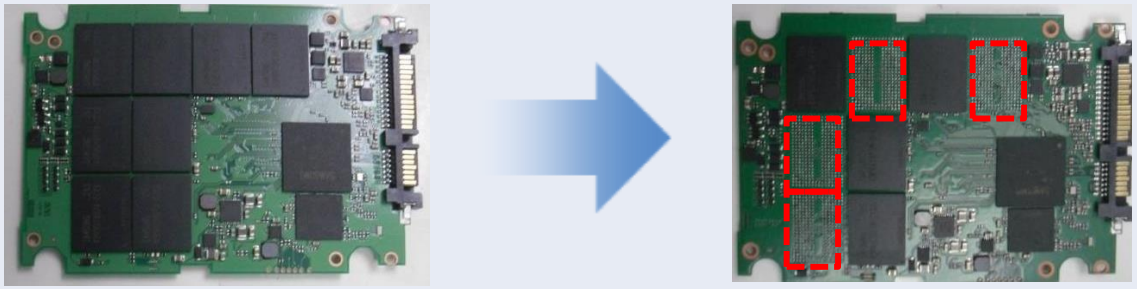
✓ **Over 50% Lower Energy Advantage is achieved**
→ Increasing overall SSD Performance
by using 8-way Interleaving NAND Operation



Enterprise SSD Comparison

Smaller Real Estate

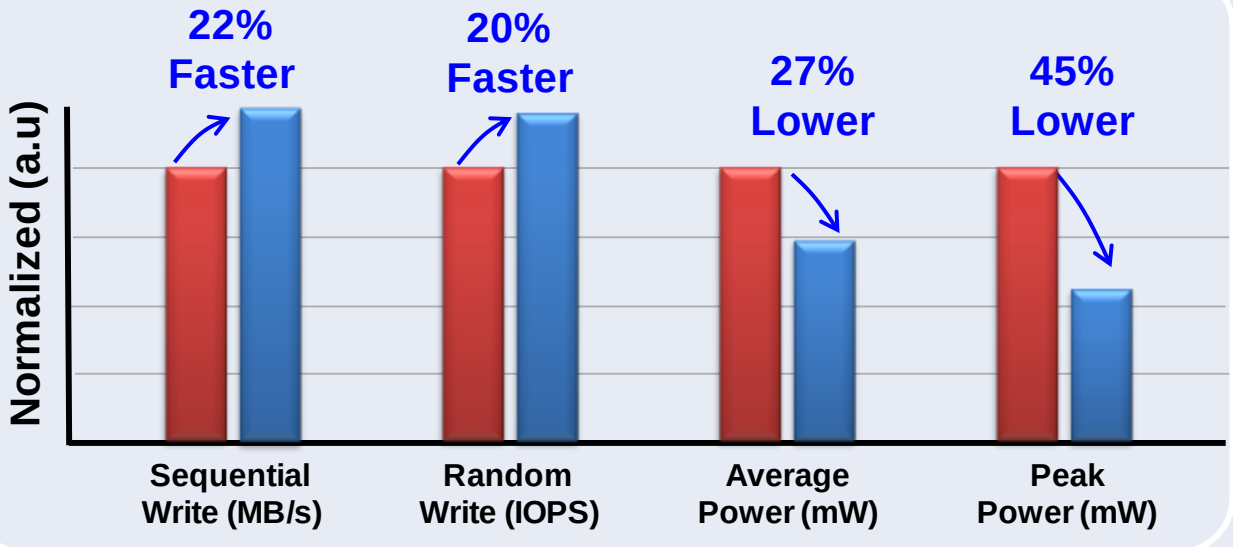
512GB Ep-SSD



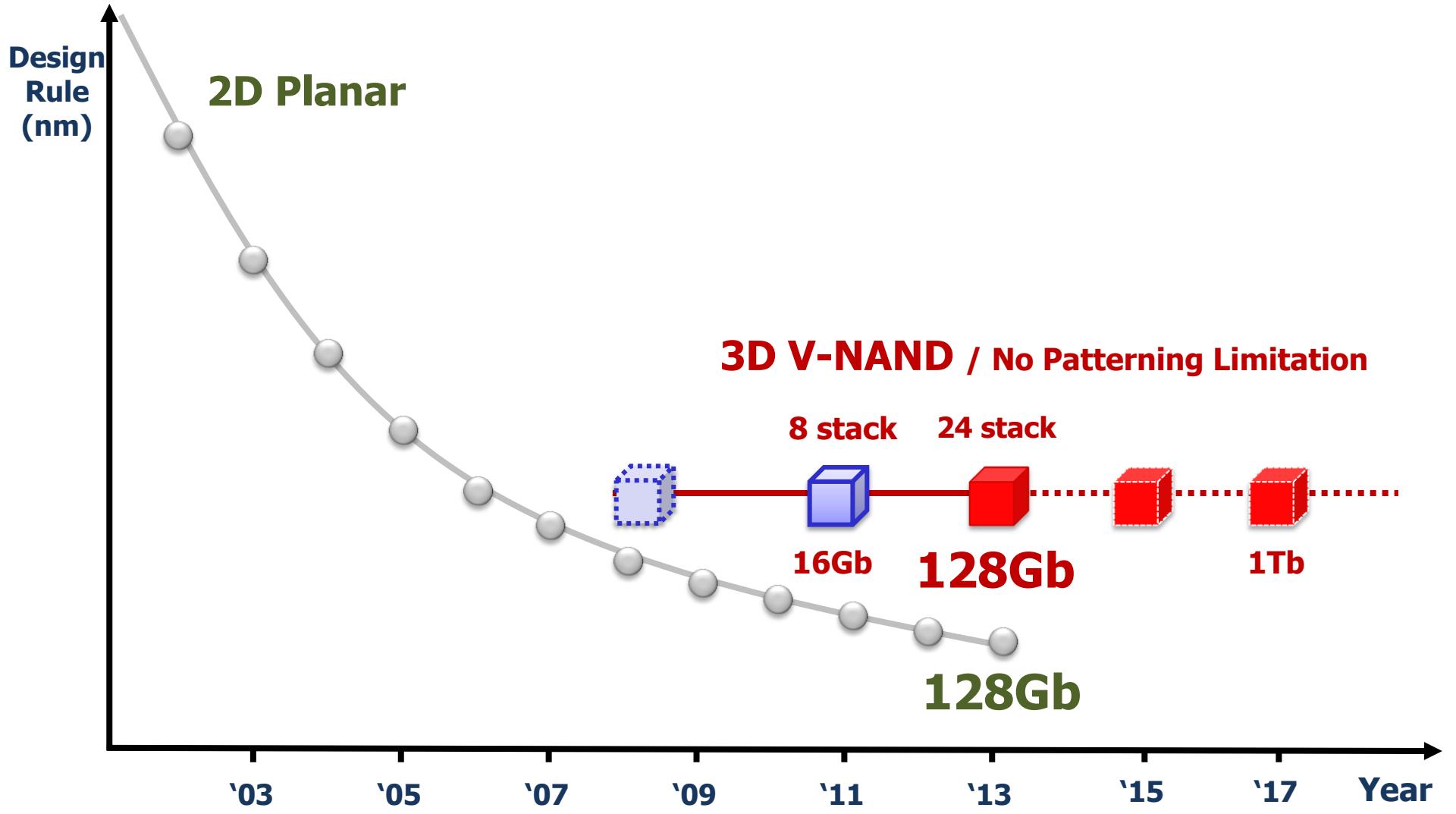
Planar NAND SSD
(8-ch, 8-way)

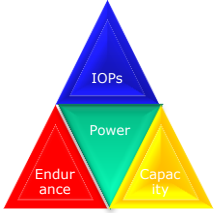
3D V-NAND SSD
(8-ch, 4-way)

Higher Performance



Flash: Scaling Continues

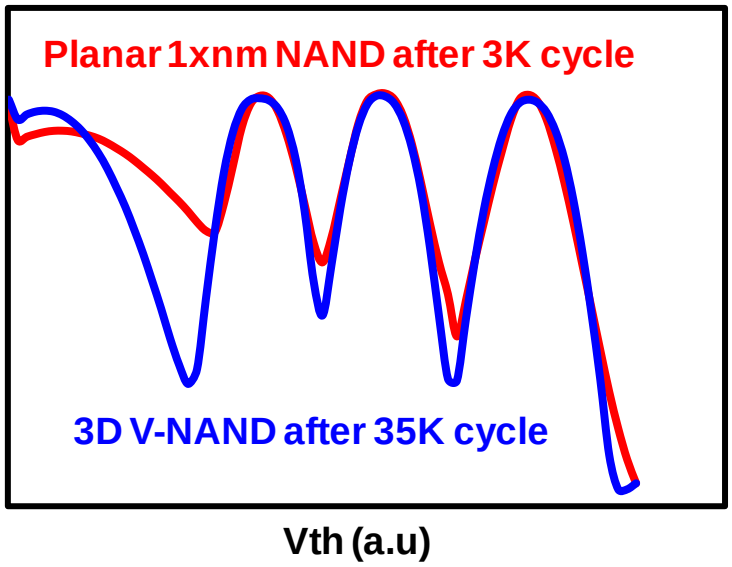


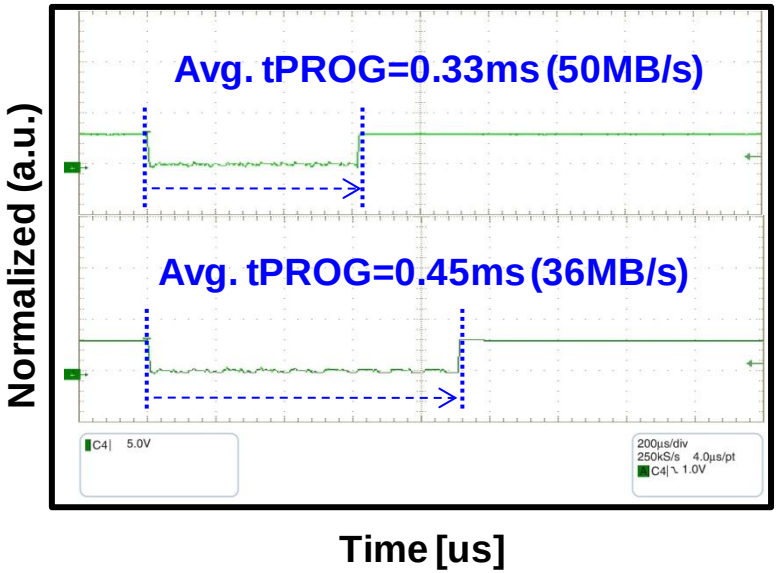


NAND Flash Endurance

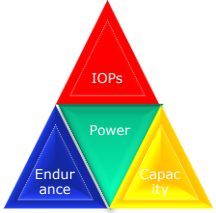
Flash: MLC Endurance

✓ 36MB/s + 35K Endurance
for Data-center & Enterprise SSD Applications
✓ 50MB/s + 3K Endurance for Mobile Applications



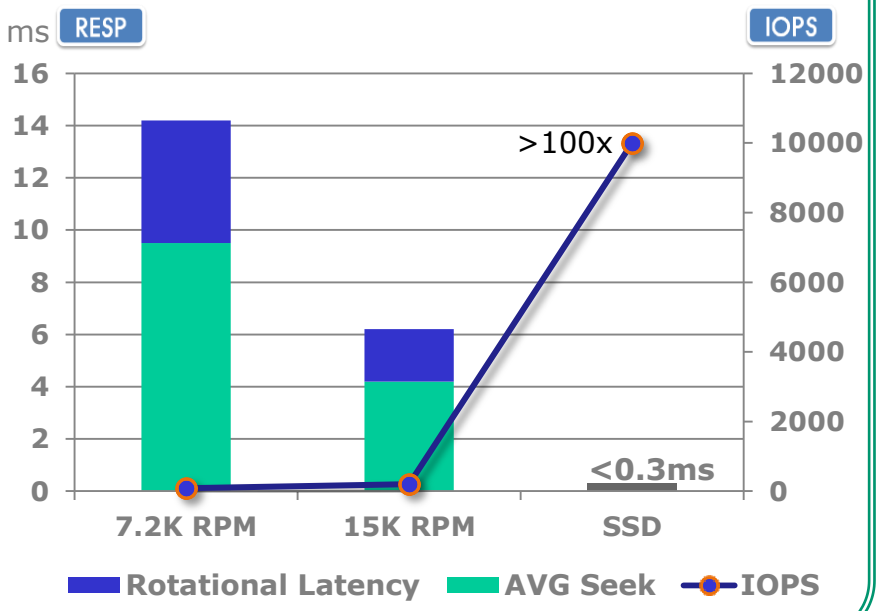


Endurance improved dramatically

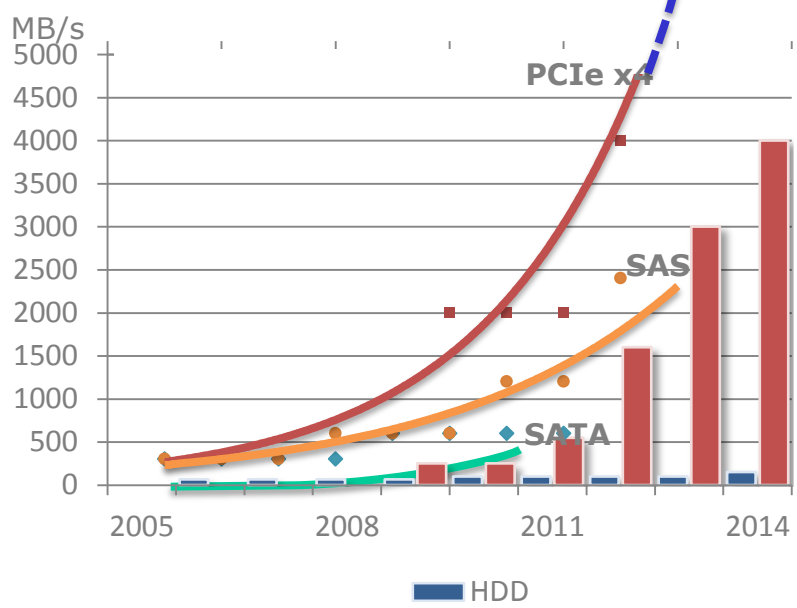


Flash: Performance

Latency & IOPS



Interface & Performance



Interface Unlocks Bandwidth: PCIeG2->G3->G4

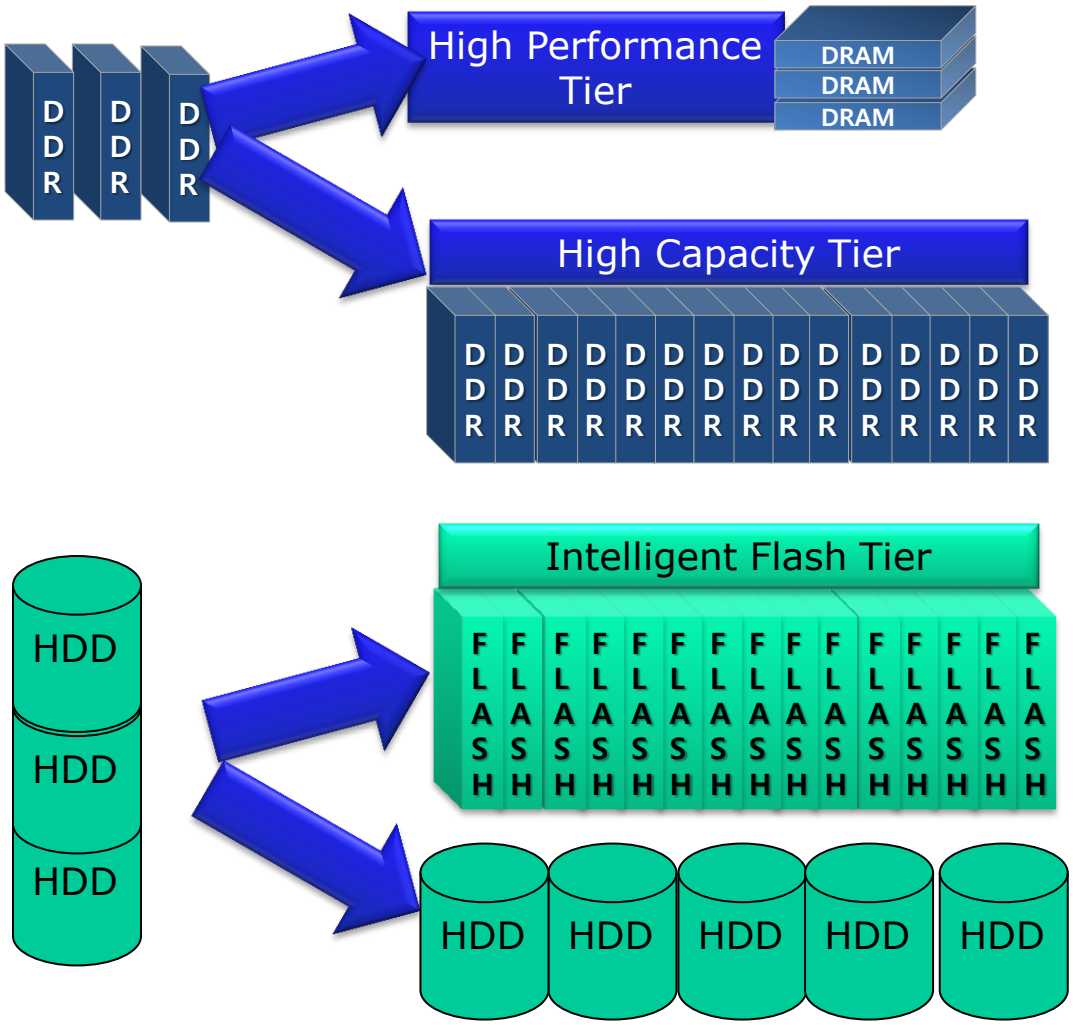
Solution needs to scale: Controllers, Algorithms, & Flash Organization

Flash: Inherent Intelligence

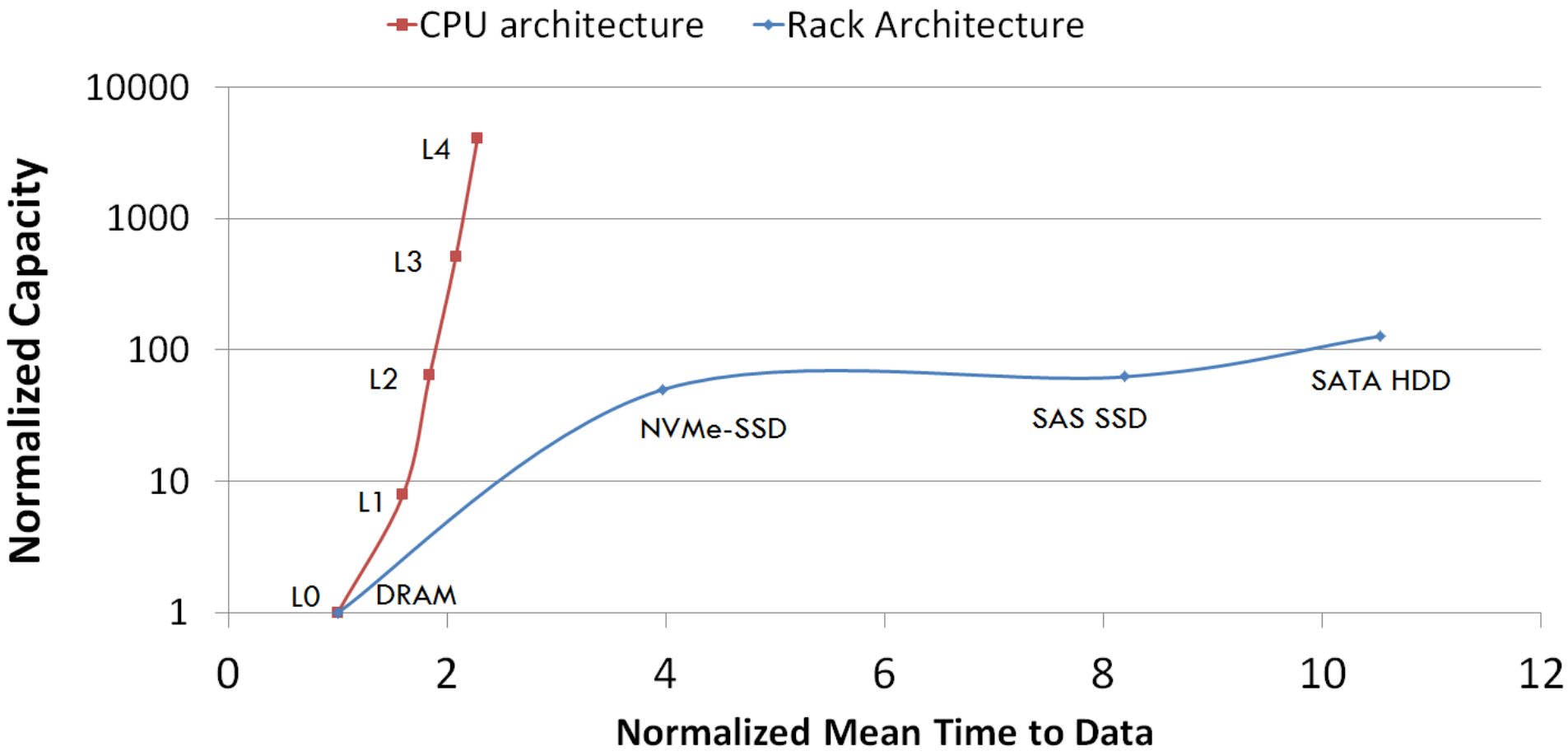
Endurance

3D Scaling

2nd Step: System Tiering Flash/HDDs



Today's Rack Scaling



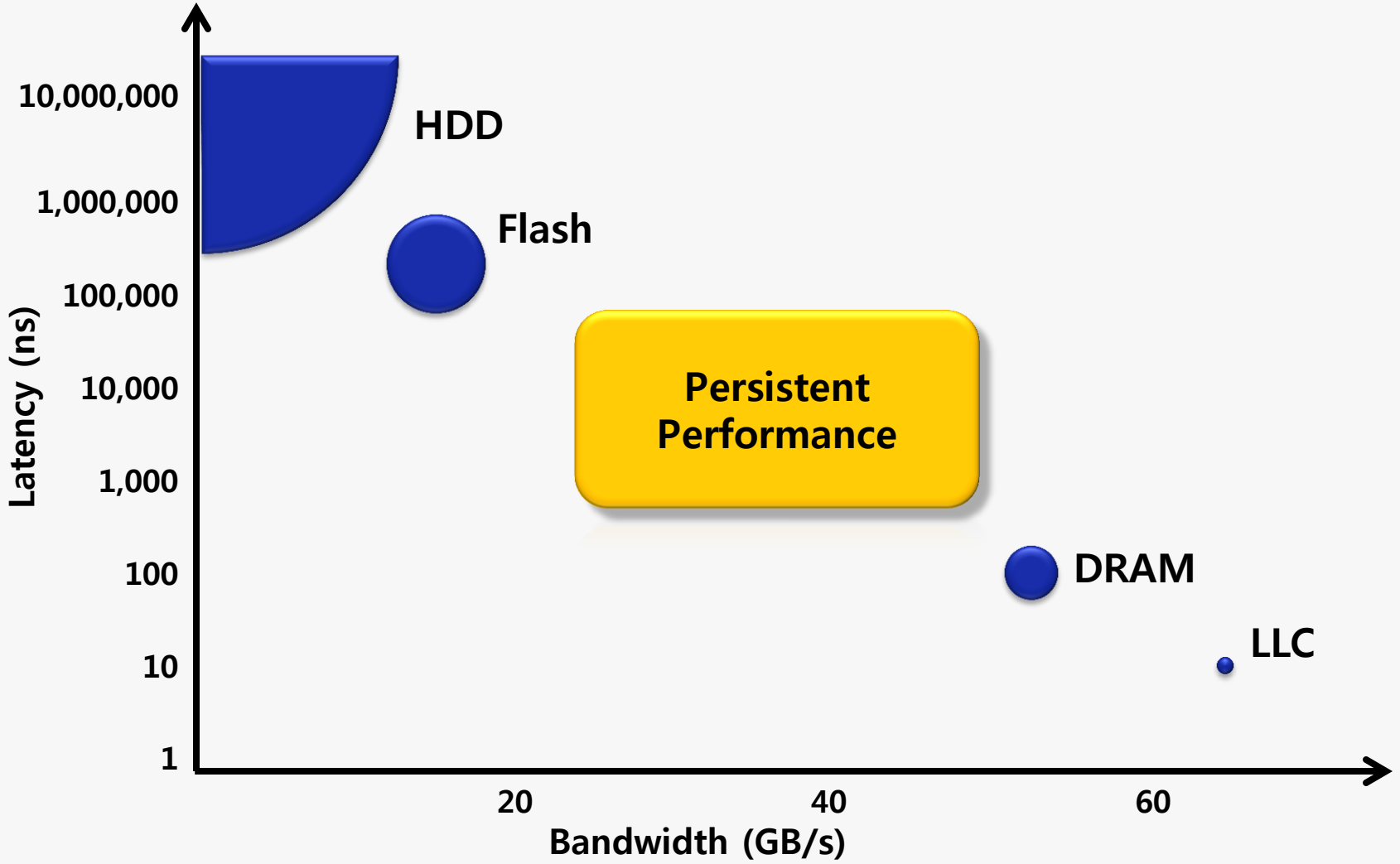
Acknowledgement: Krishna Malladi.
Disclaimer: conceptual model only. CPU data on different scale.

Flash Significantly Improves the DRAM-Disk Gap

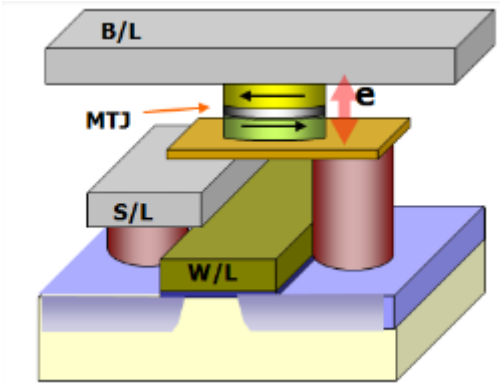
Agenda

- Environment – BW & Capacity growth
- DRAM – BW & Capacity -> Tiering
- Flash –Scales, Becomes Intelligent, Tiers
- **New "Persistent Performance"**

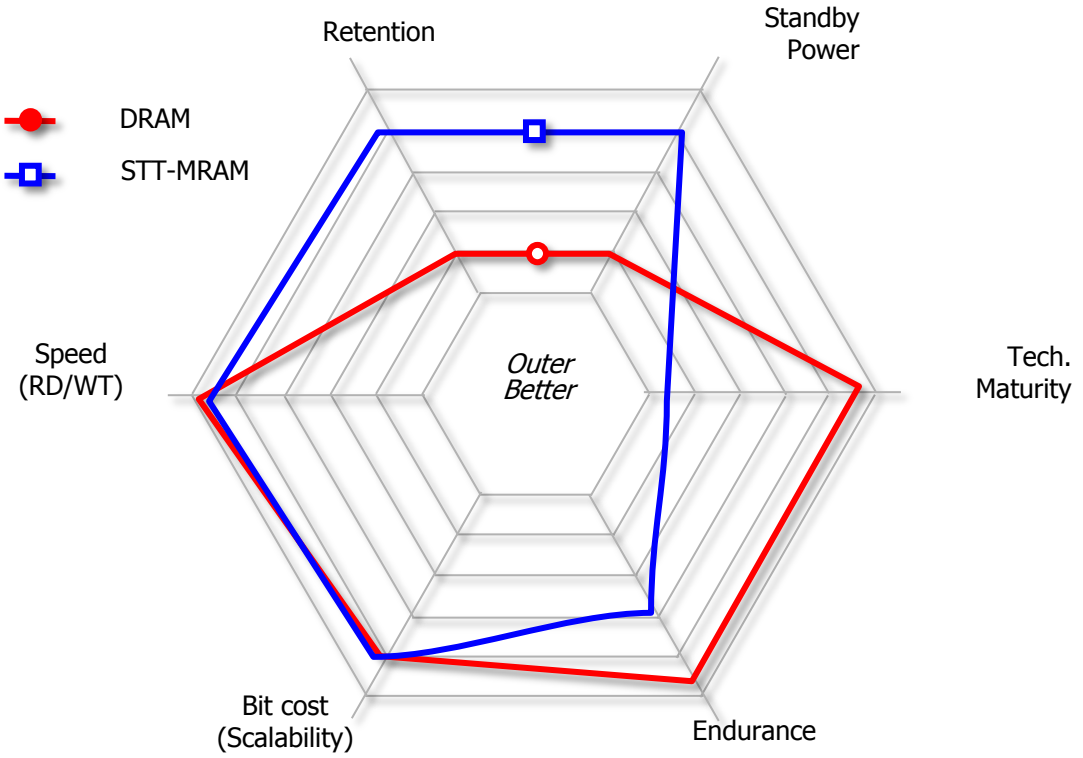
Opportunity for New Technology



STT-MRAM

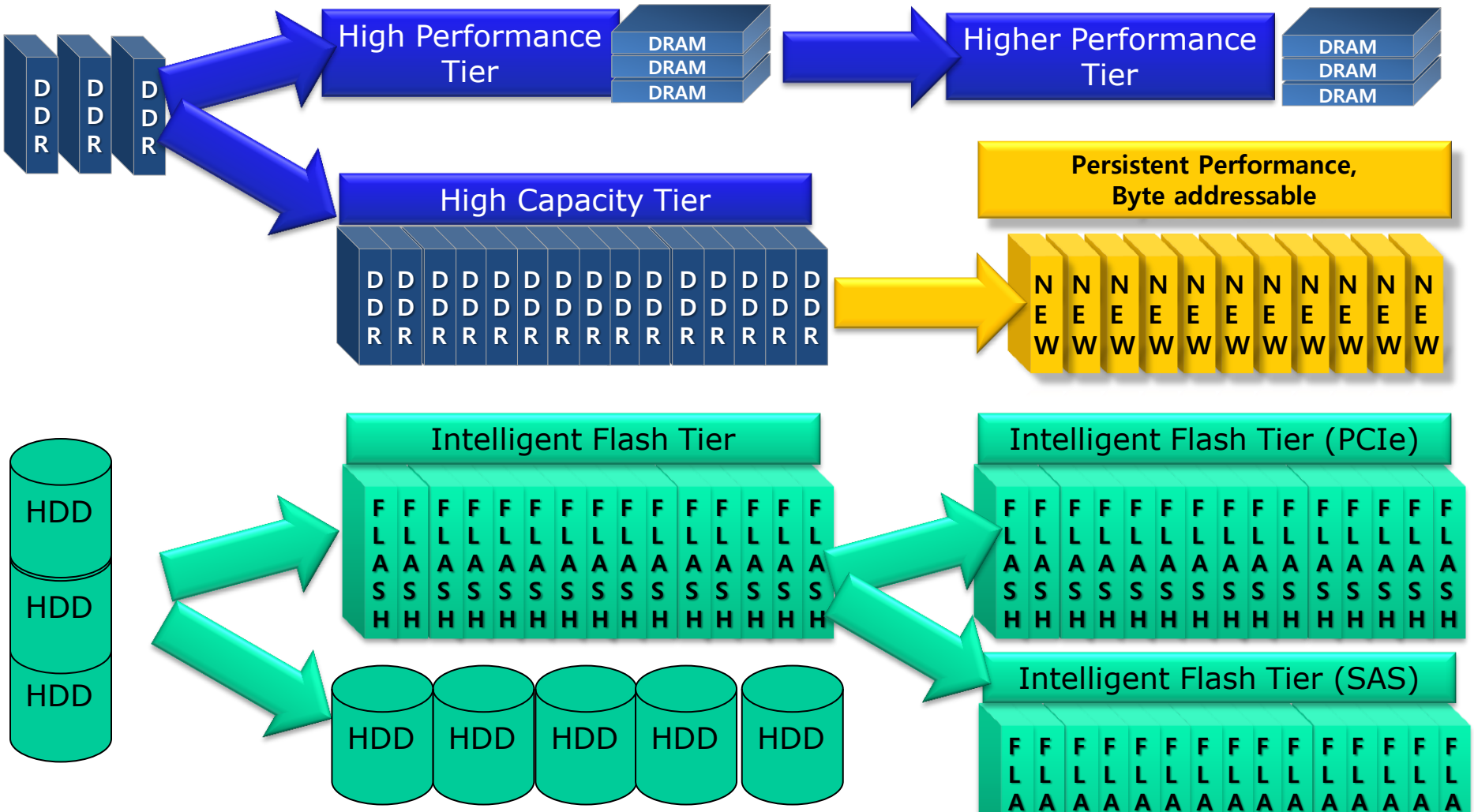


STT-MRAM Cell Structure

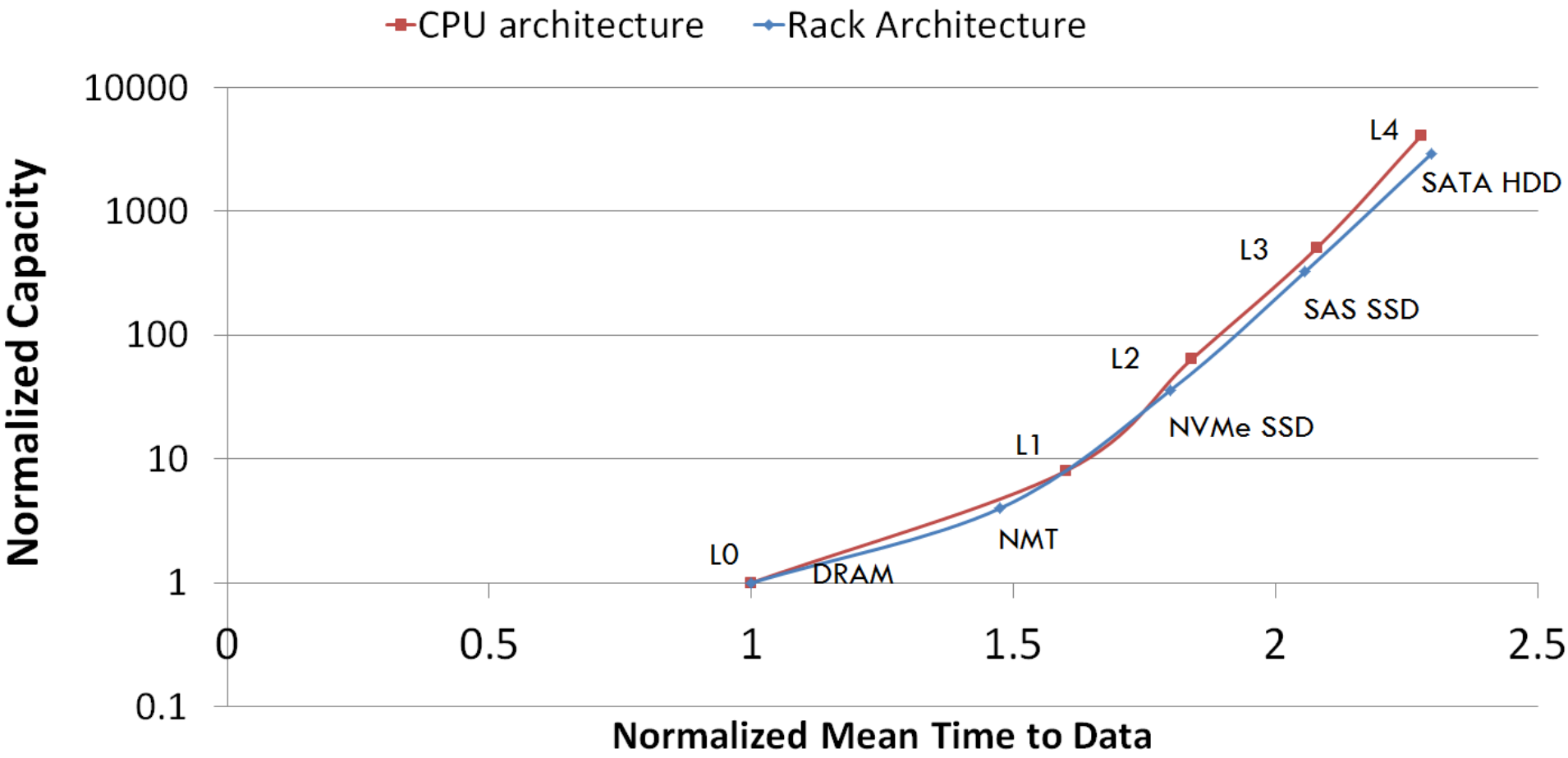


Promising Technology, Not Mature Yet

3rd Step: New possibilities



Future Rack Scaling Vision



Acknowledgement: Krishna Malladi. Disclaimer: conceptual model only.

Ideal Scaling: 1. V-NAND 2. NMT 3. System SW

Thank you!

Questions: Bob.Brennan@Samsung.com