High Performance and Highly Reliable SSD

-Proposal of the Fastest Storage with B4-Flash -

Moriyoshi Nakashima
GENUSION,Inc

http://www.genusion.co.jp/
info@genusion.co.jp

Big Data comes from tons of “Small Data”

- High speed computing of Big Data can be based on ultra speed processing of Small Data.

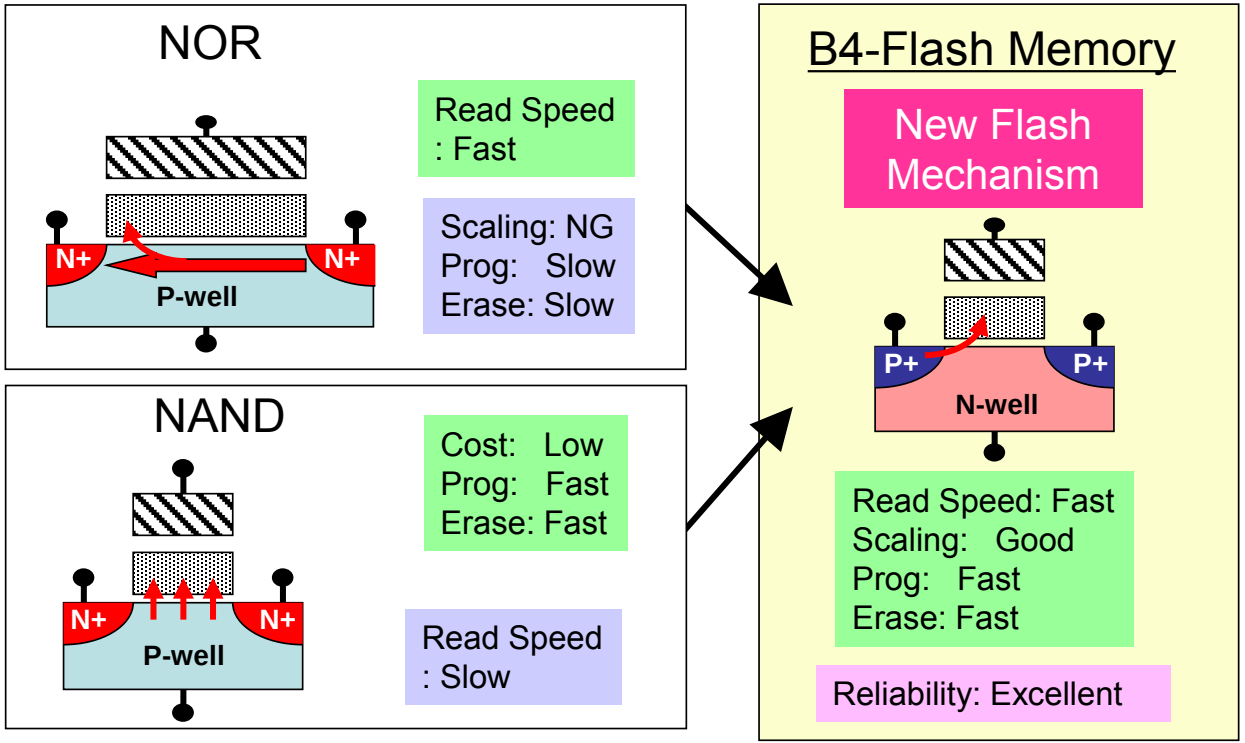
•	How small ?	; 32B, 128B, 256B, 512B or others
•	How Fast ?	; Latency 100 ns Throughput 1GB-6GB/sec
•	How Frequent ?	; IOPS > 1000K
•	Non Volatile ?	; Yes, of course!

- How to correspond to these kinds of requirements ?

B4-Flash Memory can offer excellent features for Small Data processing.

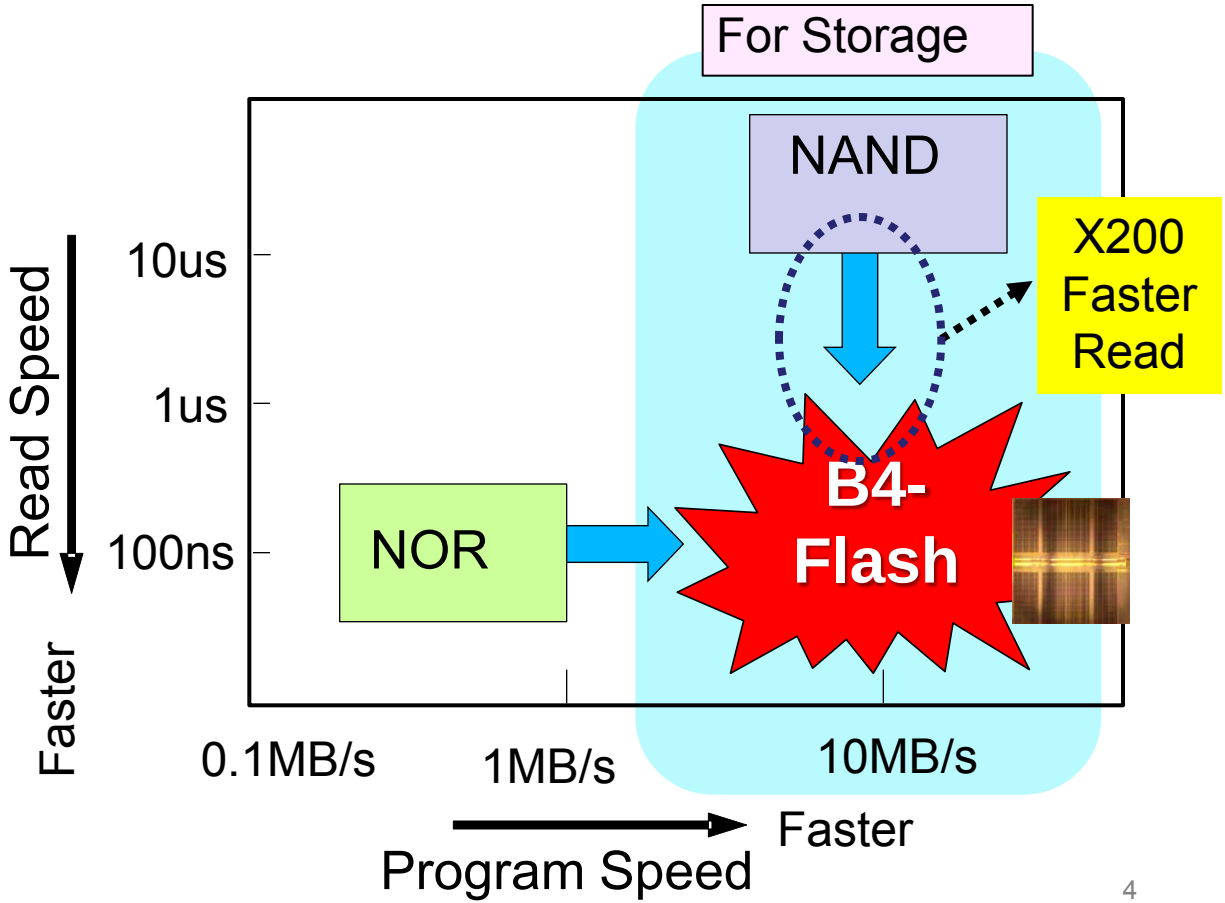
- Random read latency ; 100 ns
- Re-write performance ; equivalent to that of NAND
- Smallest Unit Data Size; 32B / die with on-chip ECC for read page and write page
- No need of complicated F/W for wear leveling thanks to its high reliability

B4-Flash brings New Features for Storage



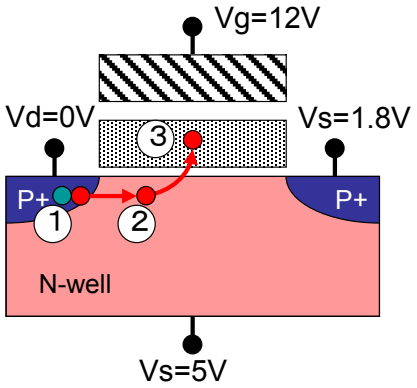
100ns Read Latency can be achieved

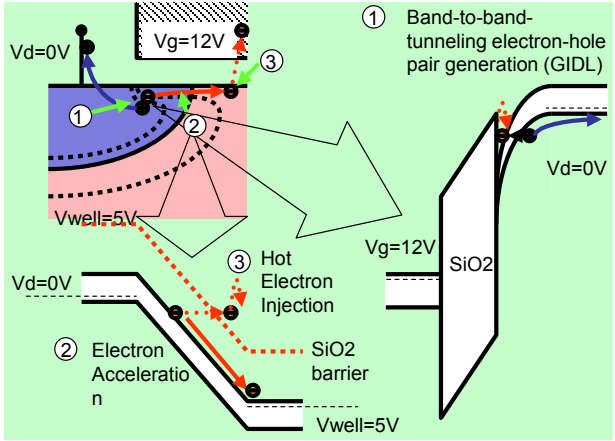
- B4-Flash achieves 100ns NOR read speed and NAND level re-write speed simultaneously.
- Read Latency 100ns can be offered for Storage.



B4-Flash is a Realistically new kind of NVM

Back Bias assisted Band to Band tunneling (B4) - Hot Electron Injection



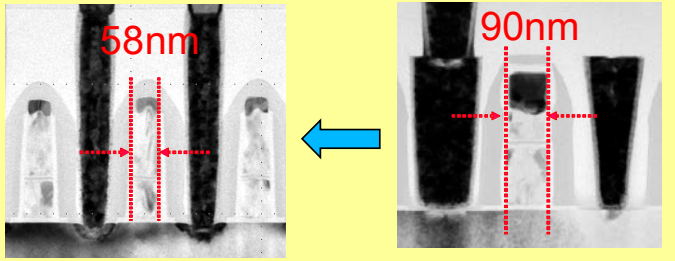


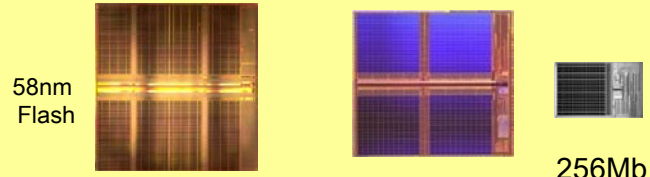
B4-HE injection + Pch MOS transistor memory cell

- Fastest Re-writable and Highly Scalable NOR Flash
- High Endurance to 100K E/W and Excellent Retention of 20years after 100K E/W

B4-Flash

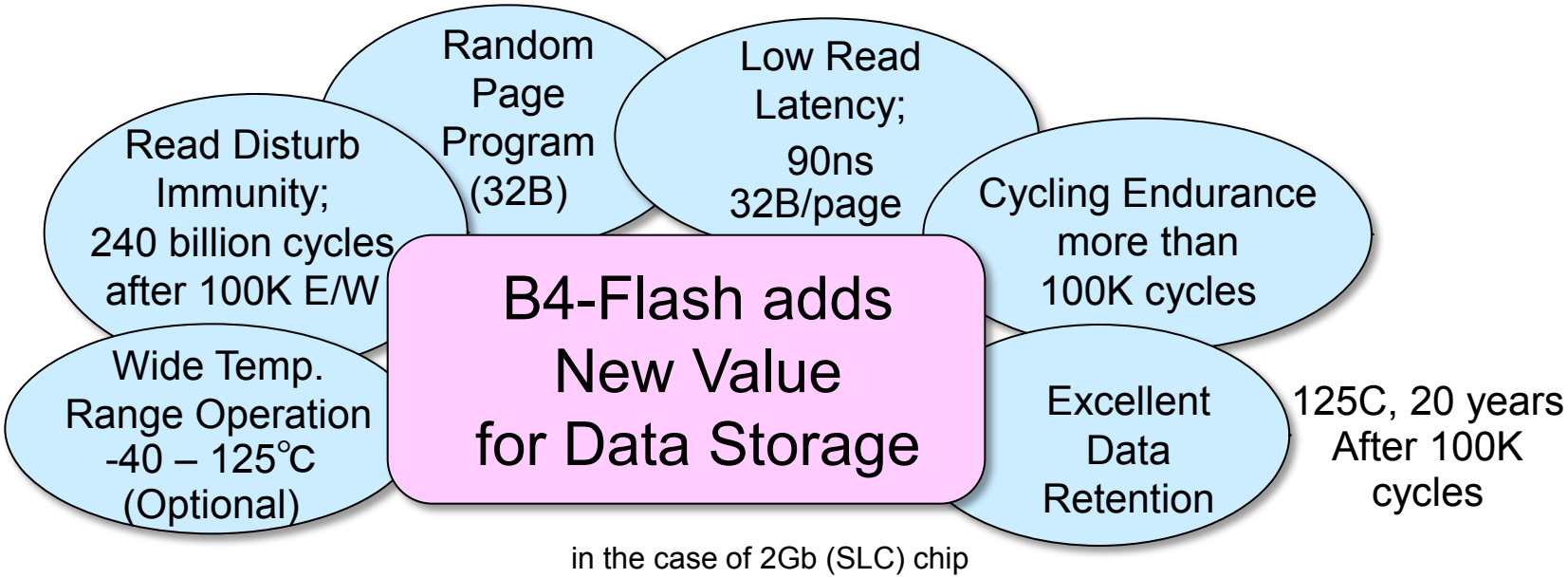
58nm 90nm





B4-Flash achieves the scalable NOR with fast re-write capability and excellent reliability.

B4-Flash can offer New Value for Storage



in the case of 2Gb (SLC) chip

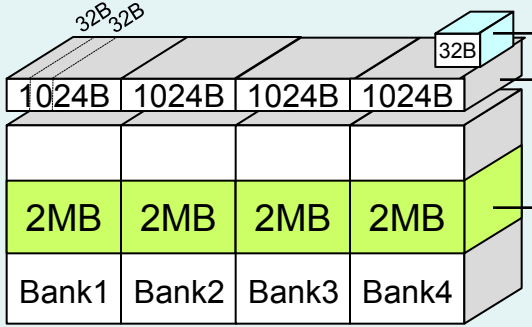
B4-Flash Memory Main Specifications

Items		B4-Flash			
		G28FVW5121S1TE	G29FXW002Kx1	G29FXW002Kx1	G29FXW004Kx1
		SLC, 1die	MLC, 2die	SLC, 1die	MLC, 1die
Memory Density		512Mbit	2Gbit	2Gbit	4Gbit
Process Node		90nm	90nm	58nm	58nm
I/O		x8/x16 parallel	x8/x16 parallel	x16 parallel	x16 parallel
Power Supply Voltage	Core (VCC)	1.7V~2.0V	1.7V~2.0V	2.7V~3.6V	2.7V~3.6V
	I/O (VCCQ)	1.7V~2.0V, 2.7V~3.6V	1.7V~2.0V, 2.7V~3.6V	1.65V~Vcc	1.65V~Vcc
Operation Temperature		-40℃~85℃	-40℃~85℃	-40℃~85℃	-40℃~85℃
Page Size	Page(program)	16W to 512W *	16W to 512W *	16W to 2KW*	16W to 4KW*
	Block(Erase)	512KW, 2MW**	1MW, 4MW**	1MW, 4MW**	2MW, 8MW**
Read	Random Access	115ns	200ns	90ns	110ns
	Page Read	30ns	30ns	15ns	15ns
	Page Size	16W	16W	16W	16W
	16Word Read	565ns	650ns	315ns	335ns
		57MB/s	49MB/s	102MB/s	96MB/s
	Read Current(typ.)	12mA	14mA	10mA	10mA
Block Erase Time(typ.)		100ms	100ms	130ms	130ms
Page Program Time(typ.)		150us	1000us	210us	1350us
Chip Erase Time(typ.)		2 sec**	2sec	5sec**	5sec**
Chip Program Time(typ.)		10 sec**	131us**	15sec**	90sec**
Chip Erase Speed (typ.)		40MB/s	160MB/s	62MB/s	123MB/s
Chip Program Speed (typ.)		7MB/s	2MB/s	19MB/s	6MB/s
Standby	Standby current(typ.)	370uA	1000uA	100uA	100uA
P/E Operation	Operation Current(typ.)	35mA	35mA	35mA	35mA
Reliability	Data Retention	20year (Min.) @125C after 100K E/W	20year (Min.) @85C after 1K E/W	20year (Min.) @125C after 100K E/W	20year (Min.) @85C after 1K E/W
	P/E cycles	100K(Min.)	10K (Min.)	100K (Min.)	100K (Min.)
	Byte Continuous Read	24 billion after 100K E/W	2.8 billion after 10K E/W	240 billion after 100K E/W	28 billion after 10K E/W

58nm 2Gb/4Gb B4-Flash Memory Architecture

The current architecture of B4-Flash is based on NOR but it can be easily applied for the Storage

58nm/2Gb SLC



1 Page（Read） = 32Byte with on-chip ECC

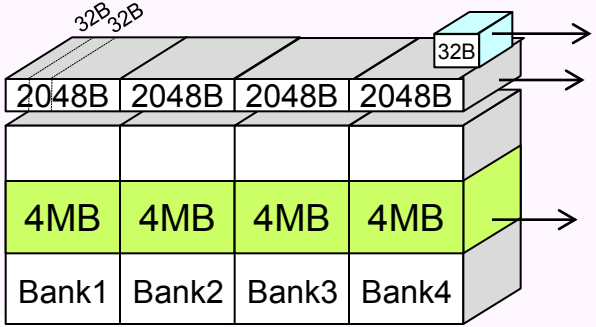
1 Page (Program) = 32B, 64B, 128B, 160B, 192B, 224B, 256B, ---- , 1024Byte, ---- , 4096Byte adjustable to same page size as read

1 Block (Erase) = 1024Bx2048page = 2MByte

1 Bank = 2MBx32Blocks = 64MByte

1 Chip = 64MBx4Banks = 256MByte

58nm/4Gb MLC



1 Page（Read） = 32Byte with on-chip ECC

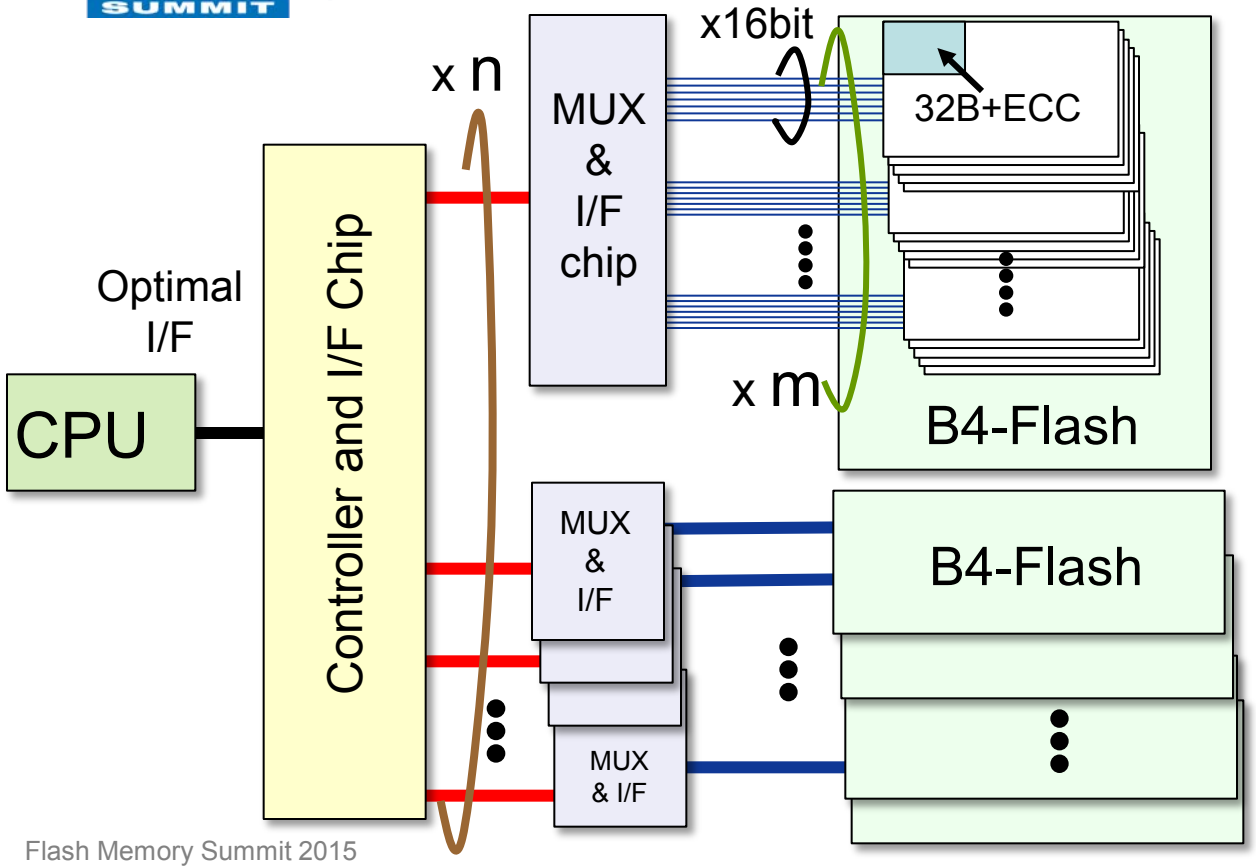
1 Page (Program) = 32B, 64B, 128B, 160B, 192B, 224B, 256B, ---- , 2048Byte, ---- , 8192Byte adjustable to same page size as read

1 Block (Erase) = 2048Bx2048page = 4MByte

1 Bank = 4MBx16Blocks = 64MByte

1 Chip = 64MBx4Banks = 512MByte

Simple Storage Model with B4-Flash



Unit Data is 32B

- ✓ x16
- ✓ 90ns read latency
- ✓ 315ns for 32B read
- ✓ 1bit on-chip ECC
- ✓ Arbitral 32B write

No need of complicated F/W for wear leveling

- ✓ 100K E/W capability
- ✓ No Read Disturb in 240 billion read cycles after 100K E/W
- ✓ Excellent retention
- ✓ internal ECC

Case Study of Simple Storage Model -1-

Case 1) n x m =8 chip parallel, System Page Size=256B
Read; latency=90ns, throughput=812MB/sec, IOPS=3.2M

Case 2) n x m =16 chip parallel, System Page Size=512B
Read; latency=90ns, throughput=1.6GB/sec, IOPS=3.2M

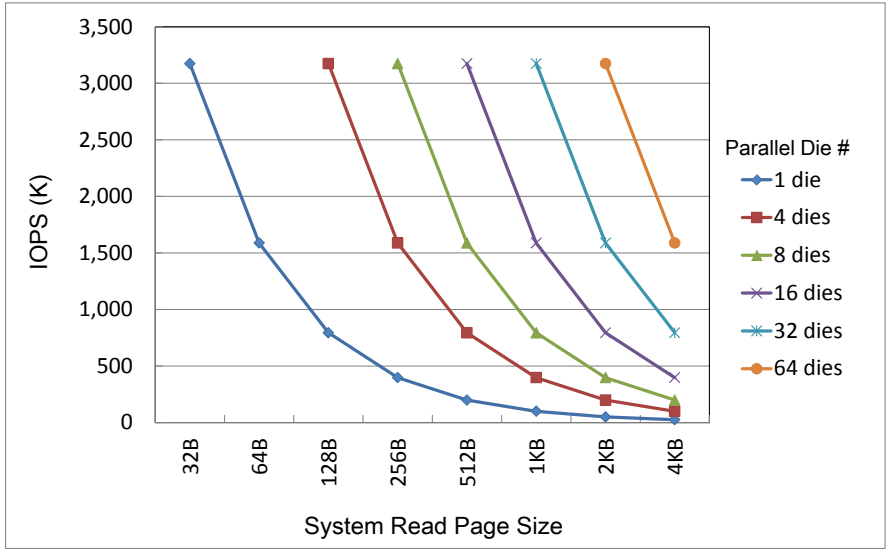
Case 3) n x m =32 chip parallel, System Page Size=1KB
Read; latency=90ns, throughput=3.3GB/sec, IOPS=3.2M

Parallel Read Die # (n x m)		1	4	8	16	32	64
Read Bus (bit)		16	64	128	256	512	1,024
System Read Page Size (Byte)		32	128	256	512	1024	2048
Read Latency (ns)		90	90	90	90	90	90
Read Speed (MB/sec)		102	406	813	1,625	3,251	6,502
Read IOPS (K) for each Data Size	32B	3,175					
	64B	1,587					
	128B	794	3,175				
	256B	397	1,587	3,175			
	512B	198	794	1,587	3,175		
	1KB	99	397	794	1,587	3,175	
	2KB	50	198	397	794	1,587	3,175
	4KB	25	99	198	397	794	1,587

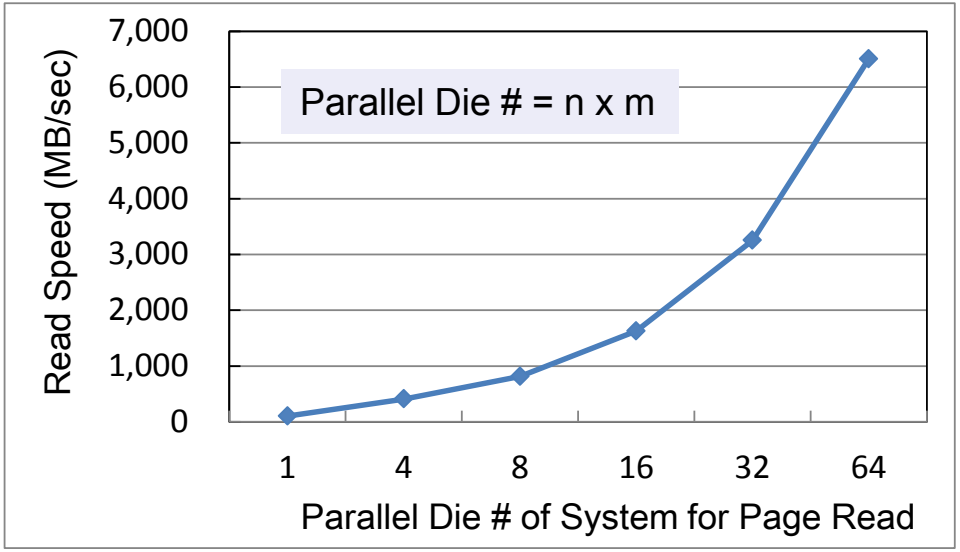
Each calculation is done using the die performance without any overhead of the system as assumption.

Case Study of Simple Storage Model -2-

According to necessary “Small Read Page Size”, Best Die Parallelism can be obtained for IOPS >3M with 90ns of read latency.



Highest Read IOPS for each Data Size



Read Throughput with Parallelism

Each calculation is done using the die performance without any overhead of the system as assumption.

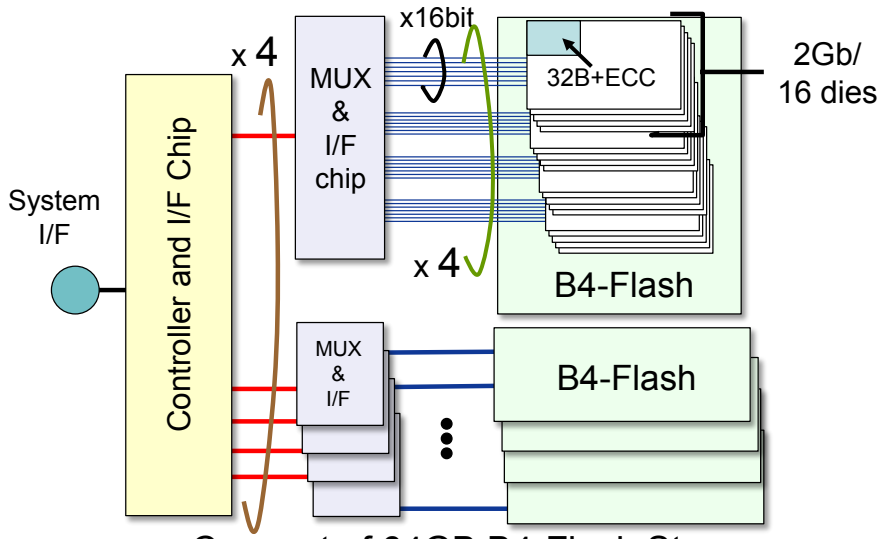
64GB Fast Storage

64GB High Performance Data Storage with B4-Flash

- ✓ System Page Size; 512B
- ✓ 2Gb SLC B4-Flash;
  ((8die MCP x 2chips)x4)
  x4modules
- ✓ Parallel 256bit bus

Expected Performance

✓ Read latency;	90ns
✓ IOPS of 512B;	3.2M
✓ Read throughput;	1.6GB/sec



Concept of 64GB B4-Flash Storage

Write performance must depend on the controller system architecture itself and it should be decided using the same level of B4-Flash SLC write performance as that of SLC NAND.

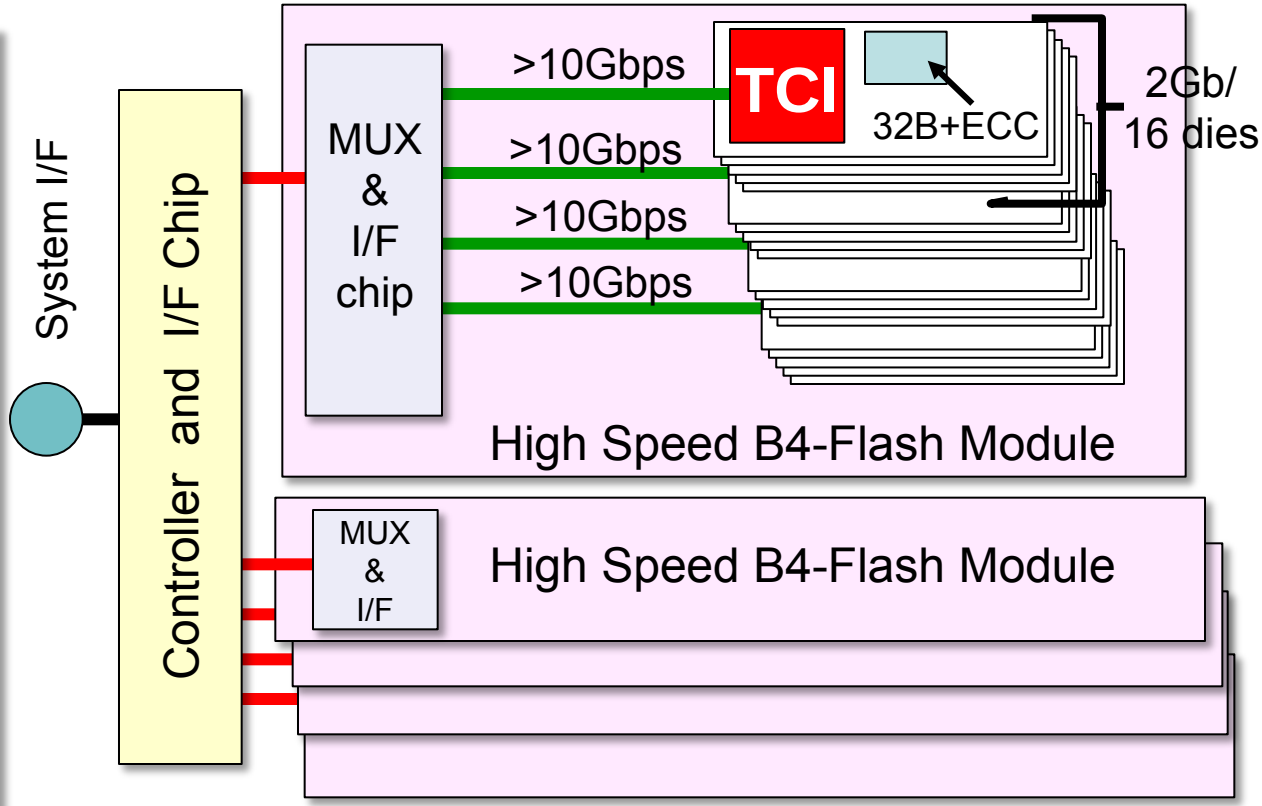
Each calculation is done using the die performance without any overhead of the system as assumption.

TCI (Thru Chip Interconnect) for Higher Speed

TCI (Thru Chip Interconnect) will be implemented on each die and eliminate so much busses and wires around dies.

It can make higher speed interconnection among memory dies in module.

And TCI will bring the benefit for the module design to achieve higher performance of the storage.



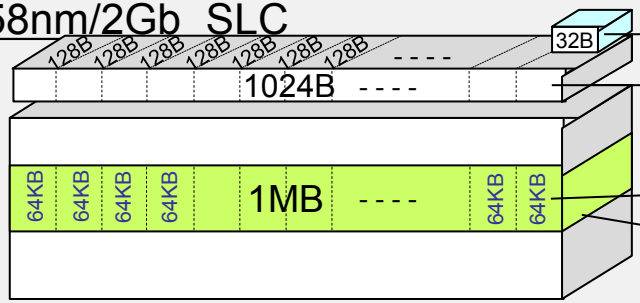
StorageB4 Main Specifications (including small sector types)

Items		NOR base Current B4-Flash products		StorageB4 (Storage ApplicationDedicated) (Under Planning)	
		G29FXW002Kx1	G29FXW004Kx1		
		Type A		Type B	
		SLC, 1die	MLC, 1die	SLC, 1die	MLC, 1die
Memory Density		2Gbit	4Gbit	2Gbit	4Gbit
Process Node		58nm	58nm	58nm	58nm
I/O		x16 parallel	x16 parallel	x16 parallel	x16 parallel
Power Supply Voltage	Core (VCC)	2.7V～3.6V	2.7V～3.6V	2.7V～3.6V	2.7V～3.6V
	I/O (VCCQ)	1.65V～Vcc	1.65V～Vcc	1.65V～Vcc	1.65V～Vcc
Operation Temperature		-40℃～85℃	-40℃～85℃	-40℃～85℃	-40℃～85℃
Page Size	Page(program)	16W to 2KW*	16W to 4KW*	16W to 64W	16W to 128W
	Sector (Erase)	1MW, 4MW**	2MW, 8MW**	32KW	64KW
Read	Random Access	90ns	110ns	180ns	200ns
	Page Read	15ns	15ns	20ns	20ns
	Page Size	16W	16W	16W	16W
	16Word Read	315ns	335ns	480ns	500ns
		102MB/s	96MB/s	67MB/s	64MB/s
Sector Erase	Sector Erase Time(typ.)	130ms	130ms	2ms	2ms
	Block Erase (x16 Sector) Time(typ.)	-	-	6ms	6ms
	Back Ground Erase	No	No	Yes	Yes
Page Program	Page Program Time(typ.)	210us	1350us	150us	1000us
	16 Page Program Time(typ.)	-	-	300us	2000us
Chip Speed	Erase Time(typ.)	5sec (4Bank)	5sec (4Bank)	7.7sec (16sector)	7.7sec(16 sector)
	Program Time(typ.)	15sec (4Bank)	90sec (4Bank)	39sec (16sector)	262sec (16sector)
	Erase Speed (typ.)	62MB/s	123MB/s	333MB/s	333MB/s
	Program Speed (typ.)	19MB/s	6MB/s	7MB/s	2MB/s
Reliability	Data Retention	20year (Min.) @125C after 100K E/W	20year (Min.) @85C after 1K E/W	Under Estimation	Under Estimation
	P/E cycles	100K (Min.)	100K (Min.)	100K (Min.)	100K (Min.)
	Byte Continuous Read	240 billion after 100K E/W	28 billion after 10K E/W	24 billion after 100K E/W	2.8 billion after 10K E/W

“StorageB4” 2Gb/4Gb Architecture (small sector)

“Storage B4” will be developed as Fast Storage Application dedicated Memory with B4-Flash architecture.

58nm/2Gb SLC



1 Page（Read） = 32Byte with on-chip ECC

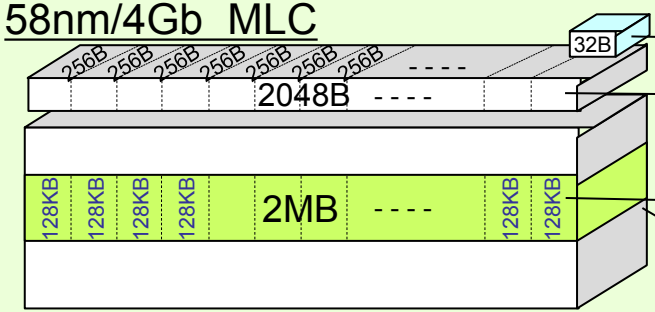
1 Page (Program) = 32B, 64B, 96B , 128B, ---- , 256B, 512B, ---- , 1024Byte, ---- , 2048Byte adjustable to same page size as read

1 Sector (Erase) = 128Bx512page = 64KByte

1 Block (Erase) = 64KBx16sector = 1MByte

1 Chip = 1MBx256blocks = 256MByte

58nm/4Gb MLC



1 Page（Read） = 32Byte with on-chip ECC

1 Page (Program) = 32B, 64B, 96B, 128B, ----, 256B, ---- , 512B, ----, 4096Byte adjustable to same page size as read

1 Sector (Erase) = 256Bx512page = 128KByte

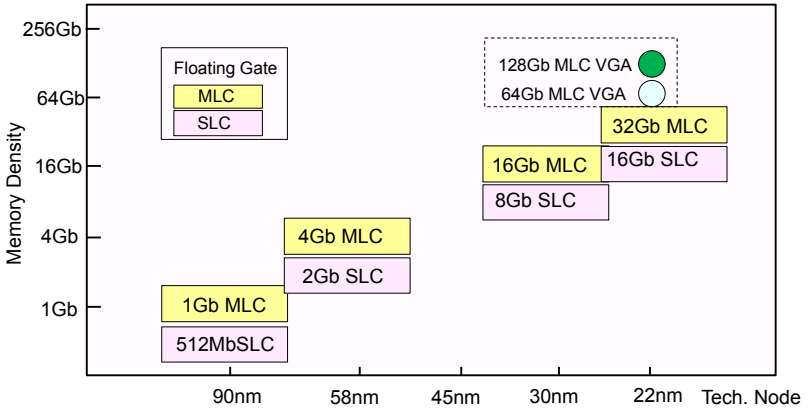
1 Block (Erase) = 128KBx16sector = 2MByte

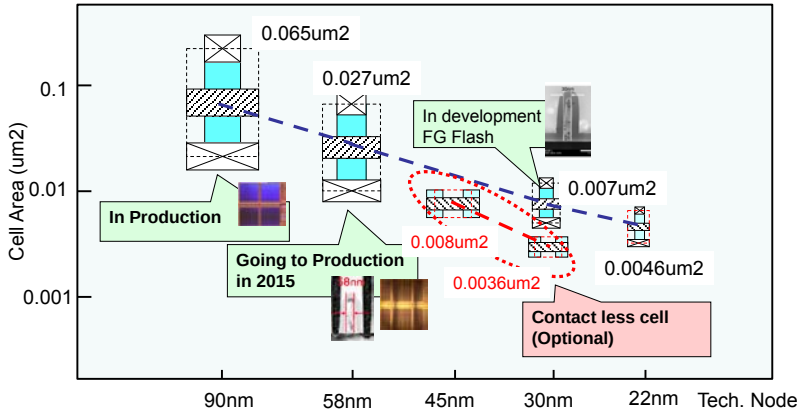
1 Chip = 2MBx256blocks = 512MByte

Higher Density with Low Cost can be achieved

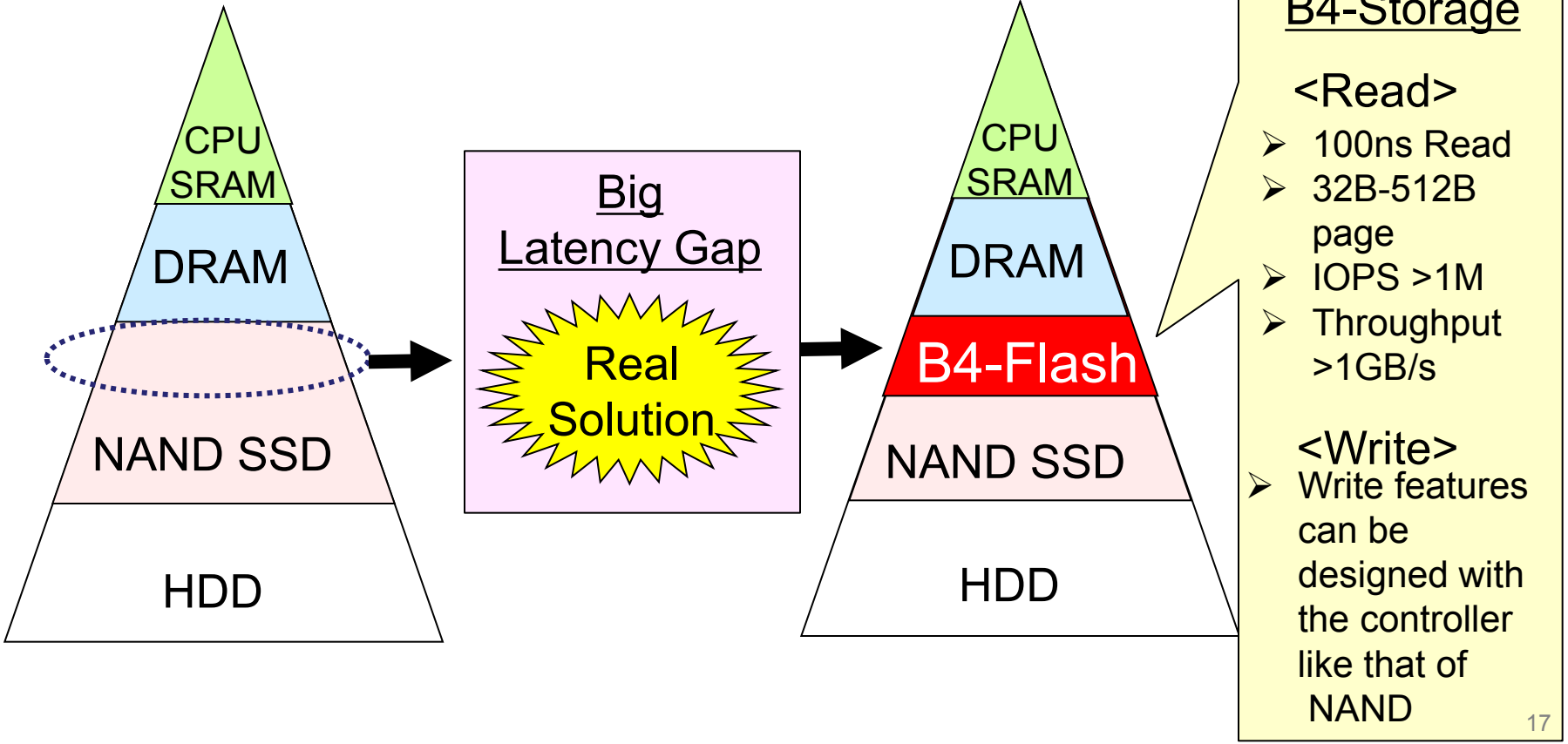
How to achieve Higher Density B4-Flash

- ✓ Scaling Down to 20nm
- ✓ MLC capability
  - MLC B4-Flash has been realized in 90nm/58nm process
  - 4F2 Cell with MLC
  - Smaller than DRAM (6F2)
  - MLC may require a little precise F/W of Wear Leveling for write
  - Read performance must be almost same as that of SLC
- ✓ Virtual Grand Array (VGA)will be the challenge for 2F2 MLC same as 2D NAND in future

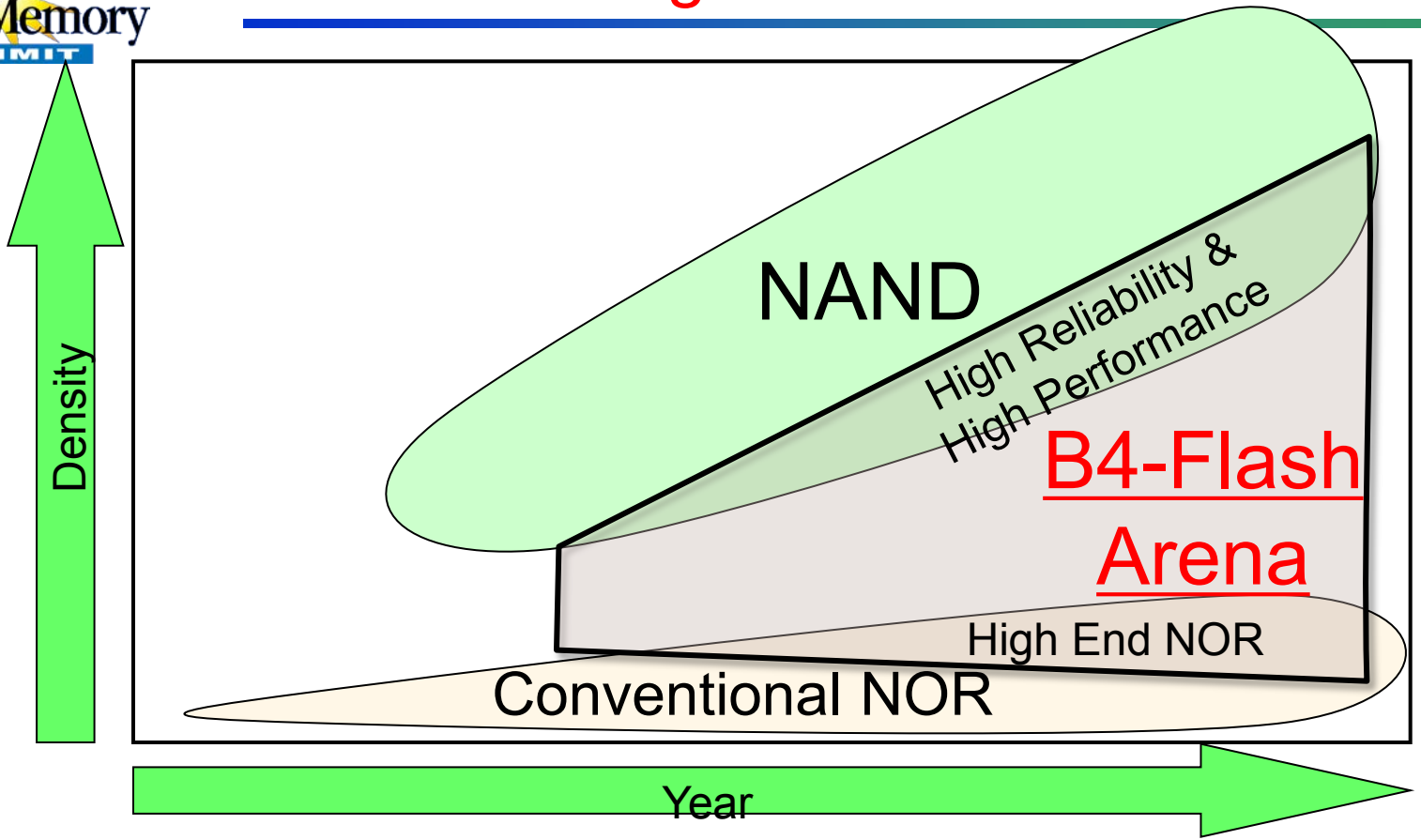




B4-Flash for Storage Class Memory Application

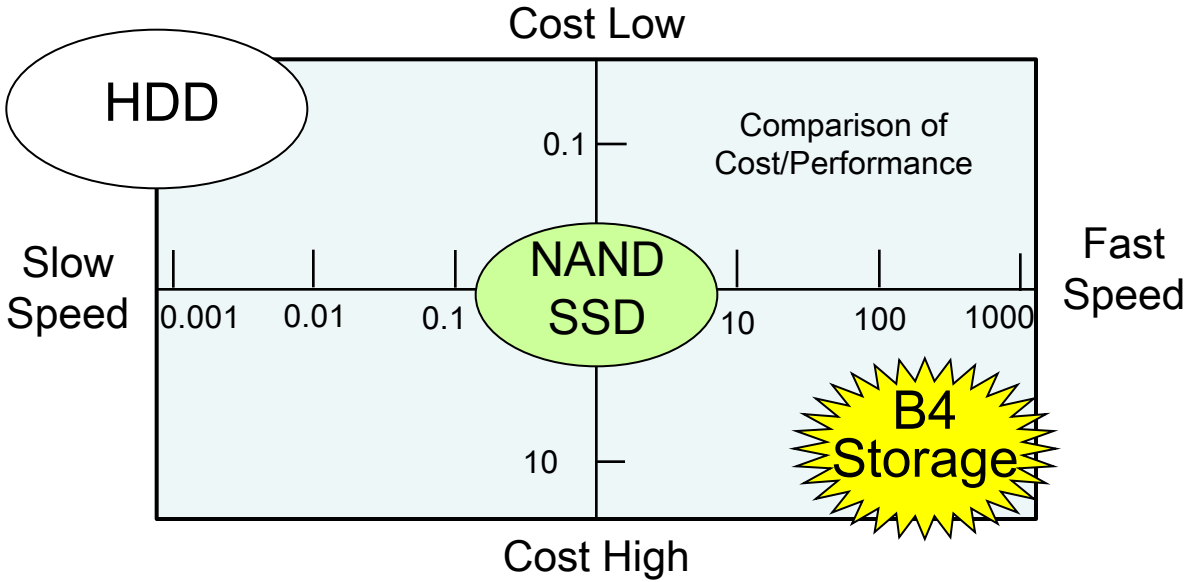


Creating New Flash Arena



x10 of cost performance

- B4 Storage can offer x100 Speed with more cost.
- We hope there exist the applications requiring x10 cost performance



If you have any kind of interest in B4 Storage, please let us know as;

Flash Memory Summit Booth; **#536**
E-mail; fms2015@genusion.co.jp
Web.; http://www.genusion.co.jp/index_e.html

Thank you for your attention!

GENUSION exhibiting at FMS 2015 is partially supported by Japan External Trade Organization (JETRO).