



# Memory technology evolution: an overview of system memory technologies

*technology brief, 8<sup>th</sup> edition*

Abstract.....	2
Introduction.....	2
Basic DRAM operation.....	2
DRAM storage density and power consumption.....	4
Memory access time.....	4
Chipsets and system bus timing.....	4
Memory bus speed.....	5
Burst mode access.....	5
SDRAM technology.....	6
Bank interleaving.....	7
Increased bandwidth.....	7
Registered SDRAM modules.....	7
DIMM Configurations.....	8
Single-sided and double-sided DIMMs.....	8
Single-rank, dual-rank, and quad-rank DIMMs.....	8
Rank interleaving.....	9
Memory channel interleaving.....	10
Advanced memory technologies.....	11
Double Data Rate SDRAM technologies.....	11
DDR-1.....	11
DDR-2.....	13
DDR-3.....	14
Module naming convention and peak bandwidth.....	14
Fully-Buffered DIMMs.....	15
FB-DIMM architecture.....	16
Challenges.....	17
Rambus DRAM.....	18
Importance of using HP-certified memory modules in ProLiant servers.....	19
Conclusion.....	19
For more information.....	20
Call to action.....	20

## Abstract

The widening performance gap between processors and memory along with the growth of memory-intensive business applications are driving the need for better memory technologies for servers and workstations. Consequently, there are several memory technologies on the market at any given time. HP evaluates developing memory technologies in terms of price, performance, and backward compatibility and implements the most promising technologies in ProLiant servers. HP is committed to providing customers with the most reliable memory at the lowest possible cost.

This paper summarizes the evolution of memory technology and provides an overview of some the newest memory technologies that HP is evaluating for servers and workstations. The purpose is to allay some of the confusion about the performance and benefits of the dynamic random access memory (DRAM) technologies on the market.

## Introduction

Processors use system memory to temporarily store the operating system, mission-critical applications, and the data they use and manipulate. Therefore, the performance of the applications and reliability of the data are intrinsically tied to the speed and bandwidth of the system memory. Over the years, these factors have driven the evolution of system memory from asynchronous DRAM technologies, such as Fast Page Mode (FPM) memory and Extended Data Out (EDO) memory, to high-bandwidth synchronous DRAM (SDRAM) technologies. Yet, system memory bandwidth has not kept pace with improvements in processor performance, thus creating a “performance gap.” Processor performance, which is often equated to the number of transistors in a chip, doubles every couple of years. On the other hand, memory bandwidth doubles roughly every three years. Therefore, if processor and memory performance continue to increase at these rates, the performance gap between them will widen.

Why is the processor-memory performance gap important? The processor is forced to idle while it waits for data from system memory. Thus, the performance gap prevents many applications from effectively using the full computing power of modern processors. In an attempt to narrow the performance gap, the industry vigorously pursues the development of new memory technologies. HP works with Joint Electronic Device Engineering Council (JEDEC) memory vendors and chipset developers during memory technology development to ensure that new memory products fulfill customer needs in regards to reliability, cost, and backward compatibility.

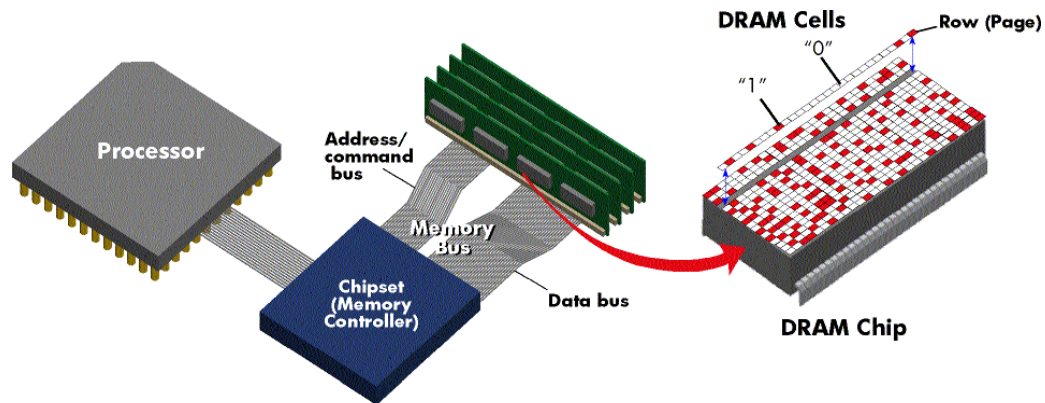
This paper describes the benefits and drawbacks regarding price, performance, and compatibility of DRAM technologies. Some descriptions are very technical. For readers who are not familiar with memory technology, the paper begins with a description of basic DRAM operation and terminology.

## Basic DRAM operation

Before a computer can perform any useful task, it copies applications and data from the hard disk drive to the system memory. Computers use two types of system memory—cache memory and main memory. Cache memory consists of very fast static RAM (SRAM) and is usually integrated with the processor. Main memory consists of DRAM chips that can be packaged in a variety of ways on dual inline memory modules (DIMMs) for the notebook, desktop PC, and server markets.

Each DRAM chip contains millions of memory locations, or cells, which are arranged in a matrix of rows and columns (Figure 1). On the periphery of the array of memory cells are transistors that read, amplify, and transfer the data from the memory cells to the memory bus. Each DRAM row, called a page, consists of several DRAM cells. Each DRAM cell on a page contains a capacitor capable of storing an electrical charge for a very short time. A charged cell represents a "1" data bit, and an uncharged cell represents a "0" data bit. The capacitors discharge over time so they must be recharged, or refreshed, thousands of times per second to maintain the validity of the data. These refresh mechanisms are described later in this section.

**Figure 1.** Representation of a single DRAM chip on a DIMM



The memory subsystem operates at the memory bus speed. Typically, a DRAM cell is accessed when the memory controller sends electronic address signals that specify the row address and column address of the target cell. The memory controller sends these signals to the DRAM chip by way of the memory bus. The memory bus consists of two sub-buses: the address/command bus and the data bus.

The data bus is a set of lines (traces) that carry the data to and from DRAM. Each trace carries one data bit at a time. The throughput (bandwidth) of the data bus depends on its width (in bits) and its frequency. The data width of a memory bus is usually 64-bits, which means that the bus has 64 traces, each of which transports one bit at a time. Each 64-bit unit of data is called a data word.

The address portion of the address/command bus is a set of traces that carry signals identifying the location of data in memory. The command portion of the address/command bus conveys instructions such as read, write, or refresh.

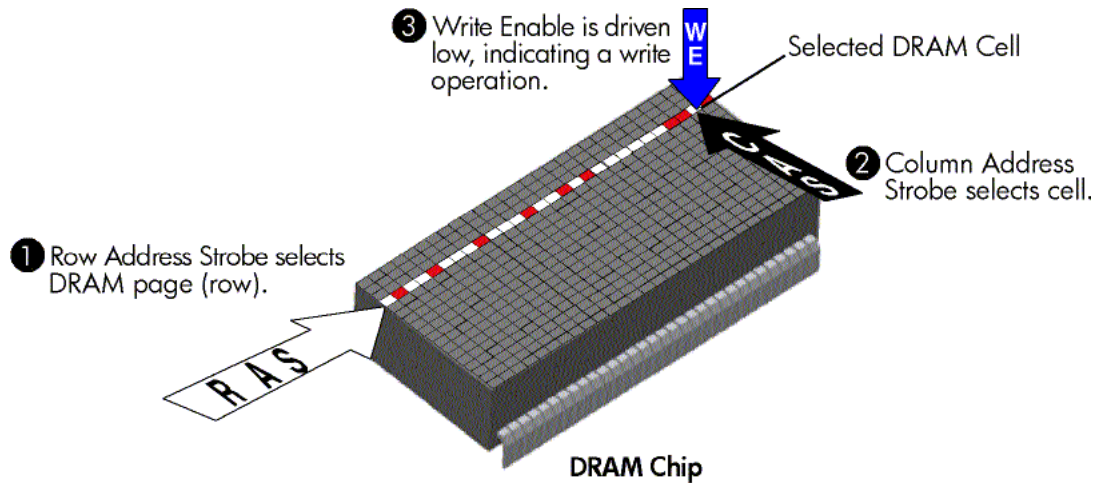
When FPM or EDO memory *writes* data to a particular cell, the location where the data will be written is selected by the memory controller. The memory controller first selects the page by strobing the Row Address onto the address/command bus. It then selects the exact location by strobing the Column Address onto the address/command bus (see Figure 2). These actions are called Row Address Strobe (RAS) and Column Address Strobe (CAS). The Write Enable (WE) signal is activated at the same time as the CAS to specify that a write operation is to be performed. The memory controller then drives the data onto the memory bus. The DRAM devices latch the data and store it into the respective cells.

During a DRAM read operation, RAS followed by CAS are driven onto the memory bus. The WE signal is held inactive, indicating a read operation. After a delay called CAS Latency, the DRAM devices drive the data onto the memory bus.

While DRAM is being refreshed, it cannot be accessed. If the processor makes a data request while the DRAM is being refreshed, the data will not be available until after the refresh is complete. There

are many mechanisms to refresh DRAM, including RAS only refresh, CAS before RAS (CBR) refresh, and Hidden refresh. CBR, which involves driving CAS active before driving RAS active, is used most often.

**Figure 2.** Representation of a write operation for FPM or EDO RAM



## DRAM storage density and power consumption

The storage capacity (density) of DRAM is inversely proportional to the cell geometry. In other words, storage density increases as cell geometry shrinks. Over the past few years, improvements in DRAM storage density have increased capacity from almost 1 kilobit (Kb) per chip to 2 gigabit (Gb) per chip. In the near future, it is expected that capacity will increase even further to 4 Gb per chip.

The industry-standard operating voltage for computer memory components was originally at 5 volts. However, as cell geometries decreased, memory circuitry became smaller and more sensitive. Likewise, the industry-standard operating voltage has decreased. Today, computer memory components operate at 1.8 volts, which allows them to run faster and consume less power.

## Memory access time

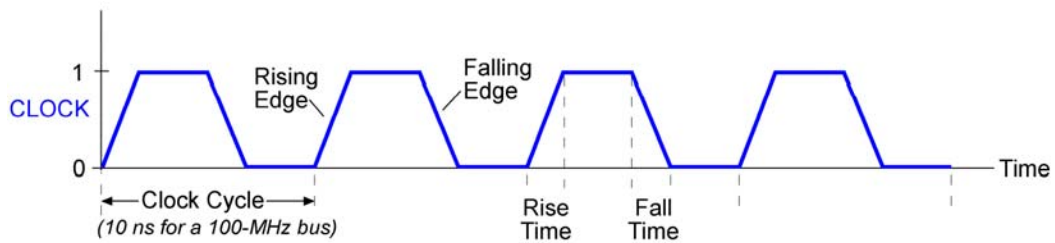
The length of time it takes for DRAM to produce the data, from the CAS signal until the data is available on the data bus, is called the memory access time or CAS Latency. Memory access time is measured in billionths of a second (nanoseconds, ns) for asynchronous DRAM. For synchronous DRAM, the time is converted to number of memory bus clocks.

## Chipsets and system bus timing

All computer components that execute instructions or transfer data are controlled by a system bus clock. The system chipset controls the speed, or frequency, of the system bus clock and thus regulates the traffic between the processor, main memory, PCI bus, and other peripheral buses.

The bus clock is an electronic signal that alternates between two voltages (designated as "0" and "1" in Figure 3) at a specific frequency. The bus frequency is measured in millions of cycles per second, or megahertz (MHz). During each clock cycle, the voltage signal transitions from "0" to "1" and back to "0". A complete clock cycle is measured from one rising edge to the next rising edge. Data transfer along the memory bus can be triggered on either the rising edge or falling edge of the clock signal.

**Figure 3.** Representation of a bus clock signal



The clock cycle time is measured in nanoseconds (ns,  $1 \times 10^{-9}$  second) and the clock frequency is measured in megahertz (MHz, millions of cycles per second). The clock frequency is the reciprocal of the clock cycle time. For example, if the clock cycle is 10 ns then the clock frequency is  $1/(0.000000010)=100$  MHz.

Over the years, some computer components have gained in speed more than others have. For this reason, the components in a typical server are controlled by different clocks that run at different, but related, speeds. These clocks are created by using various clock multiplier and divider circuits to generate multiple signals based on the main system bus clock. For example, if the main system bus operates at 100 MHz, a divider circuit can generate a PCI bus frequency of 33 MHz (system clock  $\div$  3) and a multiplier circuit can generate a processor frequency of 400 MHz (system clock  $\times$  4). Computer components that operate in whole multiples of the system clock are termed synchronous because they are “in sync” with the system clock.

Synchronous components operate more efficiently than components that are not synchronized (asynchronous) with the system bus clock. With asynchronous components, either the rest of the system or the component itself must wait one or more additional clock cycles for data or instructions due to clock resynchronization. In contrast, synchronized components know on which clock cycle data will be available, thus eliminating these timing delays.

## Memory bus speed

The speed of the DRAM is not the same as the true speed (or frequency) of the overall memory subsystem. The memory subsystem operates at the memory bus speed, which may not be the same frequency (in MHz) as the main system bus clock. The two main factors that control the speed of the memory subsystem are the memory timing and the maximum DRAM speed.

## Burst mode access

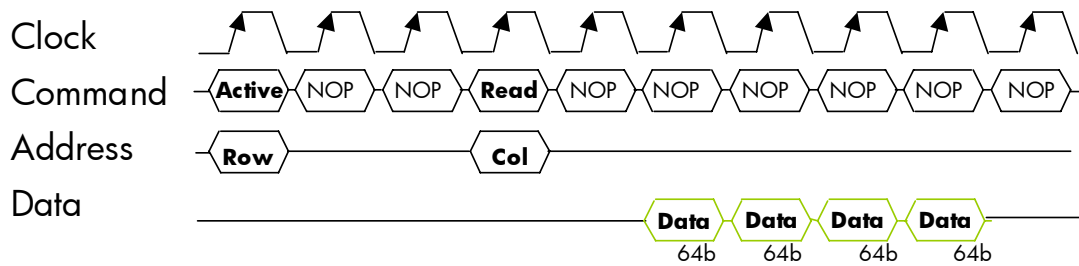
The original DRAM took approximately six system bus clock cycles for each memory access. During memory access, the RAS and CAS were sent first and then 64 bits of data were transferred through the memory bus. The next sequential address access required a repeat of the RAS-CAS-Data sequence. As a result, most of the overhead occurred while transferring row and column addresses, rather than the data.

FPM and EDO improved performance by automatically retrieving data from sequential memory locations on the assumption that they too will be requested. Using this process called burst mode access, four consecutive 64-bit sections of memory are accessed, one after the other, based on the address of the first section. So instead of taking six clock cycles to access each of the last three 64-bit sections, it may take from one to three clock cycles each (see Figure 4).

Burst mode access timing is normally stated in the format “x-y-y” where “x” represents the number of clock cycles to read/write the first 64 bits and “y” represents the number of clock cycles required for the second, third, and fourth reads/writes. For example, prior to burst mode access, DRAM took up to 24 clock cycles (6-6-6-6) to access four 64-bit memory sections. With burst mode access, three

additional data sections are accessed with every clock cycle after the first access (6-1-1-1) before the memory controller has to send another CAS.

**Figure 4.** Burst mode access. NOP is a “No Operation” instruction.



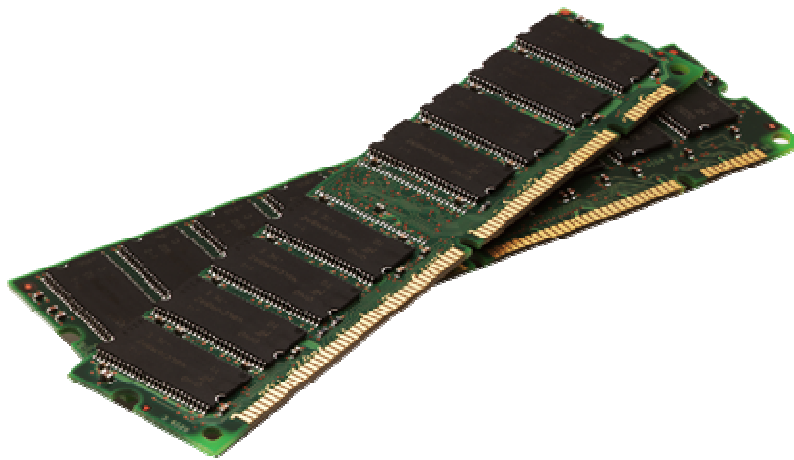
## SDRAM technology

FPM and EDO DRAMs are controlled asynchronously, that is, without a memory bus clock. The memory controller determined when to assert signals and when to expect data based on absolute timing. The inefficiencies of transferring data between a synchronous system bus and an asynchronous memory bus resulted in longer latency.

Consequently, JEDEC—the electronics industry standards agency for memory devices and modules—developed the synchronous DRAM standard to reduce the number of system clock cycles required to read or write data. SDRAM uses a memory bus clock to synchronize the input and output signals on the memory chip. This simplified the memory controller and reduced the latency from CPU to memory.

In addition to synchronous operation and burst mode access, SDRAM has other features that accelerate data retrieval and increase memory capacity—multiple memory banks, greater bandwidth, and register logic chips. Figure 5 shows SDRAM DIMMs with two key notches that prevent incorrect insertion and indicate a particular feature of the module.

**Figure 5.** SDRAM DIMM with two notches



## Bank interleaving

SDRAM divides memory into two to four banks for simultaneous access to more data. This division and simultaneous access is known as interleaving. Using a notebook analogy, two-way interleaving is like dividing each page in a notebook into two parts and having two assistants to each retrieve a different part of the page. Even though each assistant must take a break (be refreshed), breaks are staggered so that at least one assistant is working at all times. Therefore, they retrieve the data much faster than a single assistant could get the same data from one whole page, especially since no data can be accessed when a single assistant takes a break. In other words, while one memory bank is being accessed, the other bank remains ready to be accessed. This allows the processor to initiate a new memory access before the previous access has been completed, resulting in continuous data flow.

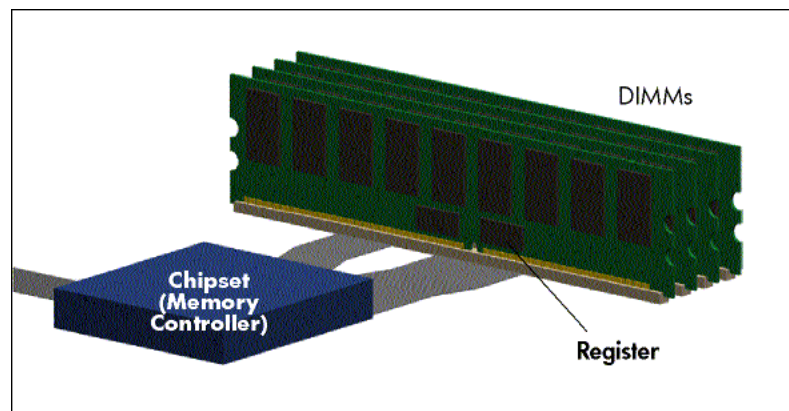
## Increased bandwidth

The bandwidth capacity of the memory bus increases with its width (in bits) and its frequency (in MHz). By transferring 8 bytes (64 bits) at a time and running at 100 MHz, SDRAM increases memory bandwidth to 800 MB/s, 50 percent more than EDO DRAMs (533 MB/s at 66 MHz).

## Registered SDRAM modules

To achieve higher memory subsystem capacity, some DIMMs have register logic chips (registers) that act as a pass-through buffer for address and command signals (Figure 6). Registers prevent the chipset from having to drive the entire arrangement of DRAM chips on each module. Rather, the chipset drives only the loading of the registers on each module. The register on each DIMM re-drives the address and command signals to the appropriate DRAM chip. Simultaneously, a phase lock loop chip on the registered DIMM generates a second clock signal that runs synchronously with the system bus clock. This prevents the system bus clock signal from having to drive all the DRAM chips, and it allows the addition of more memory modules to the memory bus to increase memory capacity.

**Figure 6.** Registered DIMMs



# DIMM Configurations

## Single-sided and double-sided DIMMs

Each DRAM chip on a DIMM provides either 4 bits or 8 bits of a 64-bit data word. Chips that provide 4 bits are called x4 (by 4), and chips that provide 8 bits are called x8 (by 8). It takes eight x8 chips or sixteen x4 chips to make a 64-bit word, so at least eight chips are located on one or both sides of a DIMM. However, a standard DIMM has enough room to hold a ninth chip on each side. The ninth chip is used to store 4 bits or 8 bits of Error Correction Code, or ECC (see “Parity and ECC DIMMs” sidebar on next page).

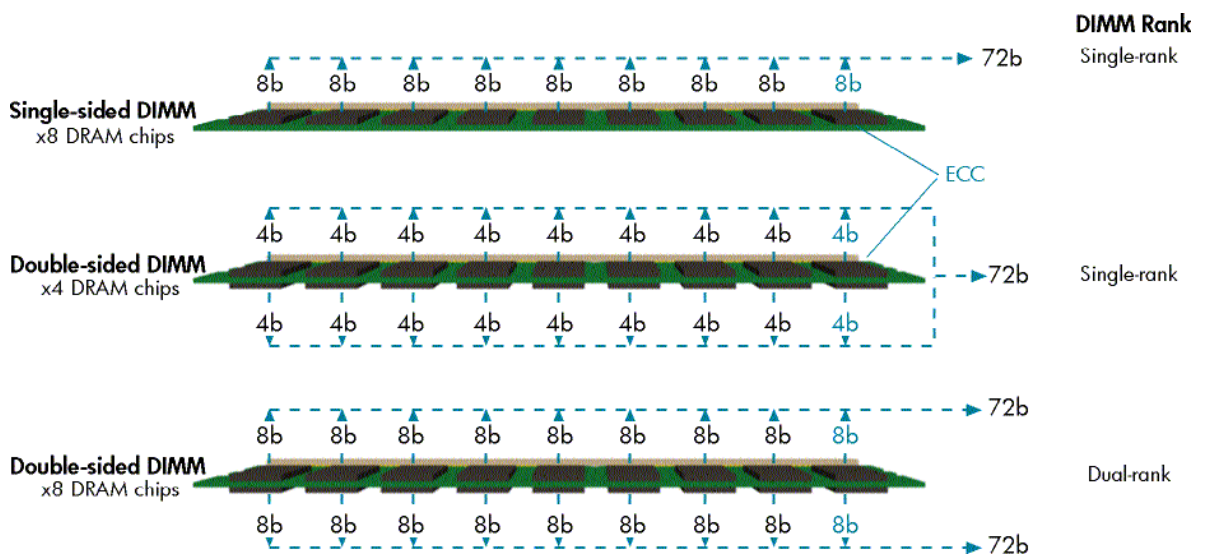
An ECC DIMM with all nine DRAM chips on one side is called single-sided, and an ECC DIMM with nine DRAM chips on each side is called double-sided (Figure 7). A single-sided x8 ECC DIMM and a double-sided x4 ECC DIMM each create a single block of 72 bits (64 bits plus 8 ECC bits). In both cases, a single chip-select signal from the chipset is used to activate all the chips on the DIMM. In contrast, a double-sided x8 DIMM (bottom illustration) requires two chip-select signals to access two 72-bit blocks on two sets of DRAM chips.

## Single-rank, dual-rank, and quad-rank DIMMs

In addition to single-sided and double-sided configurations, DIMMs are classified as single-rank or dual-rank. A memory rank is defined as an area or block of 64-bits created by using some or all of the DRAM chips on a DIMM. For an ECC DIMM, a memory rank is a block of 72 data bits (64 bits plus 8 ECC bits).

A single-rank ECC DIMM (x4 or x8) uses all of its DRAM chips to create a single block of 72 bits, and all the chips are activated by one chip-select signal from the chipset (top two illustrations in Figure 7). A dual-rank ECC DIMM produces two 72-bit blocks from two sets of DRAM chips on the DIMM, requiring two chip-select signals. The chip-select signals are staggered so that both sets of DRAM chips do not contend for the memory bus at the same time. Quad-rank DIMMs with ECC produces four 72-bit blocks from four sets of DRAM chips on the DIMM, requiring four chip-select signals. Like the dual-rank DIMMs, the memory controller staggers the chip-select signals to prevent the four sets of DRAM chips from contending for the memory bus at the same time.

Figure 7. Single-sided and double-sided DDR SDRAM DIMMs and corresponding DIMM rank





---

### **Parity and ECC DIMMs**

The ninth DRAM chip on one side of a DIMM is used to store parity or ECC bits. With parity, the memory controller is capable of detecting single-bit errors, but it is unable to correct any errors. Also, it cannot consistently detect multiple-bit errors. With ECC, the memory controller is capable of detecting and correcting single bit errors and multiple-bit errors that are contiguous. Multiple-bit contiguous errors occur when an entire x4 or x8 chip fails. The chipset (memory controller) is also capable of detecting double-bit errors that are not contiguous. The chipset halts the system and logs an error when uncorrectable errors are detected. Servers use ECC DIMMs to improve availability and reliability.

---

Memory ranks are not new, but their role has become more critical with the advent of new chipset and memory technologies and growing server memory capacities. Dual-rank DIMMs improve memory density by placing the components of two single-rank DIMMs in the space of one module. The chipset considers each rank as an electrical load on the memory bus. At slower bus speeds, the number of loads does not adversely affect bus signal integrity. However, for faster memory technologies such as DDR2-667, there are a maximum number of ranks that the chipset can drive. For example, if a memory bus on a server has four DIMM slots, the chipset may only be capable of supporting two dual-rank DIMMs or four single rank DIMMs. If two dual-rank DIMMs are installed then, the last two slots must not be populated. To compensate for the reduction in the number of DIMM slots on a bus at higher speeds, modern chipsets employ multiple memory buses.

If the total number of ranks in the populated DIMM slots exceeds the maximum number of loads the chipset can support, the server may not boot properly or it may not operate reliably. Some systems check the memory configuration while booting to detect invalid memory bus loading. When an invalid memory configuration is detected, the system stops the boot process, thus avoiding unreliable operation.

To prevent this and other memory-related problems, customers should only use HP-certified DIMMs available in the memory option kits for each ProLiant server (see the “Importance of using HP-certified memory modules in ProLiant servers” section).

Another important difference between single-rank and dual-rank DIMMs is cost. Typically, memory costs increase with DRAM density. For example, the cost of an advanced, high-density DRAM chip is typically much higher (more than 2x) than that of a conventional DRAM chip. Because large capacity single-rank DIMMs are manufactured with higher-density DRAM chips, they typically cost more than dual-rank DIMMs of comparable capacity.

### **Rank interleaving**

As described previously, bank interleaving allows the processor to initiate an access to one memory bank before the previous access to a different bank has been completed, resulting in continuous data flow. Essentially, dual-rank and quad-rank DIMMs increase the number of “banks” available for interleaving. Rank interleaving can result in improved system performance for memory-intensive applications such as high-performance computing and video rendering.

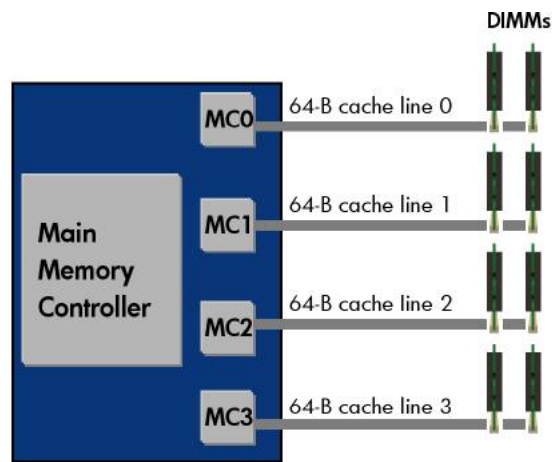
## Memory channel interleaving

Multi-core processors running multi-threaded applications pose a significant challenge to the memory subsystem. The processor cores share the bandwidth of the memory bus; therefore, the multi-core processor's performance is limited by the memory bus bandwidth. Even with sufficient memory bus bandwidth, the actual throughput of a single memory controller can create a bottleneck as it handles memory requests from multiple cores.

To overcome this bottleneck, manufacturers are designing memory controller chips with multiple integrated memory controllers (Figure 8). The chip can contain two, three, or four memory controllers that operate independently of each other to access up to two DIMMs per channel. This enables a process called channel interleaving. In channel interleaving, each integrated memory controller successively provides a 64-byte cache line of data from the first DIMM on its channel. After the last memory controller completes the data transfer, the memory controllers can provide a cache line from a second DIMM on each channel. Channel interleaving does not prevent bank or rank interleaving.

The effective throughput of the memory controller is the sum of the individual memory channels. As the number of cores on a single processor increases, the number of integrated memory controllers will need to increase accordingly to provide the necessary throughput.

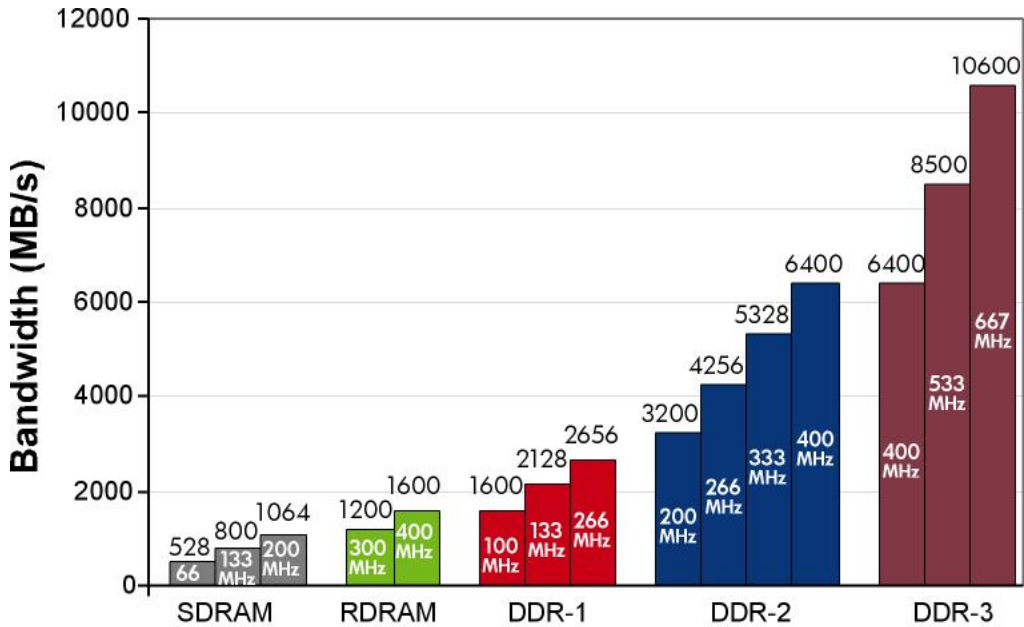
**Figure 8.** Memory channel interleaving using multiple integrated memory controllers



# Advanced memory technologies

Despite the performance improvement in the overall system due to use of SDRAM, the growing performance gap between the memory and processor must be filled by more advanced memory technologies. These technologies, which are described on the following pages, boost the overall performance of systems using the latest high-speed processors (Figure 9).

**Figure 9.** Peak bandwidth comparison of SDRAM and advanced SDRAM technologies



## Double Data Rate SDRAM technologies

Double Data Rate (DDR) SDRAM is advantageous for systems that require higher bandwidth than can be obtained using SDRAM. Basically, DDR SDRAM doubles the transfer rate without increasing the frequency of the memory clock. This section describes three generations of DDR SDRAM technology.

### DDR-1

To develop the first generation of DDR SDRAM (DDR-1), designers made enhancements to the SDRAM core to increase the data rate. These enhancements include prefetching, double transition clocking, strobe-based data bus, and SSTL\_2 low voltage signaling. At 400 MHz, DDR increases memory bandwidth to 3.2 GB/s, which is 400 percent more than original SDRAM.

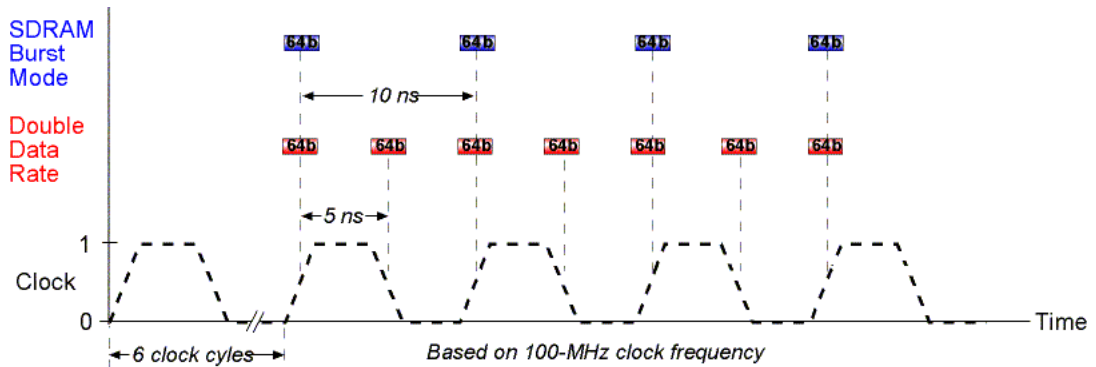
#### Prefetching

In SDRAM, one bit per clock cycle is transferred from the memory cell array to the input/output (I/O) buffer or data queue (DQ). The I/O buffer releases one bit to the bus per pin and clock cycle (on the rising edge of the clock signal). To double the data rate, DDR SDRAM uses a technique called prefetching to transfer two bits from the memory cell array to the I/O buffer in two separate pipelines. Then the I/O buffer releases the bits in the order of the queue on the same output line. This is known as a 2n-prefetch architecture because the two data bits are fetched from the memory cell array before they are released to the bus in a time multiplexed manner.

### Double transition clocking

Standard DRAM transfers one data bit to the bus on the rising edge of the bus clock signal, while DDR-1 uses both the rising and falling edges of the clock to trigger the data transfer to the bus (Figure 10). This technique, known as double transition clocking, delivers twice the bandwidth of SDRAM without increasing the clock frequency. DDR-1 has theoretical peak data transfer rates of 1.6 and 2.1 GB/s at clock frequencies of 100 MHz and 133 MHz, respectively.

**Figure 10.** Data transfer rate comparison between SDRAM (with burst mode access) and DDR SDRAM



### SSTL\_2 low-voltage signaling technology

Another difference between SDRAM and DDR-1 is the signaling technology. Instead of using a 3.3-V operating voltage, DDR-1 uses a 2.5-V signaling specification known as Stub Series-Terminated Logic<sub>2</sub> (SSTL<sub>2</sub>). This low-voltage signaling results in lower power consumption and improved heat dissipation.

### Strobe-based data bus

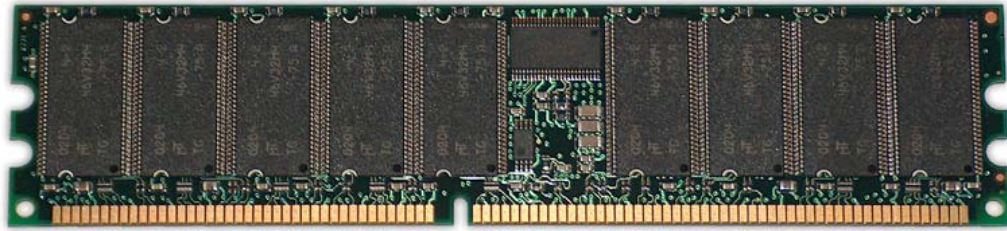
SSTL<sub>2</sub> signaling allows DDR-1 to run at faster speeds than traditional SDRAM. In addition, DDR-1 uses a delay-locked loop (one for every 16 outputs) to provide a data strobe signal as data becomes valid on the SDRAM pins. The memory controller uses the data strobe signal to locate data more accurately and resynchronize incoming data from different DIMMs.

DDR-1 operates at transfer rates of 400 Mb/s, or 3.2 GB/s. Although the data bus is capable of running at these speeds, the command bus cannot. Tight system timing requirements were alleviated on the data bus by using strobes. However, the command bus does not use a strobe and must still meet setup times to a synchronous clock. Thus, at a data rate of 400 Mb/s, the command bus must operate at 200 MHz.

## DDR-1 DIMMs

DDR-1 DIMMs require 184 pins instead of the 168 pins used by standard SDRAM DIMMs. DDR-1 is versatile enough to be used in desktop PCs or servers. To vary the cost of DDR-1 DIMMs for these different markets, memory manufacturers provide unbuffered and registered versions. Unbuffered DDR-1 DIMMs place the load of all the DDR modules on the system memory bus, but they can be used in systems that do not require high memory capacity. Registered DDR-1 DIMMs (Figure 11) place only one load per DIMM on the memory bus, regardless of how many SDRAM devices are on the module. Therefore, they are best suited for servers with very high memory capacities.

**Figure 11.** The 184-pin DDR-1 Registered DIMM. The DDR-1 DIMM has one notch instead of the two notches found on SDRAM DIMMs.



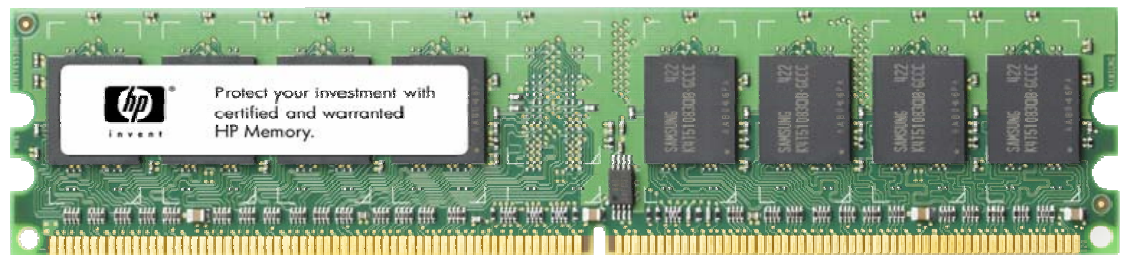
## Backward compatibility

Because of their different data strobes, voltage levels, and signaling technologies, it is not possible to mix SDRAM and DDR-1 DIMMs within the same memory subsystem.

## **DDR-2**

DDR-2 SDRAM is the second generation of DDR SDRAM. It offers data rates of up to 6.4 GB/s, lower power consumption, and improvements in packaging. At 400 MHz and 800 Mb/s, DDR-2 increases memory bandwidth to 6.4 GB/s, which is 800 percent more than original SDRAM. DDR-2 SDRAM achieves this higher level of performance and lower power consumption through faster clocks, 1.8-V operation and signaling, and simplification of the command set. The 240-pin connector on DDR-2 is needed to accommodate differential strobes signals (Figure 12).

**Figure 12.** The DDR-2 DIMM has a 240-pin interface



### DDR-3

DDR-3, the third-generation of DDR SDRAM technology, will make further improvements in bandwidth and power consumption. Manufacturers of DDR-3 will initially use 90 nm fabrication technology and move toward 70 nm as production volumes increase. DDR-3 will operate at clock rates from 400 MHz to 800 MHz with theoretical peak bandwidths ranging from 6.40 GB/s to 12.8 GB/s. DDR-3 is expected to reduce power consumption by up to 30% compared to a DDR-2 DIMM operating at the same speed. DDR-3 DIMMs are expected to use the same 240-pin connector as DDR2 DIMMs, but the key notch will be in a different position.

To increase performance and reduce power consumption of DDR-3, designers made several key enhancements, including:

- An 8-bit prefetch buffer, compared to the 4-bit buffer for DDR-2, storing more data before it is needed.
- Fly-by topology (for the commands, addresses, control signals, and clocks) improves signal integrity by reducing the number of stubs and their length. This feature requires the controller to support “write leveling” on DDR-3 DIMMs.
- 1.5-V signaling (compared to 1.8 V for DDR-2) for lower power consumption
- A thermal sensor integrated on the DIMM module signals the chipset to throttle memory traffic to the DIMM if its temperature exceeds a programmable critical trip point.

### Module naming convention and peak bandwidth

Table 1 summarizes the various types of DDR-1, DDR-2, and DDR-3 SDRAM and their associated naming conventions. Originally, the module naming convention for DDR-SDRAM was based on the effective clock rate of the data transfer: PC200 for DDR SDRAM that operates at 100 MHz; PC266 for 133 MHz; and so forth. But after confusion arose over the Rambus naming convention, the industry based the DDR-SDRAM naming convention on the actual peak data transfer rate in MB/s. For example, PC266 is equivalent to PC2100 ( $64 \text{ bit} * 2 * 133 \text{ MHz} = 2.1 \text{ GB/s}$  or 2100 MB/s).

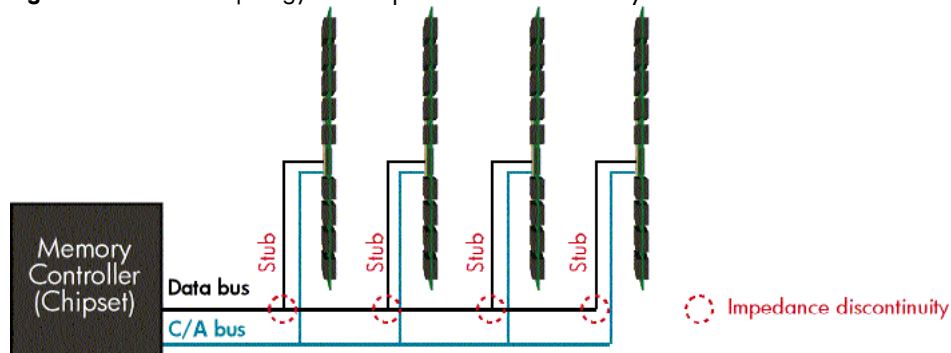
**Table 1.** Summary of DDR SDRAM technologies

Type	Component naming convention	Module naming convention	Bus speed	Peak bandwidth
DDR-1	DDR200	PC1600	100 MHz	1.6 GB/s
	DDR266	PC2100	133 MHz	2.1 GB/s
	DDR333	PC2700	166 MHz	2.7 GB/s
	DDR400	PC3200	200 MHz	3.2 GB/s
DDR-2	DDR2-400	PC2-3200R	200 MHz	3.2 GB/s
	DDR2-533	PC2-4300	266 MHz	4.3 GB/s
	DDR2-667	PC2-5300	333 MHz	5.3 GB/s
	DDR2-800	PC2-6400	400 MHz	6.4 GB/s
DDR-3	DDR3-800	PC3-6400	400 MHz	6.4 GB/s
	DDR3-1066	PC3-8500	533 MHz	8.5 GB/s
	DDR3-1333	PC3-10600	667 MHz	10.6 GB/s
	DDR3-1600	PC3-12800	800 MHz	12.8 GB/s

## Fully-Buffered DIMMs

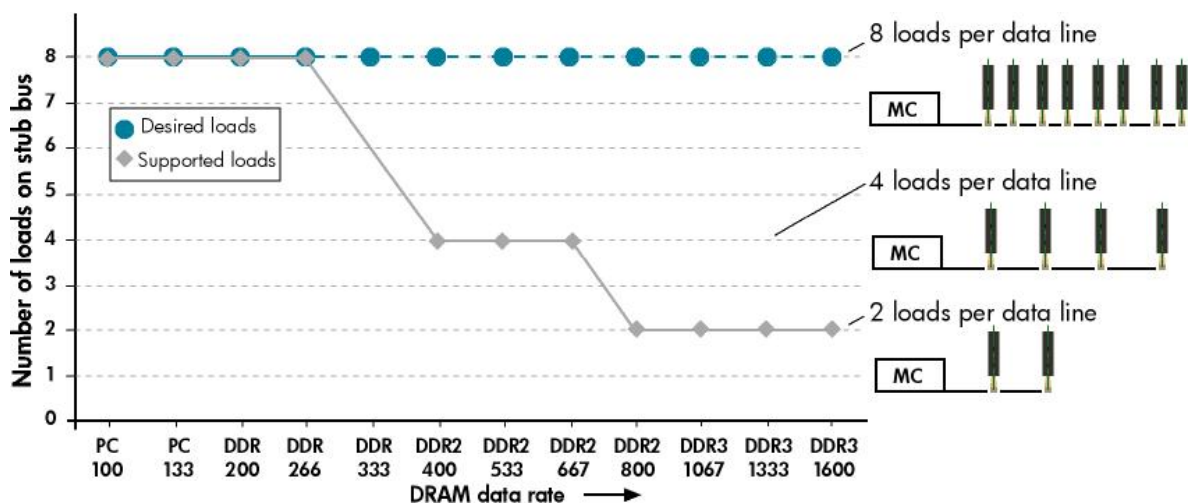
Traditional DIMM architectures use a stub-bus topology with parallel branches (stubs) that connect to a shared memory bus (Figure 13). Each DIMM connects to the data bus using a set of pin connectors. In order for the electrical signals from the memory controller to reach the DIMM bus-pin connections at the same time, all the traces have to be the same length. This can result in circuitous traces on the motherboard between the memory controller and memory slots. Both the latency resulting from complex routing of traces and signal degradation at the bus-pin connections cause the error rate to increase as the bus speed increases.

**Figure 13.** Stub-bus topology. An impedance discontinuity is created at each stub-bus connection.



Each stub-bus connection creates an impedance discontinuity that negatively affects signal integrity. In addition, each DIMM creates an electrical load on the bus. The electrical load accumulates as DIMMs are added. These factors decrease the number of DIMMs per channel that can be supported as the bus speed increases. For example, Figure 14 shows the number of loads supported per channel at data rates ranging from PC 100 to DDR-3 1600. Note that the number of supported loads drops from eight to two as data rates increase to DDR2 800.

**Figure 14.** Maximum number of loads per channel based on DRAM data rate.



Increasing the number of channels to compensate for the drop in capacity per channel was not a viable option due to increased cost and board complexity. System designers had two options: limit memory capacity so that fewer errors occur at higher speeds, or use slower bus speeds and increase the DRAM density. For future generations of high-performance servers, neither option was acceptable.

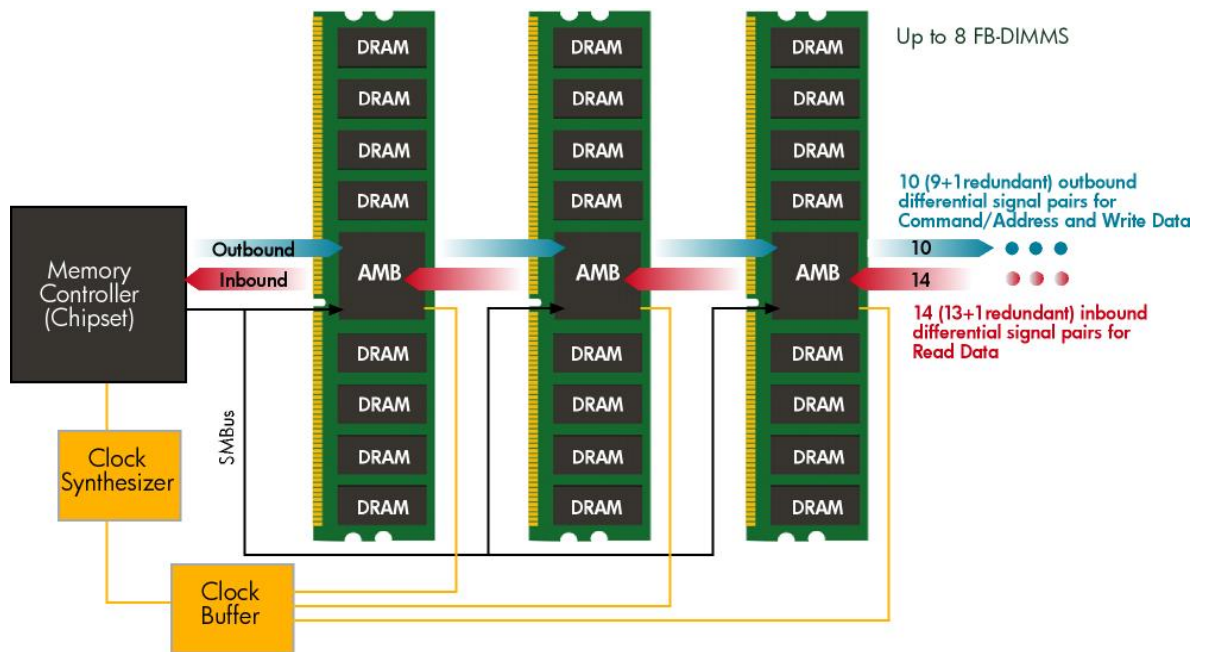
Consequently, JEDEC developed the Fully-Buffered DIMM (FB-DIMM) specification, a serial interface that eliminates the parallel stub-bus topology and allows higher memory bandwidth while maintaining or increasing memory capacity.

### FB-DIMM architecture

The FB-DIMM architecture has serial links between the memory controller and the FB-DIMMs, which are connected in a daisy chain configuration (Figure 15). Relative to the memory controller, there are 10 outbound links and 14 inbound links, also known as southbound and northbound links, respectively. These serial links connect the memory controller to an advanced memory buffer (AMB) chip that resides on each FB-DIMM, creating a point-to-point architecture. The outbound links transmit commands and write data to the FB-DIMMs while the inbound links transmit read data back to the memory controller.

The clock signal is distributed over a different set of pins. In addition to communicating over the Outbound lanes, the memory controller communicates configuration information with each AMB over the SMBus. The AMB is an intelligent chip that manages serial communication with the memory controller and parallel communication with local DRAM devices. Each AMB receives signals (address, write data, and command information) through the outbound links and re-transmits the signal to the next FB-DIMM on the channel. Each AMB decodes the command data and ignores the commands that are targeted for a different DIMM. The targeted AMB performs a read or write operation to local DRAM devices through a parallel interface. In the case of a read operation, the AMB serializes data from the DRAM devices and transmits it to the memory controller through the inbound links.

**Figure 15.** Serial communication between daisy-chained FB-DIMMs on a single channel.



When using DDR2-667 DRAM on the FB-DIMM, the peak theoretical throughput of the inbound links is 5.4 GB/s. The peak theoretical throughput of the outbound links is half the amount of data as the inbound links, approximately 2.6 GB/s.



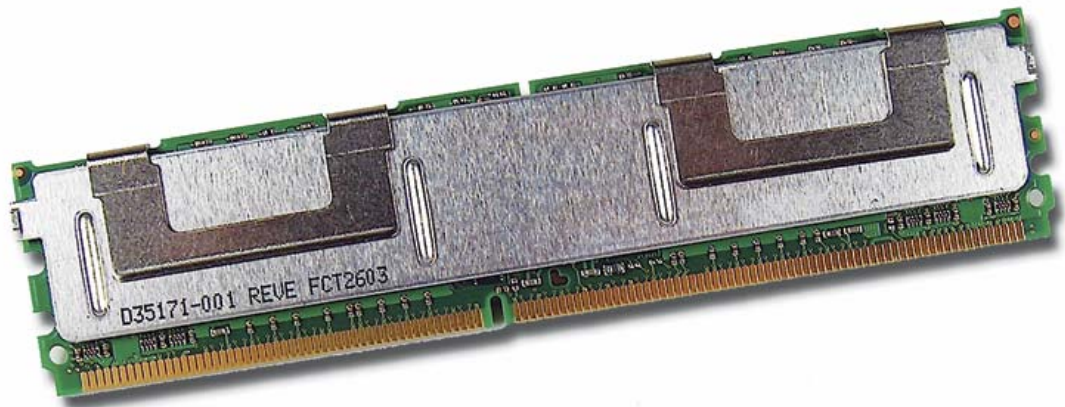
## Challenges

The challenges for the FB-DIMM architecture include latency and power use (thermal load).

Memory latency is the delay from the time the data is requested to the time when the data is available from the memory controller. The FB-DIMM architecture increases this latency in two ways: serialization and transmission. Serialization latency is added to memory access when the data is serialized prior to transmission, and then de-serialized after it reaches its destination. Additionally, the transmission latency, which exists for all memory technologies, increases due to the point-to-point interconnection of AMBs. As more DIMMs are added to the channel, the total transmission latency increases. Each FB-DIMM can add 2 to 6 nanoseconds (ns) of transmission latency; therefore, the cumulative transmission latency can be significant in a fully-scaled system.

An FB-DIMM consumes almost 5 Watts more than a typical registered DDR2 DIMM. The operation of the AMB also causes the FB-DIMM to get hotter. Therefore, a heat spreader is required to help draw heat away from the FB-DIMM so it can be cooled more efficiently by the server's internal fans (Figure 16). These concerns are driving the design of AMBs that use 15 to 20 percent less power, even though some manufacturers claim up to 40 percent power reduction.

**Figure 16.** FB-DIMM with full module heat spreader



## Rambus DRAM

Rambus DRAM (RDRAM) allows data transfer through a bus operating in a higher frequency range than DDR SDRAM. In essence, Rambus moves small amounts of data very fast, whereas DDR SDRAM moves large amounts of data more slowly. The Rambus design consists of three key elements: RDRAMs, Rambus application-specific integrated circuits, and an interconnect called the Rambus Channel. The Rambus design provides higher performance than traditional SDRAM because RDRAM transfers data on both edges of a synchronous, high-speed clock pulse. RDRAM uses a separate row and column command bus that allows multiple commands to be issued at the same time, thereby increasing the bandwidth efficiency of the memory bus. This dual command bus is a unique feature of RDRAM.

With only an 8-bit-wide command bus and an 18-bit data bus, RDRAM (Figure 17) has the lowest signal count of all of the memory technologies. RDRAM incorporates a packet protocol and is capable of operating at 800 MHz and providing a peak bandwidth of 2.4 GB/s. One packet of information is transferred in 8 ticks of the clock, which allows sending 128 bits of data in a 150-MHz clock period. Since it requires 8 ticks of the clock to transfer a packet, the internal memory controller only needs to run at a speed of 150 MHz to keep up with the packet transfer rate at 1.2 GHz. This allows for plenty of timing margin in the design of the memory controller.

**Figure 17.** Rambus DRAM



RDRAM is capable of supporting up to 32 RDRAM devices on one memory channel while maintaining a 1.2-GHz data rate. Through the use of a repeater chip, even more devices can be placed on one RDRAM channel. The repeater will interface to two different RDRAM channels and pass the data and command signals between them. One channel will communicate with the memory controller, and the other channel will communicate with the RDRAM devices. Thus, the memory controller essentially will be communicating only with the repeater chips. Up to eight repeater chips can be placed on the memory controller, and 32 RDRAM devices can be placed on each channel. This allows one channel to support a maximum of 256 devices. However, using the repeater chips will add 1 to 1.5 clocks of additional delay.

To account for differences in distance of the devices on the channel, more latency in increments of the clock can be added. This allows the memory controller to receive data from all devices in the same amount of time, thus preventing data collision on the bus when consecutive reads are performed to different devices.

Another feature of RDRAM that helps to increase the efficiency is an internal 128-bit write buffer. All write data is placed into this buffer before being sent to the DRAM core. The write buffer reduces the delay needed to turn around the internal data bus by allowing the sense amps to remain in the read direction until data needs to be retired from the buffer to the core. Essentially, a read can immediately follow a write with little bandwidth lost on the data bus.

While the RDRAM bus efficiency is high, the packet protocol increases the latency. The packet translation between the internal memory controller bus and the fast external bus requires one to two clocks of additional delay. This delay cannot be avoided when using a very fast packet protocol.

With the high data rate of Rambus, signal integrity is troublesome. System boards must be designed to accommodate the extremely stringent timing of Rambus, and this increases product time to market.

Additionally, each Rambus channel is limited to 32 devices, imposing an upper limit on memory capacity supported by a single bus. Use of repeater chips enables use of additional devices and increases potential memory capacity, but repeater chips have been very challenging to design.

Finally, the larger dies and more limited production of RDRAM compared to those of other memory technologies have increased the cost of RDRAM. RDRAM still costs up to twice as much as SDRAM.

RDRAM technology offers performance advantages and lower pin count than SDRAM and DDR SDRAM. However, SDRAM and DDR SDRAM offer more memory capacity and lower cost than RDRAM.

## Importance of using HP-certified memory modules in ProLiant servers

Customers should use only HP memory option kits when replacing or adding memory in ProLiant servers. This section describes two of the most important reasons.

First, not all DIMMs are created equal; they can vary greatly in quality and reliability. In the highly competitive memory market, some third-party memory resellers forego the level of qualification and testing needed for servers because it adds to the price of DIMMs. HP uses proprietary diagnostic tools and specialized server memory diagnostic tests that exceed industry-standards to ensure the highest level of performance and availability for ProLiant servers. The costs of system downtime, data loss, and reduced productivity caused by lower quality memory are far greater than the price difference between HP-certified memory and third-party DIMMs.

Second, use of HP memory option kits prevents improper mixing of single-rank and dual-rank DIMMs. Although single-rank and dual-rank DIMMs may have the same capacity, they differ in the way in which they are accessed by the chipset (see the "DIMM configurations" section). Therefore, to ensure that the server boots properly and operates reliably, single-rank and dual-rank DIMMs should not be used in the same bank. On the other hand, some ProLiant server platforms have configuration guidelines that allow the mixing of single-rank and dual-rank DIMMs. HP memory option kits precisely match the capabilities and requirements of the ProLiant server for which they are designated. Therefore, they prevent improper mixing of single-rank and dual-rank DIMMs.

HP memory option kits are listed in each server's user guide and in the product QuickSpecs available at [www.hp.com](http://www.hp.com).

## Conclusion

The increasing performance gap between processors and memory has generated development of several memory technologies. While some memory manufacturers prefer a revolutionary approach to memory technology development, others favor an open, evolutionary approach. Memory manufacturers must balance the cost of performance enhancements against the laws of physics and compatibility with existing technologies. HP will continue to evaluate relevant memory technologies in order to offer customers products with the most reliable, best performing memory at the lowest possible cost.

## For more information

For additional information, refer to the resources listed below.

Resource description	Web address
JEDEC Web site	<a href="http://www.jedec.org">http://www.jedec.org</a>
HP Advanced Memory Protection	<a href="http://h18004.www1.hp.com/products/servers/technology/whitepapers/adv-technology.html#mem">http://h18004.www1.hp.com/products/servers/technology/whitepapers/adv-technology.html#mem</a>
Fully-Buffered DIMM technology in HP ProLiant servers	<a href="http://h18004.www1.hp.com/products/servers/technology/whitepapers/adv-technology.html#mem">http://h18004.www1.hp.com/products/servers/technology/whitepapers/adv-technology.html#mem</a>

## Call to action

Send comments about this paper to [TechCom@HP.com](mailto:TechCom@HP.com).

© 2003, 2004, 2005, 2006, 2007, 2008 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice. The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.