



# **IBM Netfinity Server Memory Technology Directions**

## *Selecting a memory technology*

### **Executive Summary**

In the ever-changing computer industry, server memory technology continues to grow as chip designers strive to improve memory performance. Over the past 20 years, memory chips have evolved from expensive technological marvels to standardized, competitively-priced products widely available in the market place. For memory chip manufacturers to be successful in this market, their chips needed to be interchangeable. Chips from one manufacturer had to interchange with those of another in a transparent fashion. This allowed computer manufacturers to buy memory chips from any supplier who met their requirements for price, schedule and quantity. The result was a reduction in the cost of memory chips, which meant computers could be manufactured and sold at a lower price, directly benefiting consumers.

But such standardization has brought with it a reluctance by chip manufacturers to change memory chip architecture as they continually search for ways to extend the life of memory chips, protect their own investments and try to maintain razor-thin profit margins. Memory system designers are also affected any time memory chip architecture undergoes substantial change. The preferred architectural model so far has been one of continual but controlled evolution.

With the planned introduction of Direct Rambus™ memory technology later this year, many industry claims are being made about its technological achievements and benefits to end users. Indeed, Rambus memory represents something of a break with the memory technology that has evolved over the past 20 years.

Because of the large amounts of memory required in a typical server, memory cost is probably the leading factor in making a technology choice. In addition, computer manufacturers need to consider potential impacts to reliability, availability and serviceability (RAS) as they implement emerging memory technology into future systems.

This paper describes the evolution of memory technology. It explores some attributes of server memory systems and compares Rambus memory to synchronous DRAM (SDRAM). Given the current state of SDRAM and Rambus technologies and the marketplace conditions projected for both, IBM's server system design teams believe that SDRAM is the best and logical choice for server memory designs, at least for the next few years. Obviously, as technology changes and marketplace dynamics occur, this design choice could change in the future. This paper addresses the design choice for IBM server memory designs and is not a reflection of design and manufacturing plans for IBM's embedded, desktop or mobile systems; nor does it have any bearing on the present plans of IBM's Microelectronics Division to manufacture memory chips using the Direct Rambus DRAM interface. Because of the dynamic nature of this industry, changes may occur in IBM's actual implementation. IBM reserves the right to change specifications or other product information without notice.

## The Evolution of Memory

The earliest memory chips to achieve widespread use were page-mode chips. Page-mode chips used a multiplexed address bus that sent part of the memory address (the *row address*) to the chip first, then the remainder of the address (the *column address*) a short time later. Numerous enhancements to memory chips have since emerged—such as nibble mode, Extended Data Out (EDO), SDRAM and Direct Rambus. Yet remarkably, the original multiplexed-address concept still survives.

Over the years, the internal organization of memory chips has become more complex as internal banking structures emerged to support overlapping and interleaving of multiple memory operations. This complexity is a result of computer manufacturers trying to improve overall memory performance.

The current state of PC server memory technology can be summarized as follows:

- EDO memory is reaching end-of-life. Few, if any, new server systems will be provided with EDO main memory in the future.
- Because of its improved performance relative to EDO, 100MHz SDRAM memory (PC100® memory) is being used by most servers today. A 133MHz version of SDRAM (PC133 memory) is now available and is expected to begin showing up in servers before the end of 1999.
- The evolutionary model used in the commodity DRAM market suggests that the next generation of memory architecture will be Double Data Rate (DDR) SDRAM.
  - DDR architecture is similar to Single Data Rate (SDR) PC100 and PC133 memory—the main difference being that with DDR, data is transferred to and from memory on both edges of the data clock, effectively doubling memory bandwidth. With SDR architecture, data is transferred on only one clock edge.

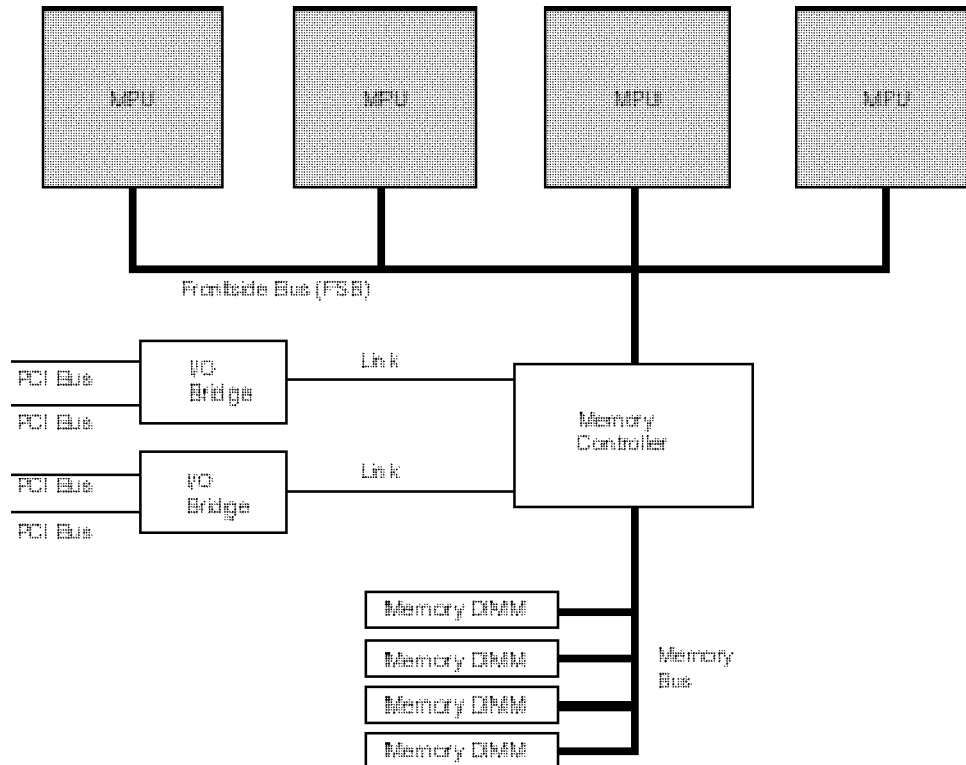
## A Look at Memory Architectures for Servers

Intel has announced it will move to Direct Rambus once 100MHz memory operation is achieved. Initially, Direct Rambus memory is expected to appear in desktop PCs, but because of Intel's position in the server industry, system designers are investigating Direct Rambus and trying to understand how to exploit its advantages in server memory systems in the future. The advantage most often attributed to Direct Rambus memory architecture is that it provides extremely high bandwidth per pin between the system memory controller and the memory itself, which allows the system designer a much wider latitude to increase memory bandwidth, or desired, to reduce pin count at the memory controller for cost savings. Server manufacturers should analyze the impact of such improved bandwidth and other characteristic in terms of performance, system reliability and cost.

### ***PC100 and PC133 SDRAM***

The core of any server is made up of memory (DRAMs), one or more microprocessors (sometimes called MPUs) and the chipset. The following diagram illustrates the internal structure of a typical server.

## Server Structure



The chipset in this illustration consists of the memory controller and two I/O bus interface chips that provide PCI buses. The type of memory used in any PC server is determined by the capabilities of the chipset that the manufacturer provides. Most chipsets for servers currently support PC100 SDRAM chips mounted on Dual Inline Memory Modules (DIMMs) and typically have an 8-byte memory interface. The memory bus width and frequency of these chipsets are usually intended to match the Frontside Bus (FSB) data rates of currently available Intel MPUs used in servers. Such a memory system provides data at a rate sufficient to keep up with the processors on the FSB.

Chipset designers are familiar with the characteristics of PC100 SDRAM, and they have optimized their chips to leverage the performance attributes of PC100 memory. While clock rates of MPUs have soared past 400MHz, the internal clock rates of the chipsets remain more closely aligned with memory and FSB frequencies, providing adequate performance. The closely matched clock frequencies of the FSB, memory controller and memory allow the chipset to access memory and return data to the processor efficiently.

Industry-standard PC133 SDRAM is an important bridge between PC100 SDRAM and future DDR SDRAM. With sample quantities available today, and with an architecture that is virtually identical to PC100 SDRAM, PC133 is expected to be incorporated into PC servers that ship in 1999. PC133 improves some specific memory timings and transfers data 33% faster between memory and the microprocessor without fundamentally changing the architecture of the memory bus. PC133 can provide a one-to-one speed match with an FSB running at up to 133MHz. This is another step in the evolution of memory technology for faster system performance. The internal architecture and clocking schemes of memory controller chips won't require radical change to take advantage of the extra performance. Plus they will maintain the efficiency of existing chipset architectures. The result is extra memory performance for the end user.

## **DDR SDRAM**

The next step beyond PC100 and PC133 SDRAM is DDR SDRAM. From a memory-chip manufacturer's perspective, PC100 and DDR are two products derived from a single core technology. The internal chip architecture of both products is virtually identical, but DDR is able to transfer two units of data per clock cycle compared to one unit for PC100 or PC133. DDR products operate at industry-standard 3.3v (initially) or 2.5v supply levels, and they employ the same four-bank internal architecture as PC100. A couple of differences in the interface facilitate the higher memory data bus speeds:

- The signal levels change from LVTTTL to SSTL 2.
- Bi-directional data strobe signals replace the memory clock to control data latching more precisely.

Memory packaging and bus topologies won't radically change with the introduction of DDR SDRAM. This will allow chipset manufacturers to incorporate DDR memory into their designs in a reasonably straightforward fashion. Because of increased memory bus frequency, however, very careful card wiring analysis and layout will be required. Compared to PC100 memory, DDR memory moves data to the memory controller twice as quickly; and for an equivalent memory bus width, it can support a processor FSB running up to twice the frequency. DDR memory will be available in 4-bit, 8-bit and 16-bit chip organizations.

DDR SDRAM is supported by many leading memory suppliers. That means it is likely to be widely available and competitively priced. Because of the large amount of memory typically present in servers, its price is a major component of the overall server's price.

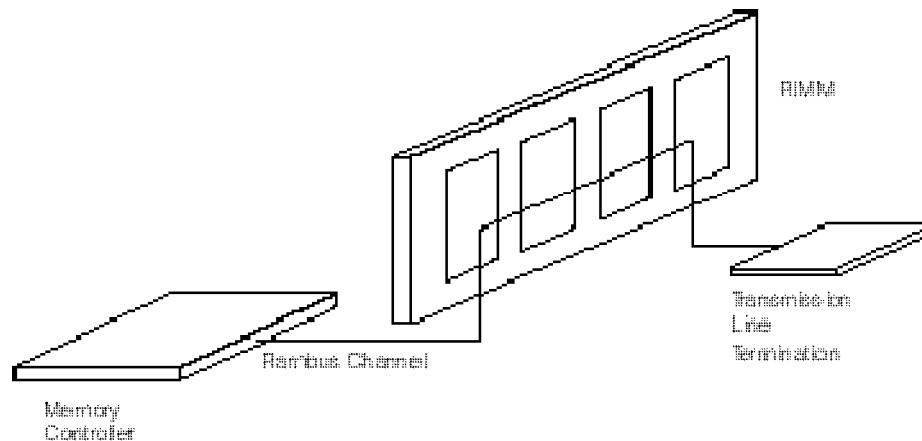
Various DDR speed sorts are expected to be available, initially at 200 and 266Mbps per pin, eventually leading to bit rates as high as 400Mbps per pin. Such an increase in speed will support a substantial increase in overall system performance as the faster chips become available.

## **Direct Rambus Channel**

Rambus, Inc. has developed a 30-signal memory interface called the *Direct Rambus Channel* memory which provides up to 800Mbps per pin data transfer rates between memory chips and the memory controller. Each Direct Rambus Channel is 16 or 18 bits wide and can transfer data at up to 1.6GBps (800Mbps x 2 bytes). This data rate is 8 times as fast as standard PC100 SDRAM. Multiple Direct Rambus Channels can be used in parallel to aggregate memory bandwidth. A Direct Rambus DRAM is also called *Direct RDRAM™*, or simply *RDRAM*. Up to 32 RDRAMs can be supported on a channel, and repeater chips are added to extend memory capacity if more memory is needed.

The RDRAM equivalent of the well-known DIMM is the *RIMM™*. A RIMM is similar in size to a DIMM, but not physically or functionally interchangeable. Up to 16 RDRAMs can be packaged on a RIMM, the width of the data bus is fixed at 16 or 18 bits, and each RDRAM on the RIMM is connected in parallel. To help achieve 800Mbps, the data bus daisy-chains through the RIMM (as shown in the following diagram) and is terminated at the end of the channel farthest from the memory controller. If fewer than the maximum number of RIMMs is present in the system, feed-through devices must be plugged into the unused sockets to maintain continuity.

## RIMM Signal Routing



Data transfers are organized into packets of 16 bytes. Consequently eight data transfers occur for each packet over a period of four clock cycles. (This takes 10 nanoseconds at the 800Mbps bit rate.) Internally the RDRAM chip is organized as a 128-bit wide data flow, allowing the entire 16-byte transfer to be generated from a single access.

Signals used to address and control operation of RDRAM memory are packetized and transmitted to the chips at the same frequency as data. Address and control packets consist of eight units of information spread across four clock cycles. Separate ROW and COL buses allow independent control of row and column functions.

Both SDRAM and RDRAM are designed with multiple internal banks to allow for pipelining of some memory operations. When a memory chip is accessed, a range of addresses becomes available for immediate accessing of data. If it is necessary to access data outside the active range, the memory requires a *precharge* operation before the next access can begin. Because the memory controller must wait for the precharge to complete before accessing part of its memory, latency is increased and bandwidth is reduced—which in turn reduces system performance. By organizing memory into multiple banks, the probability of accessing data from an already precharged section of memory is improved, and the precharge time can often be hidden while alternate banks are being accessed. Internally RDRAMs are organized into 16 quasi-independent banks, compared to four fully independent banks for SDRAM. The increased number of banks in the RDRAM chip is an advantage of RDRAM over SDRAM<sup>1</sup>.

### **RDRAM versus SDRAM Considerations**

With PC100 SDRAM firmly established as today's memory of choice for servers, PC133 SDRAM emerging, and DDR SDRAM as the next logical step in memory evolution, server manufacturers should consider four major factors in determining whether to switch to Direct RDRAM. They are:

- Memory bandwidth
- Memory latency

---

<sup>1</sup>The advantage is somewhat mitigated because the RDRAM banks share some internal circuitry between adjacent banks—the result of which is that a precharge penalty might be incurred on certain back-to-back operations involving adjacent banks.

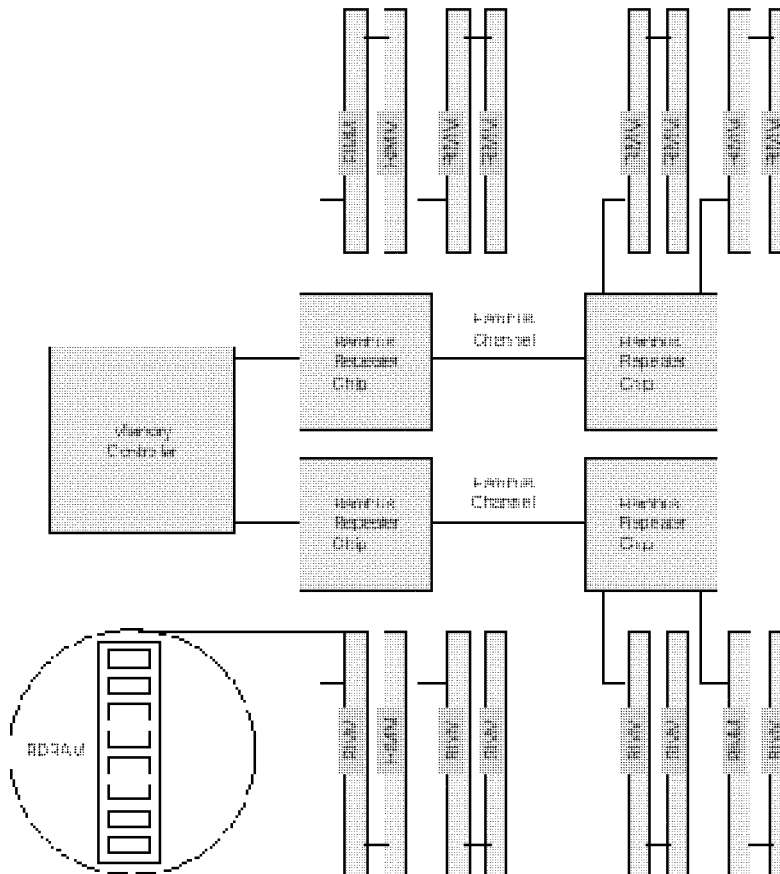
## Server memory architecture

- Reliability, availability and serviceability (RAS)
- Memory cost

### Memory Bandwidth

Memory bandwidth is one key component of memory performance. At 800Mbps per pin at the memory chip, Direct RDRAM is clearly the leader. It can claim up to four times the per-pin bandwidth of first-generation DDR SDRAM and six times the bandwidth of PC133 SDRAM. This advantage can lead to a reduction in the number of pins required on the memory controller, which could result in cost savings for that component. For desktop systems that have relatively small amounts of memory compared to that of servers, a few RIMMs can be placed near the memory controller to satisfy total memory capacity requirements. In a server application, memory capacity beyond 4GB is now the norm. With each Direct Rambus Channel having a limit of 32 RDRAMs on a bus, it is necessary to include additional components called *repeater chips* to extend memory capacity into the realm of the server. The following diagram illustrates the use of repeater chips.

### Rambus Repeater Chips



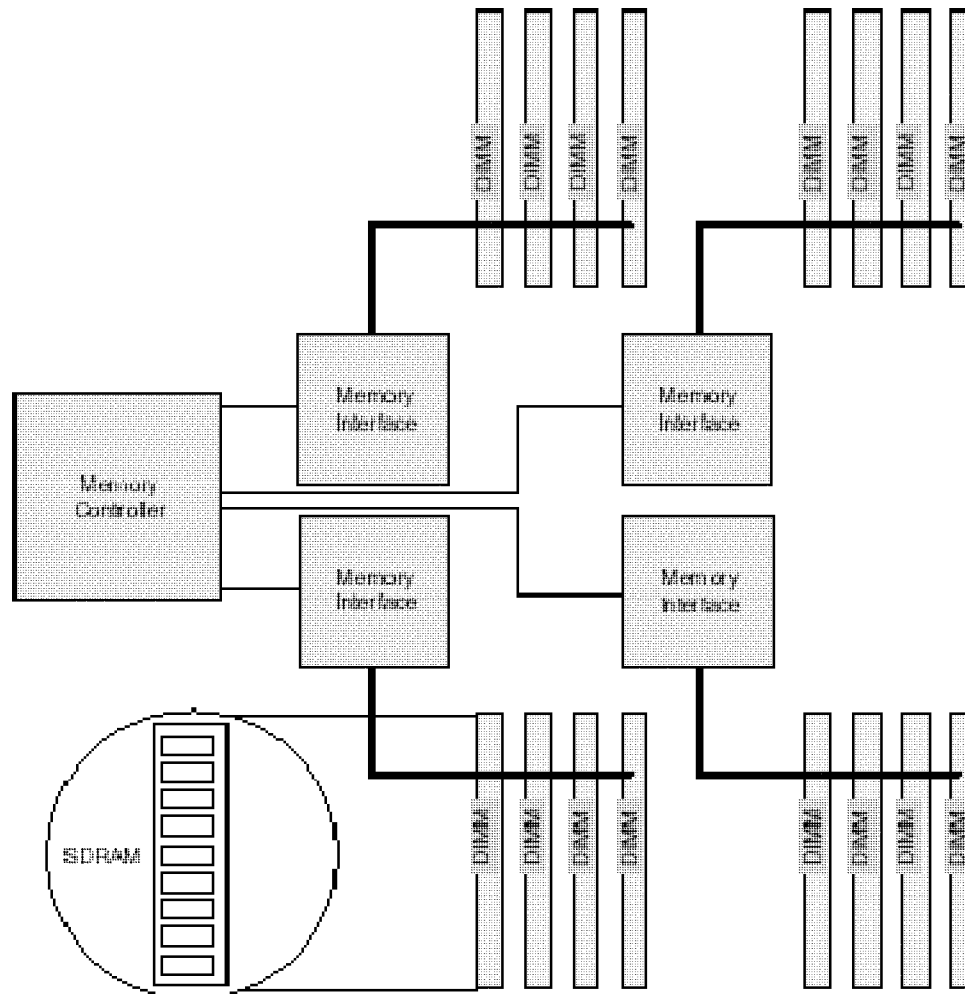
Using 128Mb memory technology, the effective memory capacity of a single RDRAM chip is 16MB. Thirty-two such chips yield 512MB of memory per Direct Rambus Channel before a

## Server memory architecture

repeater chip is required. Using two Rambus channels, four repeaters are required to arrive at a 4GB memory capacity.

Another type of support chip can be used with DDR SDRAM to improve the bandwidth per pin at the memory controller by using higher frequencies between the external support chips and the memory controller. This technique has been practiced for years in large servers and is called *interleaving*. Multiple memory banks are accessed in parallel, and data is combined into appropriate bus widths by the support chips and delivered to the memory controller at the desired frequency. The following diagram illustrates external support chips being used with DDR SDRAM.

### DDR SDRAM Memory System



DDR SDRAM operating at 100MHz (200Mbps per pin) can be interleaved four ways to provide data at 800Mbps per pin at the memory controller, the same as Direct RDRAM. Extremely careful signal analysis, very restrictive layout rules and tightly controlled transmission line impedances must be employed to achieve robust operation. In the case of Direct RDRAM, the entire memory structure needs to be designed to operate at maximum frequency. In the interleaved DDR SDRAM example, only the wires between the memory controller and the support chips need to operate at maximum frequency. Additionally, the Direct RDRAM wires are multidropped, up to 32 memory chips can be on the bus, while the wires between the memory

controller and DDR SDRAM support chips can be point-to-point, making the job of achieving maximum frequency operation somewhat easier. Thus the bandwidth of Direct RDRAM can, in theory, also be achieved in server memory controllers using interleaved SDRAM and high-performance support chips. As a practical matter, the pin count for the interface between the memory controller and the SDRAM support chips is likely to be somewhat higher than for the Direct Rambus architecture due to the typical need for multiple copies of clocks and control signals.

### **Memory Latency**

Another key component of memory performance is latency. Several factors determine the latency of the memory chip, including the internal organization of the memory chip itself and the intrinsic speed of the underlying memory technology. The two major latency categories are *initial latency* and *turnaround latency*. Initial latency measures the amount of time it takes to initiate access to the memory and read (or write) the first data. Turnaround latency is the amount of time needed to switch between different memory operations.

Initial latency of RDRAM and SDRAM is similar at the chip level, in part because the core storage area of the Direct RDRAM chip uses the same technology as SDRAM. Because of numerous speed sorts and technology variations, direct comparison of initial latency at the chip level is complex, and beyond the scope of this paper. However, one characteristic intrinsic to the Direct RDRAM design imposes an additional latency not present in SDRAM—packetization of ROW and COL address information. In SDRAM architecture, as with its predecessors, the ROW address, for example, is launched broadside at the memory chip. For Direct RDRAM memory, the ROW address requires four separate transfers to transmit the equivalent information. Although parameters such as overall wire length and electrical-signal swing levels do affect total propagation delay between the memory controller and the memory chip, the four separate transfers required by the Direct Rambus architecture could introduce some additional latency compared to the single transfer method for SDRAM.

As memory capacity is increased in the Direct Rambus system, the cascading of repeaters introduces more initial latency as the timings of all memory chips are programmed to match the latency of the farthest memory chip in the system. An analogous latency increase occurs for SDRAM memory used in systems which employ support chips such as were previously described. Signals must propagate through the support chips to reach the SDRAMs, which increases overall memory latency.

Compared with SDRAM, Direct RDRAM provides better turnaround latency between reads and writes at the memory bus level. That is, fewer cycles are lost between read and write operations. This can result in a performance increase, especially in lower cost memory controllers likely to be used in desktop systems. However, read-write turnaround latency penalties can be mitigated. High-performance server memory controllers sometimes provide buffer space to accumulate data which is to be written to memory, and memory writes are delayed until the memory system is not performing read accesses (or the buffers fill up).

The advantage of this system is that read requests received by the memory controller can be performed ahead of write operations, as long as the correct order of reads and writes appears to be maintained by the memory controller. If the write activity is correctly sequenced, the reads will be batched separately from the writes—resulting in fewer read-write turnaround latency penalties. This requires additional logic complexity, of course, and there will always be some turnaround penalty to be paid.

The latency of the memory chip and its associated interface is only part of the overall memory latency. In symmetric multiprocessing (SMP) servers, data is frequently provided to one processor from the cache of another processor. When that happens, the effective memory



latency is at least partially a function of processor cache design and FSB protocol. As processor cache sizes and the number of processors in the system increase, the probability of getting data from another MPU increases. Servers tend to have very large caches, and the effect is significant in 4-way and 8-way SMPs.

As very-high-speed memory interfaces like the Direct Rambus Channel are implemented, memory controllers must be carefully designed to avoid latency penalties for serialization and deserialization of data as it crosses frequency domains. Although companies that build microprocessors expend enormous custom-circuit-design effort to enable microprocessors to operate at 400MHz and above, some chipset developers cannot afford such investment for memory controllers. Consequently chipsets often operate at frequencies more in line with FSB frequencies, currently in the range of 100MHz. With the Direct Rambus Channel operating at essentially an 800MHz data rate that enables a relatively narrow data-width between the memory controller and memory, some serialization and deserialization of data occurs. For example, four data transfers from the Direct Rambus Channel might be accumulated to provide an 8-byte memory word, moving about inside the memory controller chip at one-quarter the frequency of the Rambus channel. Assembling that word from the individual subunits transferred across the Rambus channel incurs a latency penalty often referred to as a *trailing-edge effect*, which means the memory controller is waiting for the last piece of data to arrive. A good design minimizes this penalty, but it is almost impossible to eliminate completely.

To summarize, effective memory latency for DDR SDRAM and RDRAM is very system dependent. At the chip level, RDRAM is better than SDRAM in some aspects but not in others. The effective latency is a function not only of the memory chip, but also of the memory controller and the intervening wires. In SMPs, even the caches in other processors can significantly affect apparent memory latency. Users should be wary of performance claims based on minor differences at the memory chip level. As of today, the difference in memory latency between RDRAM and SDRAM chips is not a strong reason to choose one over the other. What matters most is how effectively the server vendor has optimized the overall performance of the memory system, and IBM is applying its extensive experience in server design to its family of Netfinity servers.

### ***Reliability, Availability and Serviceability***

Reliability, availability and serviceability (RAS) are cornerstones of a server memory system. Although a powerful microprocessor can be implemented in a single chip, memory often requires dozens of chips, and brings with it a correspondingly greater potential for failure. Fortunately memory can be made extremely reliable through Error Checking and Correction (ECC) codes, built-in redundancy or sparing techniques and concurrent repair (hot plug). The key to high memory availability is a robust ECC algorithm capable of detecting and correcting nearly all possible memory failures. Early mainframes and high-end servers provided Single Bit Error (SBE) correcting codes, but recently some servers have begun using codes that provide multibit error detection and correction. In ECC terminology, the technical name for the correctable entity is a *symbol*. A symbol can be any width. Current symbol sizes are up to 4 bits, and larger symbols are possible.

Typically, codes can correct a single-symbol error and detect all double-symbol errors. It is advantageous to be able to correct symbols that are as wide as the memory chip, because the system can remain fully operational should the entire memory chip fail. However, as the symbol size increases, the number of extra memory bits necessary for containing the ECC code increases dramatically. For example a server with a typical 8-byte-wide memory data path requires 8 checkbits to provide Single Symbol Correct/Double Symbol Detect (SSC/DSD) for a 1-bit symbol, 10 checkbits for a 2-bit symbol, 14 checkbits for a 4-bit symbol, 24 checkbits for an 8-bit symbol and 48 checkbits for a 16-bit symbol.

Conversely, as the overall size of the data in the ECC word increases, the overhead is reduced accordingly for a given symbol size. A 16-byte ECC word requires 16 checkbits for 4-bit symbol SSC/DSD (12.5% overhead), compared to 14 checkbits for 4-bit symbol SSC/DSD (21.9% overhead) on an 8-byte-wide data path, a 43% reduction in checkbit overhead per byte of data. The ideal combination to minimize the overhead for ECC is a small symbol and a large ECC word.

A basic 1-bit SSC/DSD ECC code can handle the occasional soft memory error or single-cell hard fail, but a significant portion of the overall memory failure rate can consist of multiple bit failures—the most catastrophic occurrence being when the entire memory chip ceases to function. Although 8-bit-wide memory chips are dominant in consumer PC systems, the vast majority of SDRAMs used in servers today are 4-bit-wide chips, for which it is entirely practical to provide ECC coverage for a total chip failure with modest checkbit overhead. In fact, the IBM Netfinity 7000 M10 was the first Intel processor-based server to offer Chipkill™ Memory. For more information see the white paper "IBM Chipkill Memory," available from our Web site at [www.ibm.com/netfinity](http://www.ibm.com/netfinity).

Desktop systems will typically not provide Chipkill memory error-correcting capabilities. But high-end, business-critical servers are driven to make memory extremely reliable and, although still not common in today's servers, Chipkill ECC is likely to become a strong requirement. The current Direct Rambus RDRAM implementation, with its 16-bit-wide format makes Chipkill error-correcting capabilities more difficult to provide, compared to 4-bit-wide or 8-bit-wide SDRAM chips.

Recognizing this, Rambus recently announced that their 256Mb memory chips will support a feature called *Interleaved Data Mode* (IDM) when these chips eventually become available. (IDM will not be available for 128Mb chips.) IDM will enable each RDRAM to read and write its data on two of the 16-bit positions in the Rambus channel, compared to the normal mode of operation where a single chip drives all 16 data wires. Rather than a single chip providing 16 bytes of data per packet, 8 chips per channel are selected simultaneously. Each chip is programmed to bring out only two bytes of data and uses two of the 16 available data wires in the Rambus channel. Thus, 8 chips must be selected, where previously only a single chip needed to be accessed. This new feature will reduce the number of RDRAMs required to implement chipkill correction compared to the current RDRAM design and providing similar overhead to SDRAM.

This approach can be used with different numbers of Rambus channels by trading the number of channels for access latency. For example, a system with three Rambus channels will experience a greater access latency than a system with nine channels, because the three-channel system will take longer to accumulate enough data to comprise a complete ECC word and data cannot be forwarded to the processor until that data can be checked for errors and corrected, if necessary. Due to the nature of the IDM design, where multiple ECC words are simultaneously (but serially) accumulated, data will not be available for use until the entire memory block transfer has been completed.

In contrast, SDRAM-based systems with external support chips can access the critical data for an entire ECC word in a single data transfer, permitting the processor to resume operation more quickly.

### **Memory Cost**

Among the factors that influence the decision to use a particular memory, cost is usually the first. In the past few could have predicted how rapidly and how far memory prices would decrease. Even so memory constitutes a significant portion of the cost of a PC server. Servers simply need large amounts of memory to function efficiently. As more customers move their

business-critical applications to Intel-based servers, the amounts of memory they need will continue to expand.

Intrinsically, the cost advantage lies with SDRAM for several reasons. The first is that for equivalent capacity, the die size of a Direct Rambus memory chip is somewhat larger than for SDRAM. Larger die size means fewer chips per wafer and lower yields. The end result is higher manufacturing cost. Increased costs might be attributed to fees manufacturers have in using the RDRAM specifications. The SDRAM specifications are an open standard. Beyond the memory chip itself, sophisticated module and card packages are required to maintain electrical signal integrity, along with thermal enhancements (such as heat spreaders) to dissipate increased power. Finally, chip manufacturers will likely have to invest in new test equipment to support RDRAM manufacturing.

Another major factor in memory cost is the size of the marketplace. The more chips the memory manufacturers can produce, the lower their unit costs. Direct RDRAM might well capture a major share of the desktop memory component marketplace, but market projections for SDRAM volumes driven by server requirements and other uses should be sufficient to keep SDRAMs cost competitive.

## **Conclusion**

Direct Rambus RDRAM is receiving significant publicity as Intel continues to invest in manufacturing capability for it and moves to introduce it into desktop systems in the near future. The question of whether or not server manufacturers will adopt Direct Rambus RDRAM naturally follows from these activities.

From a server RAS perspective, RDRAM introduces new challenges that need to be addressed to provide enhanced ECC features like chipkill correction. The 16-bit or 18-bit width of the current devices makes chipkill ECC coverage more costly and difficult than for the 4-bit or 8-bit wide SDRAM chips. The new Interleave Data Mode reduces this chip overhead, but at the cost of access latency.

Because of the large amounts of memory required in a typical server, memory cost is a significant factor in making a technology choice. Direct RDRAM memory chips are intrinsically larger than comparable SDRAM chips, and vendors may need to factor in additional fees and manufacturing costs.

Today's PC100 and PC133 SDRAM have the edge in server memory volumes, and with the emergence of DDR, SDRAM appears to be well-suited to provide the performance, RAS and cost attributes needed for servers.

## **Additional Information**

For more information on IBM Netfinity direction, products and services, refer to the following white papers, available from our Web site at [www.ibm.com/netfinity](http://www.ibm.com/netfinity).

*Achieving Remote Access using Microsoft Virtual Private Networking*  
*At your service...Differentiation beyond technology*  
*Capacity Planning for IBM Netfinity Servers Running Microsoft Windows NT Server 4.0, Terminal Server Edition*  
*Examples Implementing IBM Netfinity Server Management: Air Conditioning Failure*  
*IBM Chipkill Memory*  
*IBM Netfinity 5500 Server Family*  
*IBM Netfinity 5500 Server Family*  
*IBM Netfinity 7000 M10 Server*  
*IBM Netfinity 8500R Overview*  
*IBM Netfinity Advanced Systems Management*  
*IBM Netfinity Availability Extensions for Microsoft Cluster Server*  
*IBM Netfinity Backup Automation*  
*IBM Netfinity Cluster Directions*  
*IBM Netfinity ClusterProven Program*  
*IBM Netfinity Enterprise Storage Solutions*  
*IBM Netfinity ESCON Adapter*  
*IBM Netfinity Fibre Channel Solutions for Enterprise Storage*  
*IBM Netfinity Manager 5.2*  
*IBM Netfinity Manager Plus for Tivoli Enterprise Overview*  
*IBM Netfinity OnForever Initiative*  
*IBM Netfinity PCI Hot-Plug Solutions*  
*IBM Netfinity Predictive Failure Analysis*  
*IBM Netfinity RAID Technology*  
*IBM Netfinity Server Quality*  
*IBM Netfinity Server Ultra2 SCSI Directions*  
*IBM Netfinity Storage Area Networks*  
*IBM Netfinity Storage Management Using Tape Subsystems*  
*IBM Netfinity Systems Management for Servers*  
*IBM Netfinity Web Server Accelerator*  
*IBM Netfinity X-architecture*  
*IBM ServerGuide for Netfinity and PC Server Systems*  
*Implementing Microsoft IIS on Netfinity 5500 M10*  
*Integrating IBM Netfinity Manager with Intel LANDesk Server Manager*  
*Integrating IBM Netfinity Manager with Microsoft Systems Management Server*  
*Lotus Domino Clusters Installation Primer*  
*Lotus Domino Clusters Overview*

## *Server memory architecture*



© International Business Machines Corporation 1999

IBM Personal Computer Company  
3039 Cornwallis Road  
Department. LO6A  
Research Triangle Park, NC 27709

Printed in the United States of America  
9-99  
All rights reserved

References in this publication to IBM products or services do not imply that IBM intends to make them available in all countries in which IBM operates.

IBM reserves the right to change specifications or other product information without notice.

IBM Netfinity servers and PC Servers are assembled in the U.S., Great Britain, Japan, Australia and Brazil and are comprised of U.S. and non-U.S. parts.

Are you Year 2000 ready? Visit [www.ibm.com/pc/year2000](http://www.ibm.com/pc/year2000) or call 1 800 426-3395 (and request document number 10020 from our faxback database) for the latest information.

IBM, Chipkill, Netfinity and PC100 are trademarks of International Business Machines Corporation in the United States and/or other countries.

Direct Rambus, Rambus, RDRAM, RIMM are trademarks or registered trademarks of Rambus, Inc. Other company, product and service names may be trademarks or service marks of other companies.

THIS PUBLICATION MAY INCLUDE  
TYPOGRAPHICAL ERRORS AND TECHNICAL  
INACCURACIES. THE CONTENT IS PROVIDED  
AS IS, WITHOUT EXPRESS OR IMPLIED  
WARRANTIES OF ANY KIND.