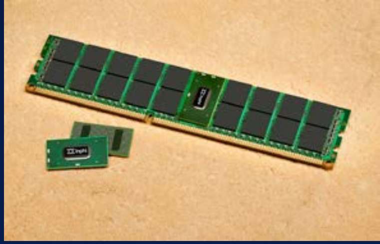
Server Design Summit

Introducing LRDIMM in Servers and Workstations

Sameer Kuppahalli, Product Manager, Inphi

Theme

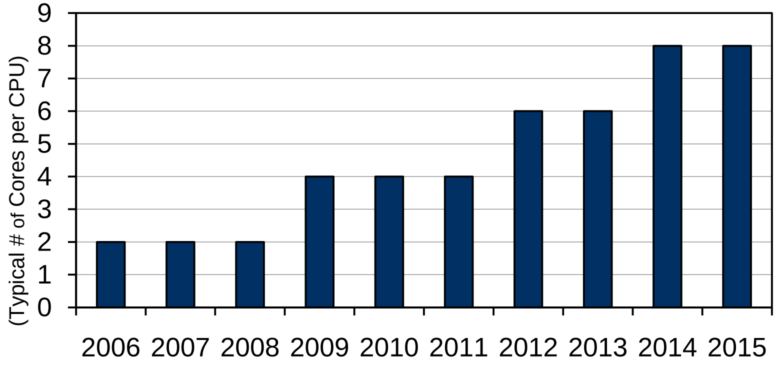


- **Challenges with Memory Bandwidth Scaling**
- How LRDIMM Addresses this Challenge
- Under the hood
- LRDIMM Value Proposition
- Key Enabler: The Memory Buffer (MB)

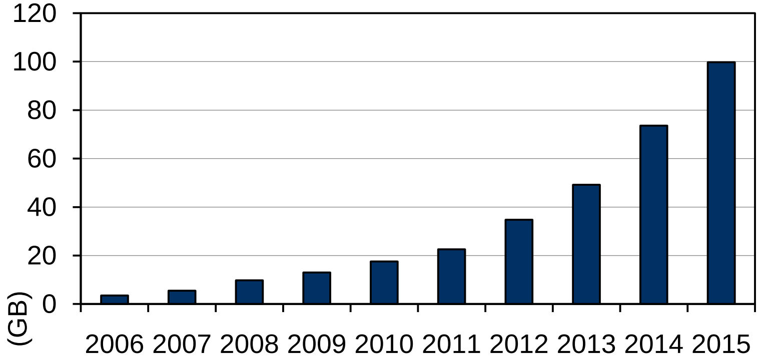
Memory Bandwidth & Capacity Requirements Growing in Servers and Workstations

- As the number of cores increase, more threads contend for the memory capacity and bandwidth
- Applications today, such as Virtualization that utilize multiple cores, require more Memory capacity as well as Memory Bandwidth
- Without increase in the memory subsystem performance, cores will get starved for data, thereby impacting the overall system performance

Source: IDC, 2011



Source: IDC, 2011



Source: IDC, 2011

Increasing the Memory Subsystem Performance

Cost

Latency

Power Consumption

Form-factor

Constraints

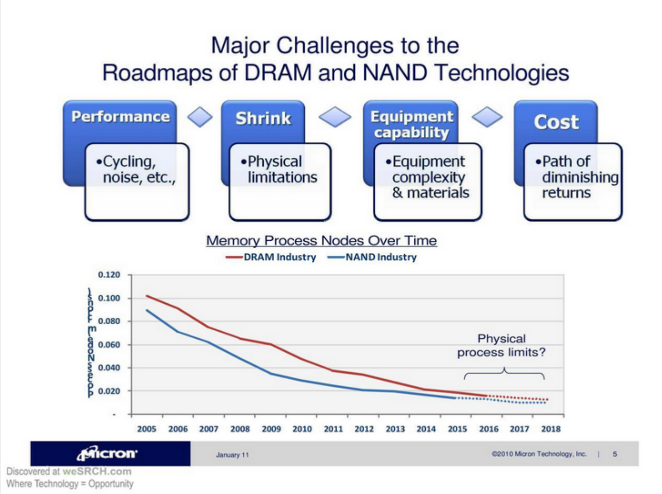
Channel Speed

Capacity

Drivers

Options for improving Memory Module Capacity and Speed

Options	Power	Latency	Cost	Speed
Using cutting edge DRAM densities with Existing Module technology	++	++	--	+
“New” Module technology	+	+	++	++



Major Challenges to the Roadmaps of DRAM and NAND Technologies

- **Performance**
  - Cycling, noise, etc.,
- **Shrink**
  - Physical limitations
- **Equipment capability**
  - Equipment complexity & materials
- **Cost**
  - Path of diminishing returns

Memory Process Nodes Over Time

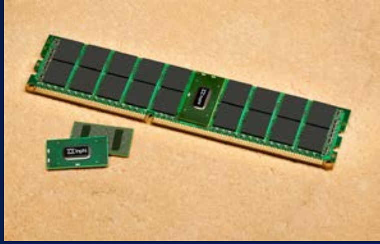
January 11

©2010 Micron Technology, Inc. | 5

Discovered at weSRCH.com
Where Technology = Opportunity

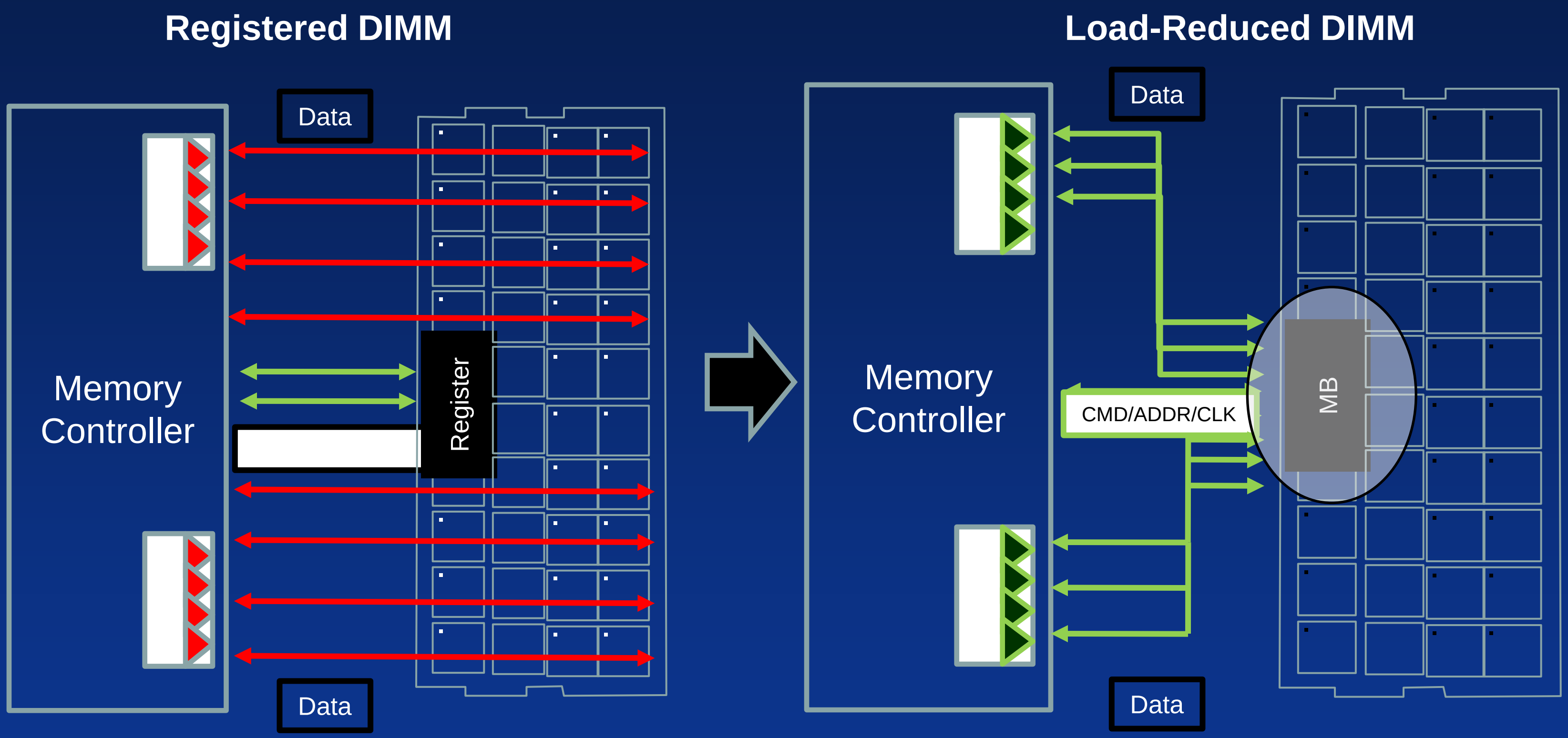
- ❑ **“New” Module technology is required for Cost Effective Solution**
- ❑ **Need to solve speed and Signal Integrity challenges**

Theme



- Challenges with Memory Bandwidth Scaling
- **How LRDIMM Addresses this Challenge**
- Under the hood
- LRDIMM Value Proposition
- Key Enabler: The Memory Buffer (MB)

Buffer at the "heart" of New Technology: Load-Reduced DIMM

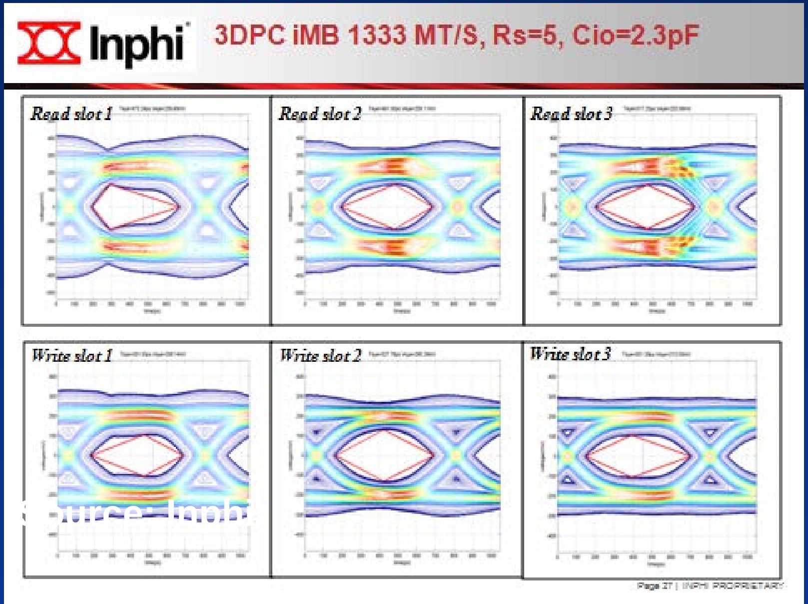


- ☐ **Memory Buffer LRDIMM solves the Channel Speed and SI challenges by "reinforcing" <u>all</u> signals for command, address, control and data on the Memory-controller interface**

Signal Integrity Improvement

- Improved SI on the Data interface with Memory controller

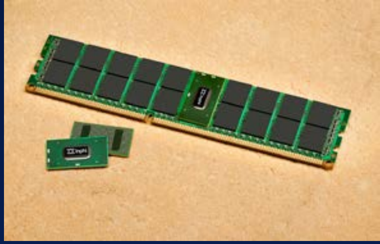
- This allows for higher channel speed and/or more DRAM ranks on the DIMM



Inphi 3DPC iMB 1333 MT/S, Rs=5, Cio=2.3pF

Source: Inphi

Theme

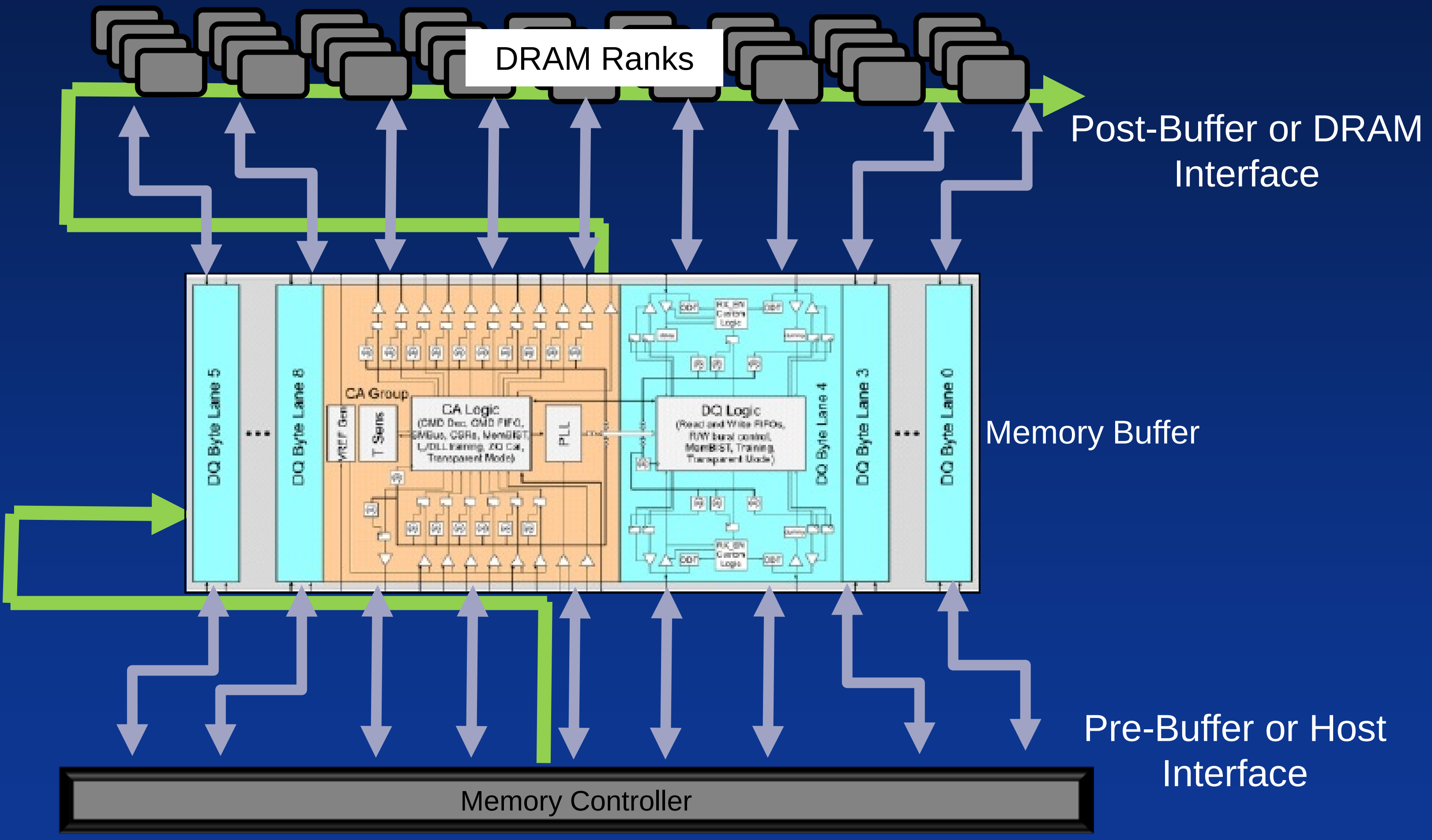


- Challenges with Memory Bandwidth Scaling
- How LRDIMM Addresses this Challenge
- **Under the hood**
- LRDIMM Value Proposition
- Key Enabler: The Memory Buffer (MB)

Memory Buffer Features & Functions

Feature	Benefit	End-user Value
Rank Multiplication, refresh support	☐ Higher DIMM/channel population	☐ Increased Capacity
Read/Write On-die-termination (ODT)	☐ Improved Signal integrity on the Host interface for	☐ Better performance
Improved receiver sensitivity		
Extended “in-band” Control Word (CW) accesses	☐ Efficient access to the control words helps reduced boot and reset times	☐ Better performance
Error Injection	☐ In system Reliability testing	☐ Improved RAS
Temperature sensor on buffer	☐ Efficient thermal management of the DIMM	
MemBIST & Transparent mode support	☐ Reduced Test time	☐ Lower DIMM cost
ODT value optimization	☐ 15% reduction in active buffer power ☐ 20% reduction in idle buffer power	☐ Optimal power consumption

Memory Buffer Block Diagram



DRAM Ranks

Post-Buffer or DRAM Interface

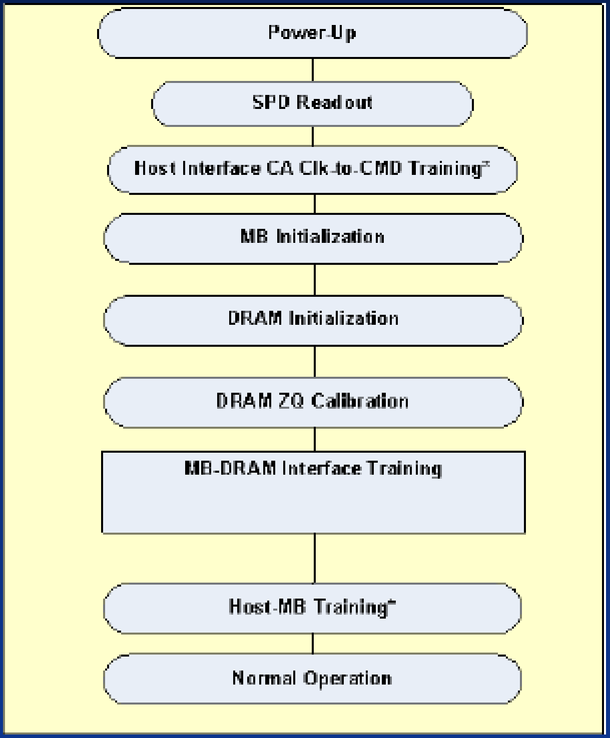
Memory Buffer

Pre-Buffer or Host Interface

Memory Controller

Command, Address, Control, Clock	→
Data	↔

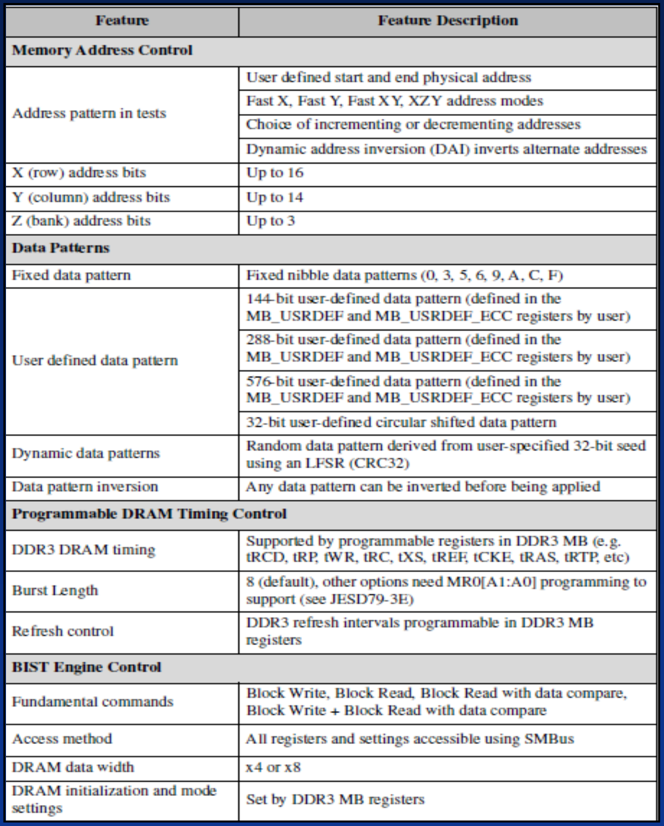
Initialization Overview



1. Power up requirements same as 882
2. Host initializes MB Control Words
3. Host initializes DRAM MRS registers
4. Host issues ZQ command to DRAM
5. Host issues Control Word write to start MB-DRAM interface training
6. Host does interface training
7. Normal Operation

MemBIST Engine

- The DDR3 Memory Buffer (MB) supports Memory Built-In Self Test (MemBIST)
- At DIMM manufacturing, MEMBIST offers a method to detect efficiently
- This testing may be initiated either in or out of band, making MEMBIST compatible with



Feature	Feature Description
Memory Address Control	
Address pattern in tests	User defined start and end physical address
	Fast X, Fast Y, Fast XY, XZY address modes
	Choice of incrementing or decrementing addresses
	Dynamic address inversion (DAI) inverts alternate addresses
X (row) address bits	Up to 16
Y (column) address bits	Up to 14
Z (bank) address bits	Up to 3
Data Patterns	
Fixed data pattern	Fixed nibble data patterns (0, 3, 5, 6, 9, A, C, F)
User defined data pattern	144-bit user-defined data pattern (defined in the MB_USRDEF and MB_USRDEF_ECC registers by user)
	288-bit user-defined data pattern (defined in the MB_USRDEF and MB_USRDEF_ECC registers by user)
	576-bit user-defined data pattern (defined in the MB_USRDEF and MB_USRDEF_ECC registers by user)
	32-bit user-defined circular shifted data pattern
Dynamic data patterns	Random data pattern derived from user-specified 32-bit seed using an LFSR (CRC32)
Data pattern inversion	Any data pattern can be inverted before being applied
Programmable DRAM Timing Control	
DDR3 DRAM timing	Supported by programmable registers in DDR3 MB (e.g. tRCD, tRP, tWR, tRC, tXS, tREF, tCKE, tRAS, tRTP, etc)
Burst Length	8 (default), other options need MR0[A1:A0] programming to support (see JESD79-3E)
Refresh control	DDR3 refresh intervals programmable in DDR3 MB registers
BIST Engine Control	
Fundamental commands	Block Write, Block Read, Block Read with data compare, Block Write + Block Read with data compare
Access method	All registers and settings accessible using SMBus
DRAM data width	x4 or x8
DRAM initialization and mode settings	Set by DDR3 MB registers

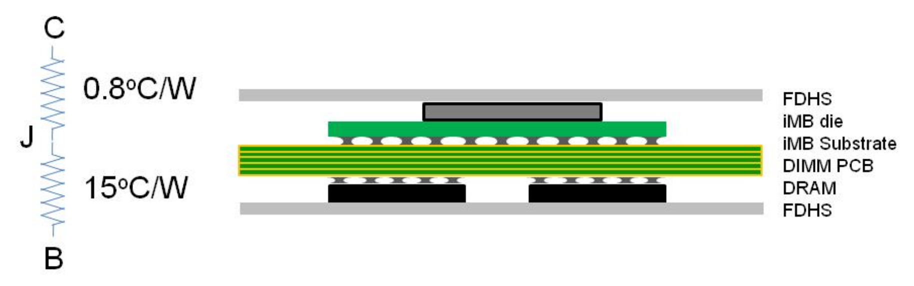
Low Power Modes

ACPI mode	ACPI description	MB / SSTE32882 mode	MB / SSTE32882 description
S1	Low wake latency sleep mode. Hardware maintains all system context.	CKE powerdown	MB / SSTE32882 puts DRAM into self-refresh. Non-essential I/O are allowed to go high-Z to save power. Clocks continue to run in order to maintain status of timed signal paths.
S3	Low wake latency sleep mode where all context is lost except system memory.	CK stopped powerdown	MB / SSTE32882 puts DRAM into self-refresh. Non-essential I/O are allowed to go high-Z to save power. Clocks are stopped to save power. Data paths are idled, and analog timing circuits are powered down.

❑The MB power-managed modes are identical to the power modes that exist in the SSTE32882

Thermal FC-BGA for Memory Buffer

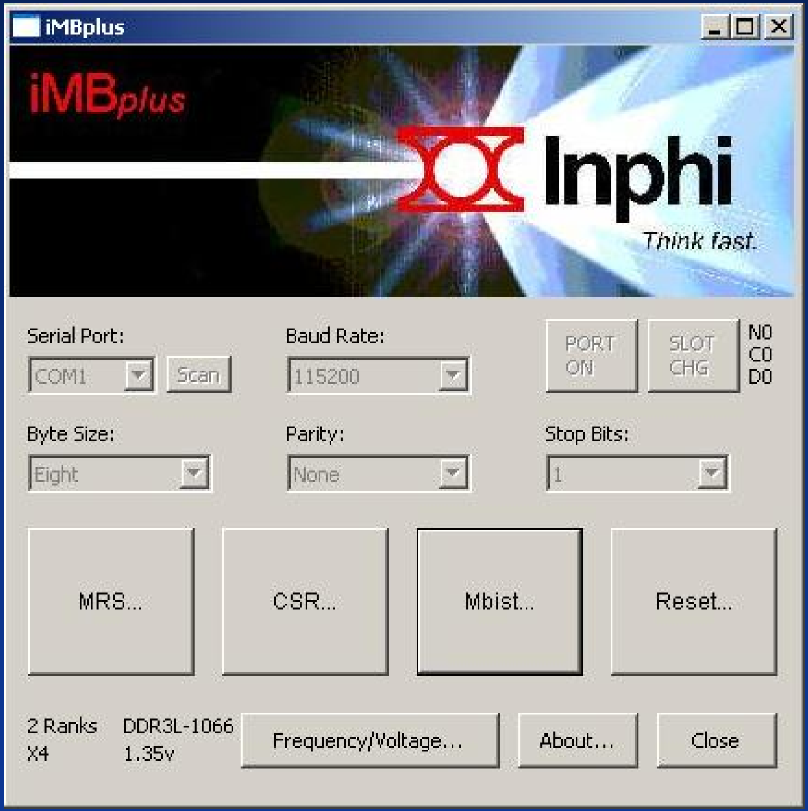
- Memory buffer package design is Flip-Chip BGA
- Thermal-up path through silicon to heat spreader



❑**FC-BGA is optimal for buffer thermal management**

LRDIMM Configuration and Analysis Tool

- In LRDIMM, there is no visibility for the host to the DRAM interface
- iMB™*plus* developed by Inphi engineering team
- Tool provides visibility to post-buffer registers, data, commands
- Provides frequency and voltage Selection
- Utilizes MemBISTOperation
- CSR Reads and Writes
- MRS Settings



Serial Port: COM1

Baud Rate: 115200

NO
C0
D0

Byte Size: Eight

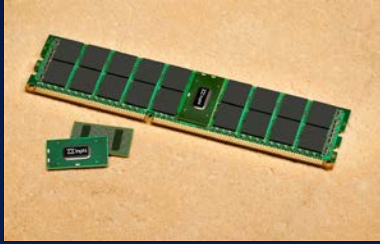
Parity: None

Stop Bits: 1

2 Ranks
X4

DDR3L-1066
1.35v

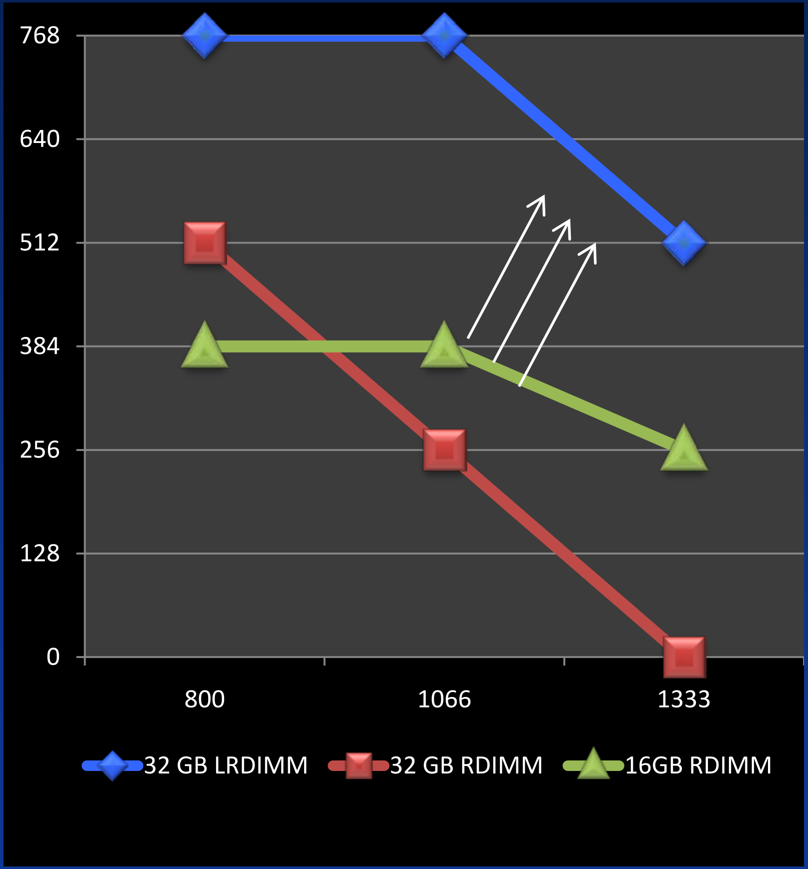
Theme



- Challenges with Memory Bandwidth Scaling
- How LRDIMM Addresses this Challenge
- Under the hood
- LRDIMM Value Proposition
- Key Enabler: The Memory Buffer (MB)

LRDIMM Value Prop

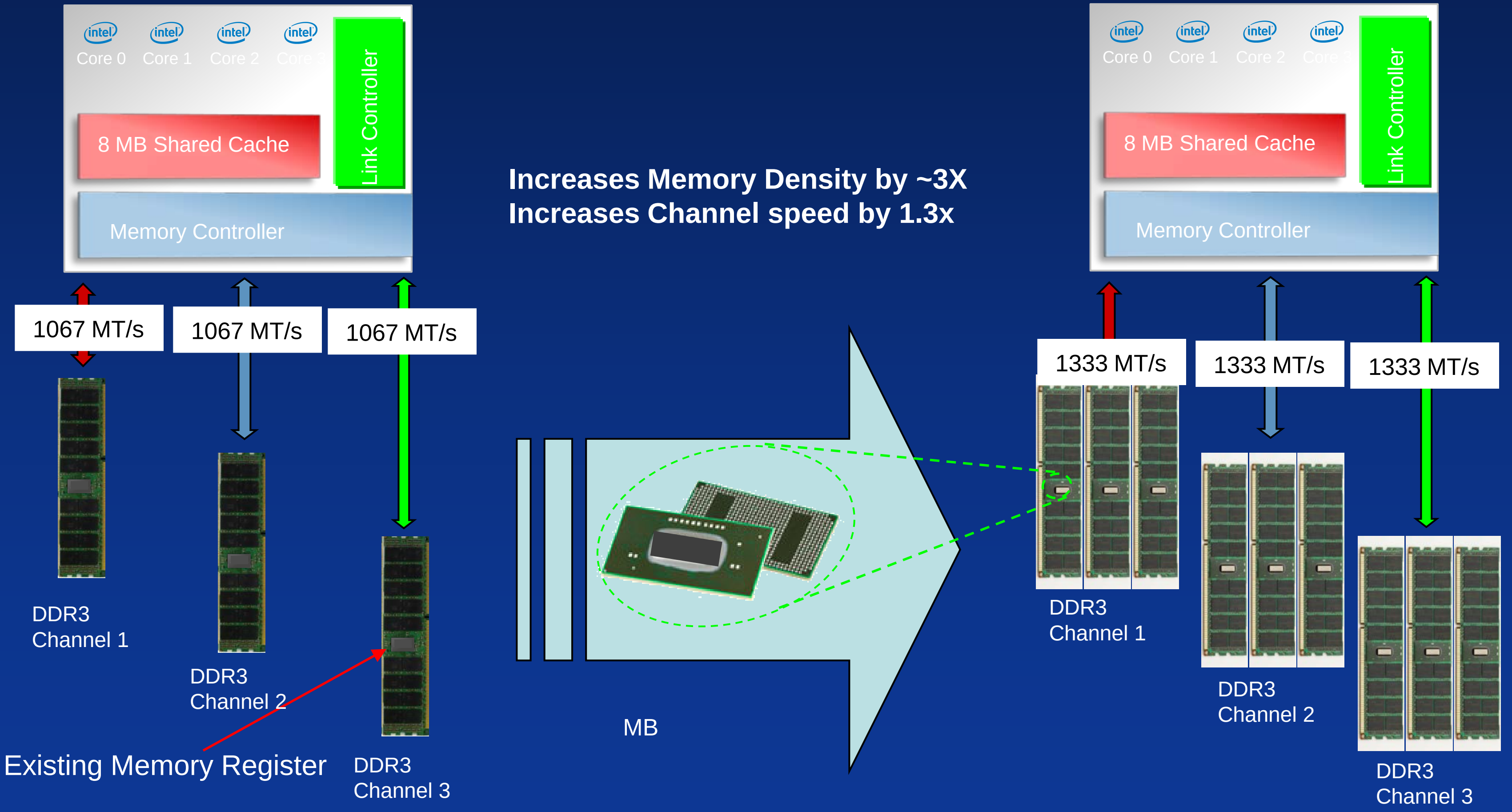
- Applications like Virtualization, in-memory databases and High performance computing are driving higher core count and associated memory
- LRDIMM enables 2x-3x capacity @ 1066 and 1333 MT/s as compared to RDIMM
- This provides significantly higher memory per core for compute and memory intensive applications



MB Significantly Expands Memory Capacity

RDIMM Solutions

MB Powered LRDIMM Solution



CPU 1

intel Core 0
intel Core 1
intel Core 2
intel Core 3

8 MB Shared Cache

Link Controller

Memory Controller

1067 MT/s
1067 MT/s
1067 MT/s

DDR3 Channel 1
DDR3 Channel 2
DDR3 Channel 3

Existing Memory Register

Increases Memory Density by ~3X
Increases Channel speed by 1.3x

MB

CPU 1

intel Core 0
intel Core 1
intel Core 2
intel Core 3

8 MB Shared Cache

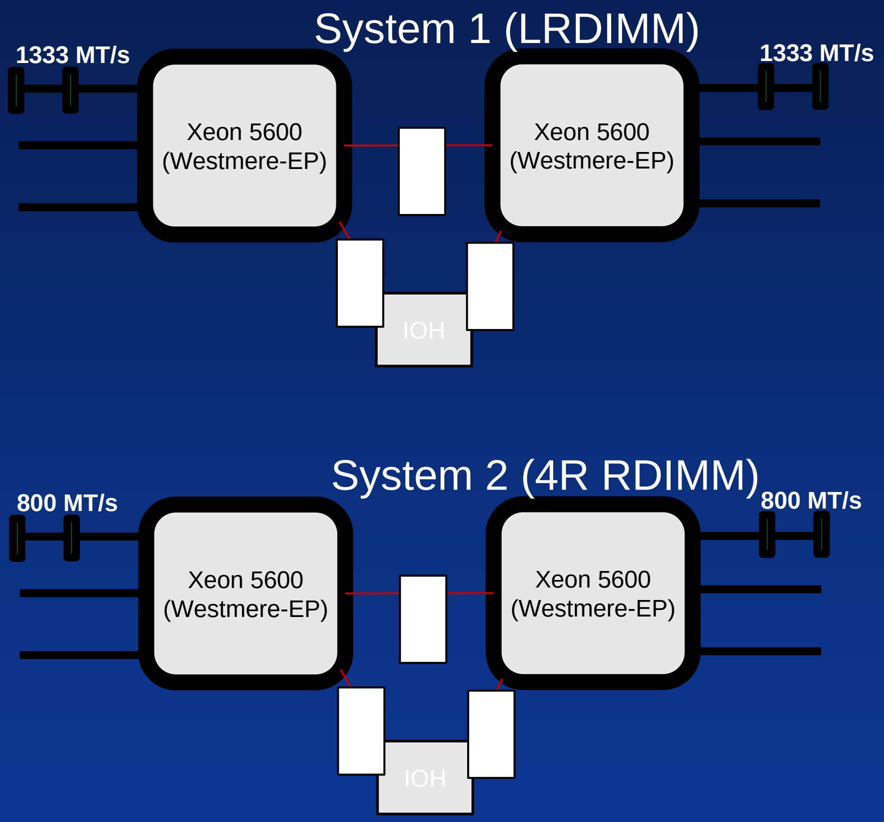
Link Controller

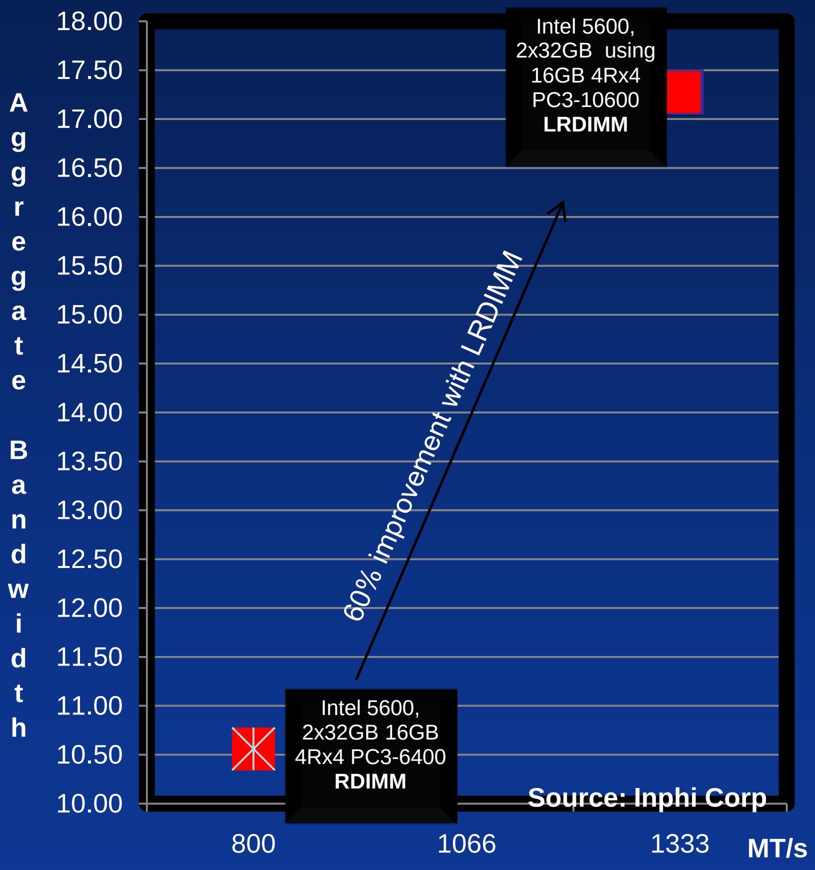
Memory Controller

1333 MT/s
1333 MT/s
1333 MT/s

DDR3 Channel 1
DDR3 Channel 2
DDR3 Channel 3

Performance Scaling of LRDIMM on Intel x5600 Dual Xeon System

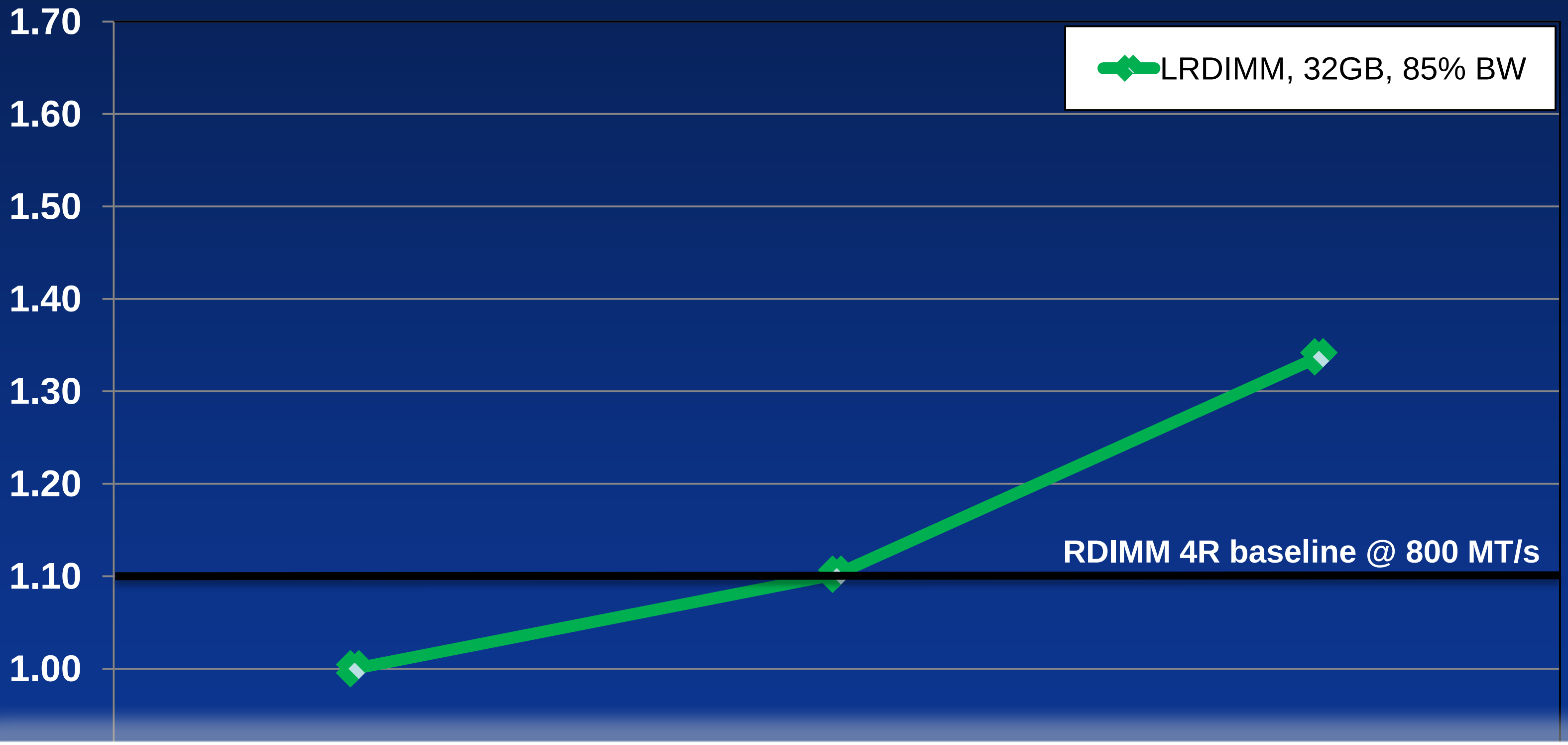




Source: Inphi Corp

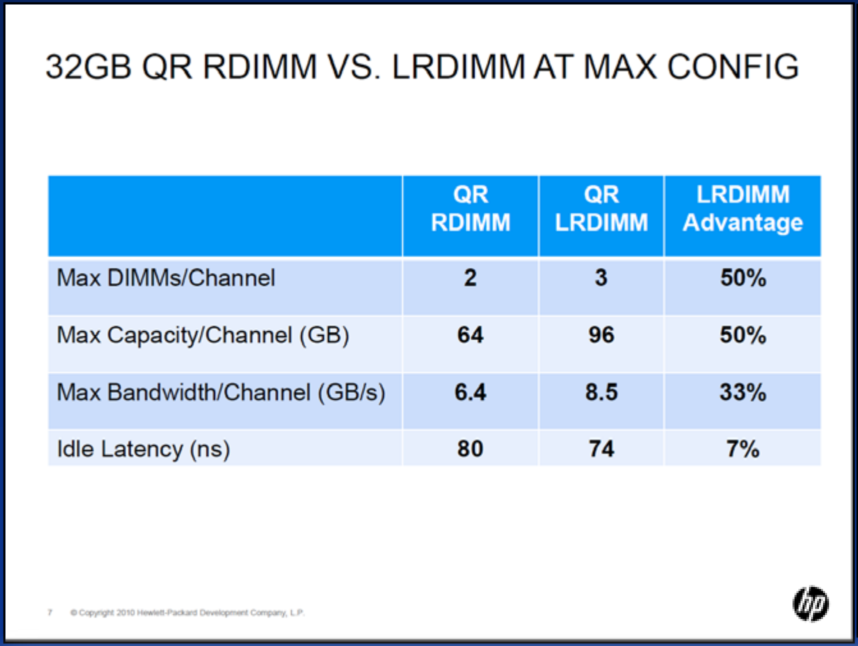
1. **System Configuration – Intel Dual Xeon 5600 populated with 2 channel (1 per CPU) of 2*16GB 4Rx4 DIMM per Channel**
2. **RDIMM and LRDIMM measured values using SANDRA Sisoft Benchmark on Intel 5600 system**

Normalized LRDIMM Power, 2 DIMM/channel



- ☐ **LRDIMM power converging with 4R RDIMM**
- ☐ **Possible due to MB features such as R-on and ODT control**

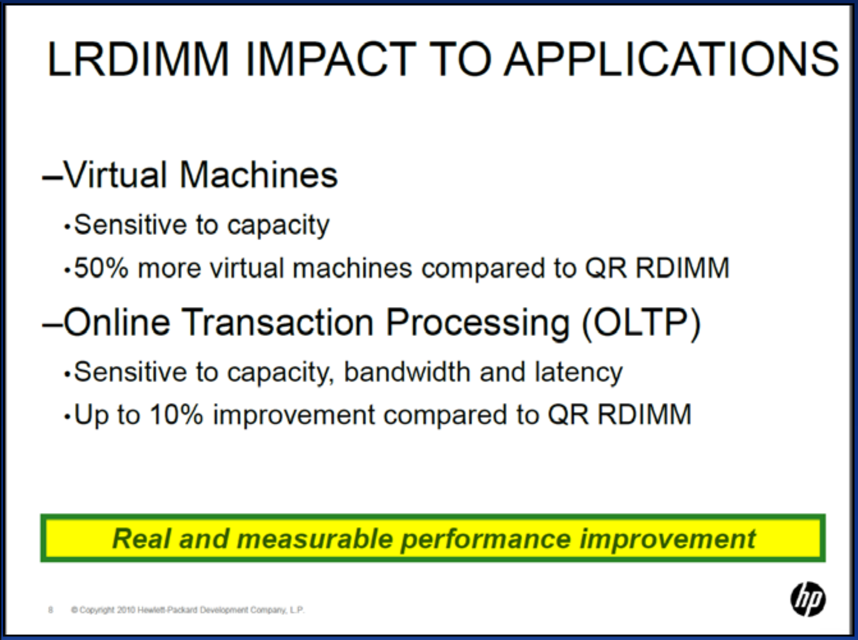
Server Vendor Perspective of LRDIMM



32GB QR RDIMM VS. LRDIMM AT MAX CONFIG

	QR RDIMM	QR LRDIMM	LRDIMM Advantage
Max DIMMs/Channel	**2**	**3**	**50%**
Max Capacity/Channel (GB)	**64**	**96**	**50%**
Max Bandwidth/Channel (GB/s)	**6.4**	**8.5**	**33%**
Idle Latency (ns)	**80**	**74**	**7%**

7 © Copyright 2010 Hewlett-Packard Development Company, L.P.



LRDIMM IMPACT TO APPLICATIONS

–Virtual Machines
- Sensitive to capacity
- 50% more virtual machines compared to QR RDIMM

–Online Transaction Processing (OLTP)
- Sensitive to capacity, bandwidth and latency
- Up to 10% improvement compared to QR RDIMM

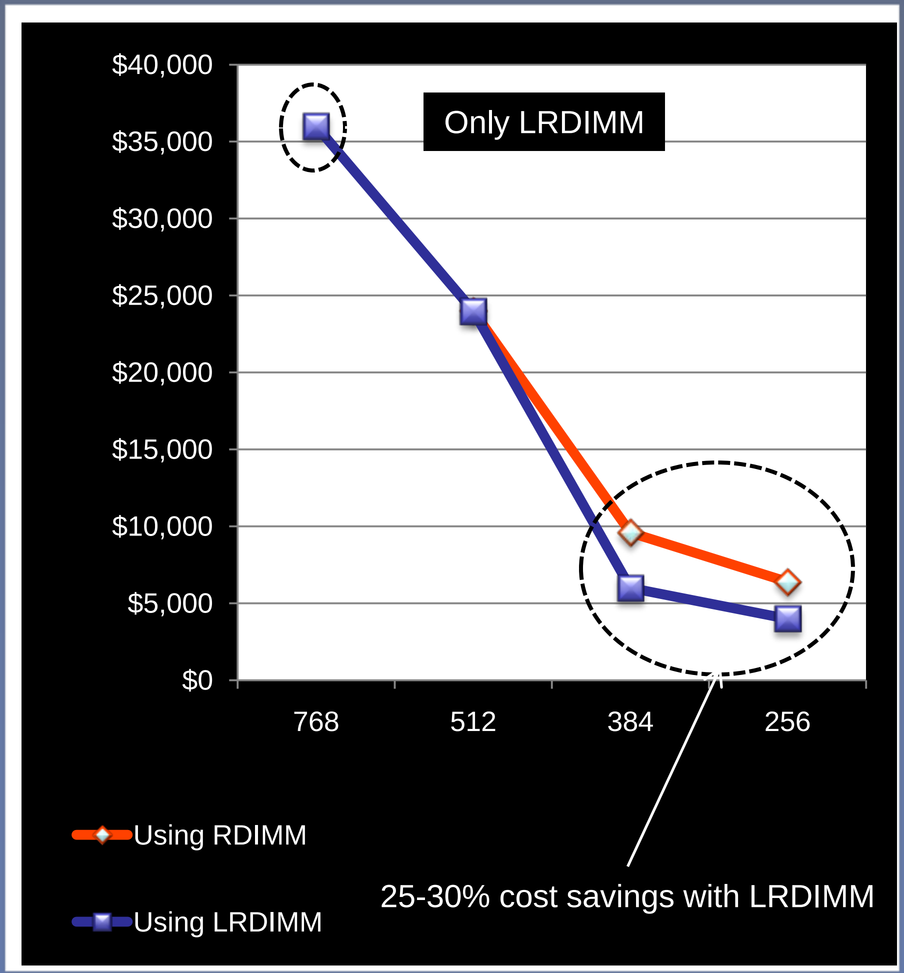
Real and measurable performance improvement

8 © Copyright 2010 Hewlett-Packard Development Company, L.P.

☐ **Significant application level benefit from LRDIMM to Virtualized, Financial and HPC applications**

LRDIMM Cost Advantages

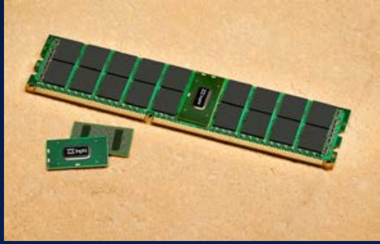
- Memory Buffer facilitates cost savings on LRDIMM for same system capacities
- This is possible due to the “rank multiplication” feature of Memory Buffer
  - Main benefit is from LRDIMM using lower density mainstream x4 DRAM technology (vs. leading edge) until cross-over
  - Once cross-over of DRAM technology occurs, incremental benefit is from LRDIMM using cheaper x8 DRAM



25-30% cost savings with LRDIMM

Note: Chart is for illustration only, based on Inphi’s market pricing estimate

Theme



- Challenges with Memory Bandwidth Scaling
- How LRDIMM Addresses this Challenge
- Under the hood
- LRDIMM Value Proposition

Key Enabler: The Memory Buffer (MB)

JEDEC LRDIMM Collaboration

Intel

AMD

HP

Inphi

Samsung

IBM

Dell

Elpida

IDT

Montage

Hynix

Nanya

Micron