Fully-Buffered DIMM Technology Moves Enterprise Platforms to the Next Level

Jon Haas
FB-DIMM Initiative Manager
Digital Enterprise Group
Intel Corporation

Pete Vogt
Principal Engineer
Digital Enterprise Group
Intel Corporation

Table of Contents

(Click on page number to jump to sections)

Fully-Buffered DIMM Technology Moves Enterprise Platforms to the Next Level

Jon Haas
FB-DIMM Initiative Manager
Digital Enterprise Group
Intel Corporation

Pete Vogt
Principal Engineer
Digital Enterprise Group
Intel Corporation

Overview: Enabling Memory to Keep Pace

Changes in enterprise platforms don't happen that often, but an exciting innovation—Fully-Buffered DIMM (FB-DIMM) technology—is just around the corner. To appreciate the extent of the new Fully-Buffered DIMM technology on upcoming enterprise platforms, you need to know a little about Amdahl's Law. Basically, a platform has three key pillars: processor, I/O and memory. Amdahl's Law states that it is ideal to have all three performing at the same level to achieve a perfectly balanced platform.

Recent advancements in processor performance driven by revolutionary breakthroughs such as Hyper-Threading Technology, multiple processing cores, and 64-bit computing have created an imbalance within the three-pillar structure. With the introduction of PCI Express* technology as the industry standard for local I/O interconnect, the I/O subsystem has established a new performance curve, keeping pace with processing advancements. However, in enterprise platforms, the memory subsystem is becoming a lagging system resource. That's where FB-DIMM technology comes in.

How did we get to this point? Basically, fewer DIMMs (dual inline memory modules) per channel are being supported as signaling rates are increasing. Given the need for cost-effective manufacturing, system vendors are faced with limited design options to support required memory sizes. Though the capacity per DIMM has been increasing due to improved DRAM (dynamic random access memory) density over time (**Figure 1**), unfortunately, total capacity per channel (the number of DIMMs per channel multiplied by the capacity per DIMM) has stayed flat at best and has even decreased in certain instances.

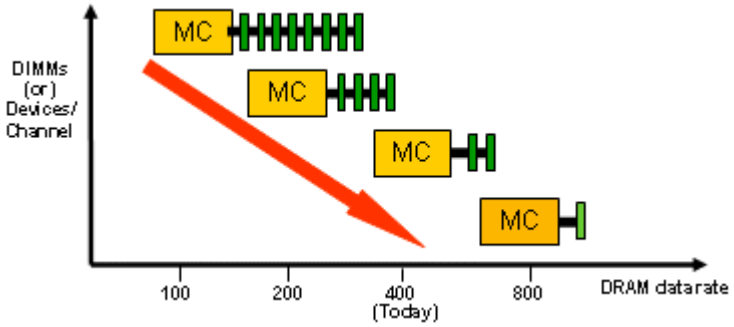


Figure 1. Scaling number of channels with memory hubs. Two ranks of DRAM devices per DIMM is assumed. In the case of single rank per DIMM, while the number of DIMMs per channel may be doubled, the declining trend shown in the figure remains the same.

Why not just increase the number of channels supported by the memory controller? The current parallel channel interface consumes a large number of pins (240) per channel. Increasing the number of channels supported would result in unacceptable cost increases to the platform.

Because of the enterprise platform's requirements for increased memory capacity to keep pace with both processor and I/O improvements, the industry has opted for a new approach called FB-DIMM technology. This new enterprise memory channel interface technology moves away from the parallel memory channel, furthering a total "serialization" of the platform.

Building the Recipe for FB-DIMM Technology

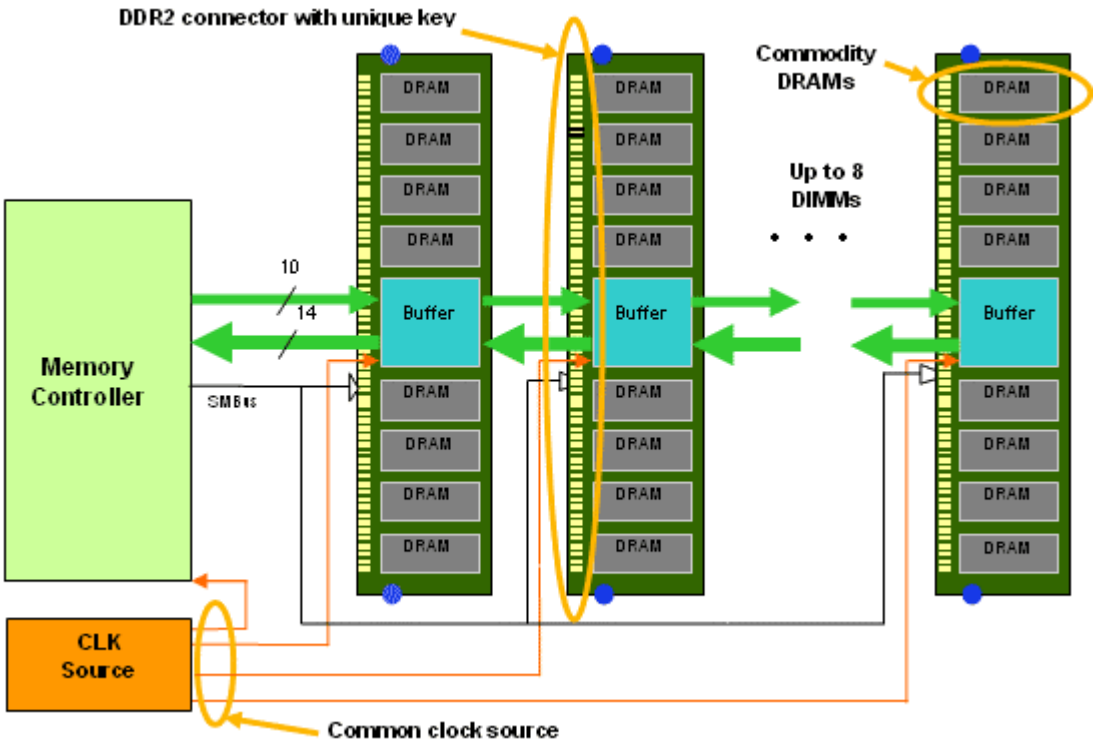
When the architects and designers sat down and discussed the requirements of this new memory architecture, it was decided that first and foremost, FB-DIMM technology would have to address the scaling needs of both capacity and bandwidth. It would also need to address current latency issues with memory hub architectures.

Then there's the business equation that needs to be considered: FB-DIMM technology would have to allow for low-cost and high-volume manufacturing to be affordable for mainstream platforms. To do that, it made sense to limit the changes to the memory industry and extend the industry investment in DRAM technology, so today's commodity DRAM chips and current DIMM form factor could be supported. Systems with the new DIMM interface standard must boot existing operating systems (OSes) without change, removing a significant barrier for adoption.

Additionally, the new FB-DIMM technology standard must embrace the next decade's requirements for enterprise servers, storage subsystems, workstations, and communications platforms. Because of the enterprise market segment focus, FB-DIMM technology also must raise the bar on overall memory RAS (reliability, availability, serviceability). It's important to note that because of its market segment focus, FB-DIMM technology is not expected to be optimized for 1-2 DIMM systems such as desktops, notebooks, and entry-level workstations.

FB-DIMM Memory Architecture

The FB-DIMM technology direct signaling interface between the memory controller and the DRAM chips is split into two independent signaling interfaces with a buffer between them. The interface between the buffer and DRAM chips is the same as today, supporting DDR2 (DDR stands for double data rate, a type of SDRAM memory; DDR2 is the second generation) in early FB-DIMM platforms and DDR3 in the future. However, the interface between the memory controller and the buffer is changed from a shared parallel interface to a point-to-point serial interface (see the figure below).



The buffer is referred to as the AMB (advanced memory buffer) and a number of suppliers, including Intel, are already making these. The AMB is designed to only take action in response to memory controller commands. The AMB is expected to deliver DRAM commands from the memory controller over the FB-DIMM interface without any alteration to the DRAM devices over the parallel DDR-based interface.

The end result is impressive scalability and throughput: FB-DIMM technology offers scalability of 192 gigabytes – 6 channels, 8 DIMMs/channel, 2 ranks/DIMM, 1 gigabyte DRAMs and offers bandwidth of 6.7 gigabytes per second (GBps) sustained data throughput per channel.

Improving Board Layouts

The FB-DIMM channel pin count is approximately 69 pins per channel, compared with about 240 pins for today's parallel channel. This results in less routing complexity and less routing area between the memory controller and DIMMs (**Figures 3 and 4**), thereby saving board cost to system manufacturers. For small factor systems such as 1U and blade systems, board real estate is in short supply, and the savings represented by the FB-DIMM technology transition are big.

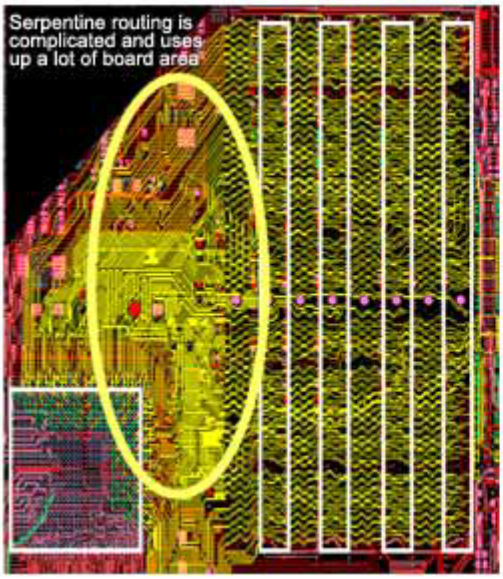


Figure 3. DDR2 Registered DIMMs: 1 Channel, 2 Routing Layers with 3rd layer required for power

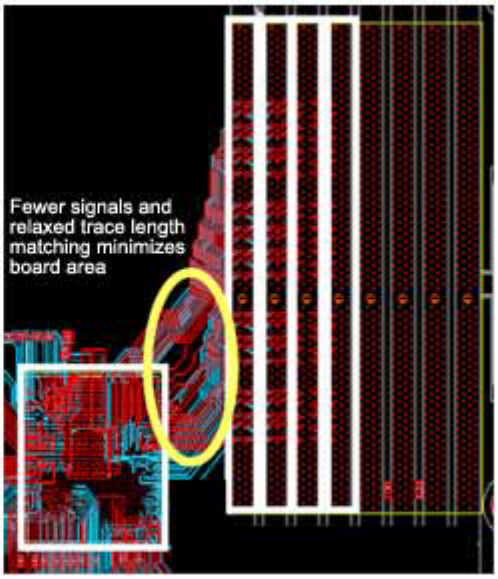


Figure 4. FB-DIMMs: 2 Channels, 2 Routing Layers (includes power delivery)

Reliability Now Built In

FB-DIMM technology offers better RAS (reliability, availability, serviceability) by extending the currently available ECC (error-correcting code, a method of checking the integrity of data in DRAM) to include protection of commands and address data. Additionally, FB-DIMM technology automatically retries when an error is detected, allowing for uninterrupted operation in case of transient errors.

Built-in Headroom for the Future

Since the FB-DIMM interface is based on serial differential signaling (similar to Serial ATA, Serial SCSI, PCI Express, and others), a memory controller can support multiple generations of FB-DIMM technology–based components. Today's platforms can support backward compatibility of memory devices (for example, both DDR and DDR2), extending the choice to on-site memory replacements and increasing system flexibility for IT environments. Bottom line, with FB-DIMM systems, an end user could have the flexibility of using first-generation FB-DIMMs with DDR2 DRAM or second-generation FB-DIMMs with DDR3 DRAM.

Reduced Total Cost of Ownership

FB-DIMM technology delivers better TCO (total cost of ownership) to IT in a number of ways:
Compatibility of FB-DIMMs across generations means that IT can extend the overall lifespan of DIMM investment through field swapping of DIMMs for new systems. Over time, IT will be able to use a newer generation of DIMMs for better performance or cost.

Due to headroom on capacity and bandwidth and Gen-X compatibility, IT can have more flexibility to repurpose a system for compute-intensive, data-intensive, or I/O-intensive applications, thereby providing better flexibility and range in reprovisioning.

With unprecedented RAS features on memory interfaces such as CRC (cyclical redundancy check) protection on address, retry, bit lane fail-over, hot add while active, and so on, IT would have fewer reasons to bring down the system, resulting in reduced down-time. Because FB-DIMM technology is transparent to OSes and applications, there are no significant barriers to adoption—or realizing—the benefits of the new technology.

Status of FB-DIMM Technology

The FB-DIMM technology standard is currently being authored within the JEDEC Solid State Technology Association industry standards body (www.jedec.org). Readers are encouraged to contact JEDEC directly for the latest status and access to industry standard specifications.

The memory industry has announced broad support of FB-DIMM components in time for the expected launch of FB-DIMM-enabled platforms in the first half of 2006. Intel has been working closely with the industry on FB-DIMM product support plans and is also working with industry tools vendors to ensure a healthy industry ecosystem of enabling tools and programs to fuel industry development.

Intel also founded the Memory Implementers Forum, an industry group cosponsored by Dell, Hewlett-Packard, and IBM, with a mission of accelerating industry development of memory technologies such as FB-DIMM for support of Intel® architecture platforms.

Summary

Fully-Buffered DIMM technology, a new memory architecture, addresses the scaling needs of both capacity and bandwidth and enables memory to keep pace with processor and I/O improvements in enterprise platforms.

More Info

At the Spring 2005 Intel Developer Forum, Intel unveiled a product development kit (PDK) to help accelerate the adoption of Fully-Buffered-DIMM technology. Read the official press release at the Intel Web site to learn more.

In 2004, Intel announced the formation of the Memory Implementers Forum. You can also read that announcement on the Intel Web site.

More information about Intel’s plans for FB-DIMM support can be found at the Memory Implementers Forum Web site.

Author Bios

Jon Haas is the initiative marketing and enabling manager for the Fully-Buffered DIMM Initiative in Intel's Digital Enterprise Group. Haas has managed a number of other initiative programs while at Intel, including PCI Express, InfiniBand* and the transitions to faster and wider parallel PCI implementations (32/33 to 64/66). Prior to managing I/O initiatives, he held both technical and product marketing engineering positions for several Intel business units. Haas joined Intel in 1986 as a field support engineer. He holds degrees from the Devry Institute of Technology and the University of Phoenix.

Pete Vogt is an architect in the Platform Architecture, Planning and Technology team, part of the Digital Enterprise Group. He has 35 years of experience in server architecture and memory subsystem design. He joined Intel in 1987. His achievements include the creation of numerous cache coherent bus architectures and multiprocessor system developments, including NuBus (used on Macintoshes*), C-Bus* (the world's first PC compatible multiprocessor), C-Bus II* (the world's first commodity multiprocessor chipset), and the 8-way Profusion* architecture. Vogt's current focus is on the architecture and development of future memory technologies for server platforms.

—End of Technology@Intel Magazine Article—