



**ARCHITECTURAL
OVERVIEW**

Document No. 60291

© Copyright 1994 NEC Electronics Inc. All rights reserved.

No part of this document may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means without the prior written permission of Rambus Inc.

Rambus, RDRAM, RSocket, RTransceiver, RModule and the Rambus Logo are trademarks of Rambus Inc.

Table of Contents

Introduction.....	4
The Physical Layer	6
The Logical Layer	8
Inside a NEC Rambus DRAM (RDRAM)	10
Attaching Masters and Slaves to the Rambus Channel	13
Packaging for Speed.....	16
System Packaging.....	17
Main Memory Application	18
Bit-Mapped Graphics in a Rambus World	20
The Advantages of Rambus Technology	22
Applying the Advantages of Rambus Technology.....	23

Preface

NEC and Rambus have developed a revolutionary new technology that is used to build high-performance, cost-effective memory subsystems. These memory subsystems are ideal for a broad range of systems—from consumer digital video products to desktop computers to supercomputers.

This *Architectural Overview* introduces NEC Rambus technology and the devices that implement it. These topics are presented in a series of sections. Each sec-

tion describes related aspects of NEC Rambus technology and its application.

The *Introduction* lists the architectural limitations that Rambus set out to solve and summarizes the resulting solution. *The Physical Layer* and *The Logical Layer* define, respectively, the electrical interface and the protocol used between devices on an NEC Rambus channel. *Inside an NEC Rambus DRAM* describes the internals of a DRAM constructed using NEC Rambus technology. *Attaching Masters and Slaves to the NEC Rambus Channel* presents the ways in which devices can connect using the NEC Rambus channel. *Packaging for Speed* and *System Packaging* cover the physical packaging of NEC Rambus devices. *Main Memory Application* and *Bit-Mapped Graphics in a Rambus World* consider two key applications of NEC Rambus technology. And finally, *The Advantages of NEC Rambus Technology* and *Applying the Advantages of NEC Rambus Technology* list the key points discussed throughout this *Architectural Overview*.

Technical specifications, user guides, data sheets and application notes are also available from both NEC and Rambus Inc. These provide further information about NEC Rambus technology, how to implement systems that use it, and individual NEC Rambus-compatible devices.

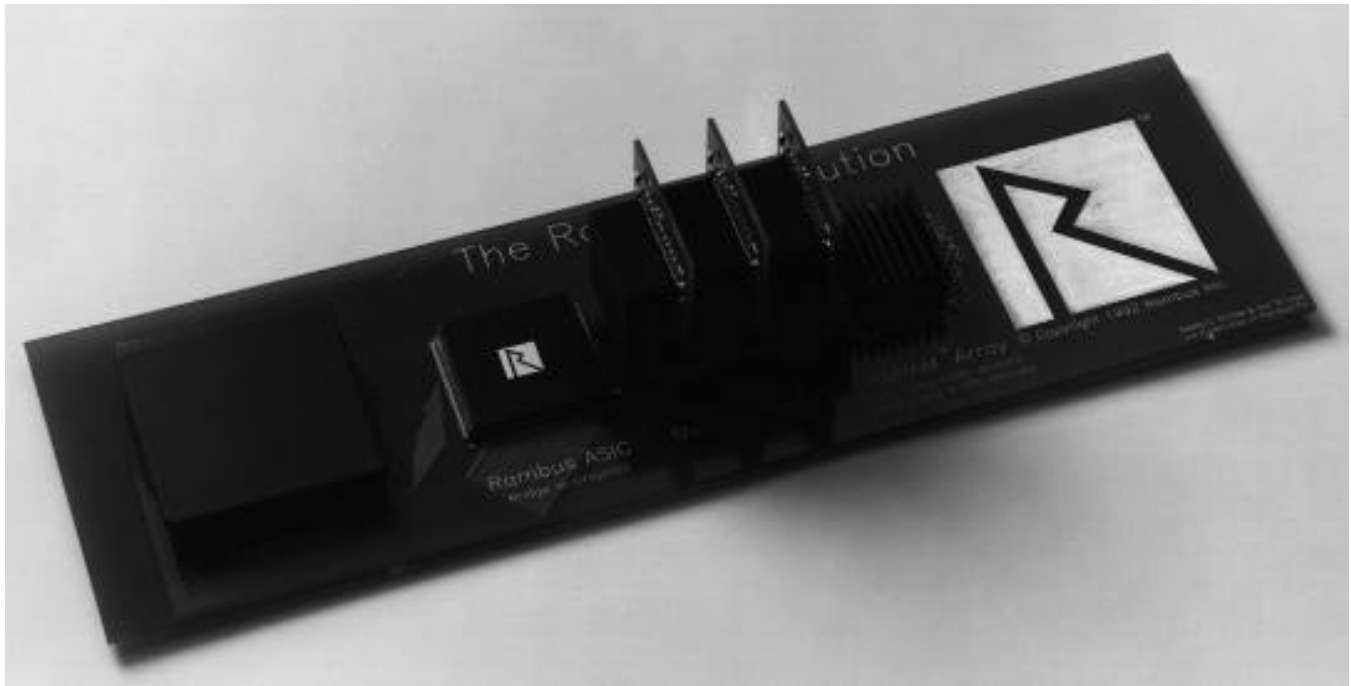


Figure 1: An NEC Rambus-Based System (Two-Thirds Actual Size)

Introduction

During the last two decades, DRAM technology has progressed dramatically. Device densities have increased from 1 Kbit per chip to 256 Mbits per chip, a factor of 256,000. DRAM performance has not kept pace with these density changes, since access times have only increased by a factor of 10. Over the same 20-year period, microprocessor performance has jumped by two orders of magnitude. This growing disparity between the speed of microprocessors and that of DRAMs has forced system designers to create complicated and expensive hierarchical memory systems using SRAM caches and DRAM arrays. In addition, now that users demand high-resolution graphics, systems often rely on expensive VRAM frame buffers to provide the necessary bandwidth.

To address this processor to memory performance gap, NEC and Rambus Inc. have developed a revolutionary chip-to-chip bus, the NEC Rambus™ Channel, that operates up to 10 times faster than conventional DRAMs. The NEC Rambus Channel directly connects memories to computational engines and interface devices such as microprocessors, digital signal processors, graphics processors, and RAMDACs. The channel uses a small number of very high-speed signals to carry all address, data, and control information. Because it is able to transfer data at 500 MBytes per second at a moderate cost, the NEC Rambus Channel is ideal for high-performance/low-cost systems.

The NEC Rambus solution eliminates external buffers and controllers while assuring a modular and scalable system solution. Each NEC Rambus DRAM, called an NEC RDRAM™, transfers data at 500 MHz over the NEC Rambus Channel. Multiple channels can be used in parallel to achieve even higher throughput.

The NEC Rambus Channel is implemented using standard PC board layout and manufacturing techniques. Devices connecting to the channel contain dedicated circuitry, the NEC Rambus interface, that can be produced on standard sub-micron CMOS processes.

Electrically, the NEC Rambus channel relies on controlled impedance-terminated transmission lines. These lines carry low-voltage-swing signals. Clock signals and data always travel in parallel to virtually eliminate clock-to-data skew.

NEC and Rambus Inc. has assured device independence now and in the future by defining a high-level protocol that moves data in blocks and by using a large 36-bit address space. Designing new generations of hardware is simplified since the signals comprising the channel will not change from generation to generation. And, since systems increasingly move data in multi-byte chunks, a block-oriented protocol makes efficient use of bus bandwidth.

The NEC Rambus Solution: the NEC Rambus Interface, the NEC Rambus Channel, and the NEC RDRAM

The NEC Rambus solution replaces costly memory subsystems and interconnect with a single, standard, high-performance, chip-to-chip bus and revolutionary DRAMs. This solution has three main elements: the NEC Rambus Interface, the NEC Rambus Channel, and the NEC RDRAM. The NEC Rambus Interface is implemented on both master and slave devices on the channel.

Masters contain intelligence and are the only devices that generate requests. NEC Rambus masters can be conventional microprocessors, peripheral chips, ASIC devices, memory controllers, or graphics engines that have an NEC Rambus Interface.

Slaves are devices that only respond to requests and therefore require a low level of intelligence. As a result, the memory die size overhead is minimized to keep them highly cost effective.

The most important slave is an NEC RDRAM, which is a CMOS DRAM incorporating unique architectural modifications and the NEC Rambus Interface circuitry. The NEC RDRAM is arranged as 2M x 8 or 2M x 9. The definition and use of the nine bits is left to the system designer.

The NEC Rambus Channel is revolutionary in that it is only 9 data bits wide but capable of transferring data at rates up to 500 MBytes per second from a single NEC RDRAM. By contrast, today's fastest DRAMs are ten times slower. To approach the performance of a single NEC RDRAM, a traditional memory system would require a wide, complex, interleaved bus using a large number of traditional DRAMs.

In addition, the NEC Rambus Channel is a system-level specification. Thus, systems containing NEC Rambus Channel(s) can operate at full rated speed.

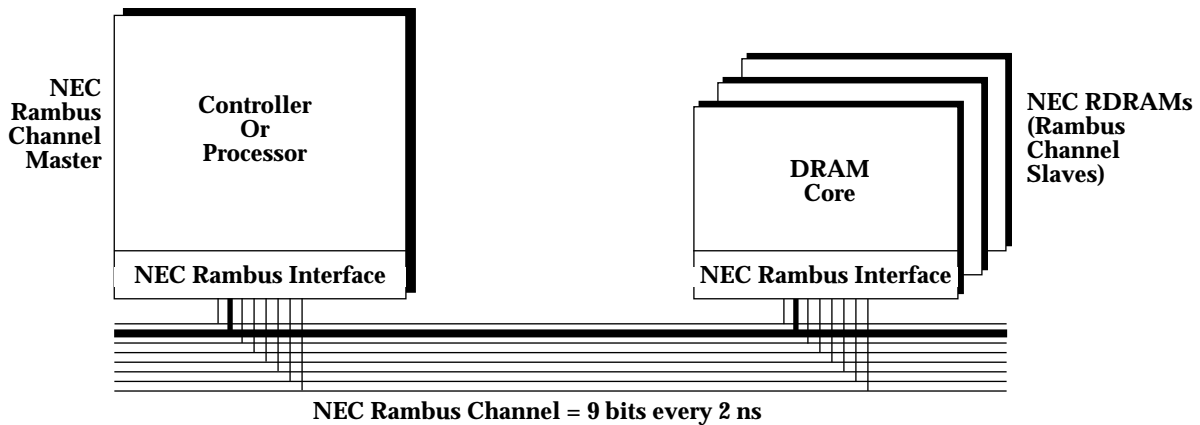


Figure 2: Logical View of An NEC Rambus-Based System

This is not the case with systems using traditional DRAMs, because component specifications such as setup times, hold times, and so forth for DRAMs, ASICs and buffers combine to reduce effective cycle times. Also, when DRAM control signals are generated from a master clock, the timing of these signals is often degraded, further slowing the system.

The NEC Rambus Channel has a well defined mechanical interface. Masters and slaves connect to the PCB with an interface that has only 15 active signals. The NEC RDRAM package itself has 32 pins, while master devices use 28 pins to connect to the NEC Rambus Channel. The package used on the NEC RDRAM is a vertical-surface-mount package, SVP, which is an EIAJ standard. Memory subsystem expansion to hundreds of NEC RDRAMs is accommodated through the use of a custom Rambus Socket (RSocket™).

Each NEC Rambus master and slave has its own NEC Rambus Interface, which is currently available as an ASIC cell. This interface converts from the low-swing-voltage levels used by the NEC Rambus Channel to ordinary CMOS logic levels.

All critical, high-speed system design issues that arise while designing masters, NEC RDRAMs and the NEC Rambus Channel have been resolved for the designer. Designers can implement an NEC Rambus system using the company's step-by-step pre-engineered documentation.

The Physical Layer

The NEC Rambus Channel achieves its high speed with dense packaging, high-quality transmission lines, low-voltage signaling, and precise clocking. By employing these techniques on conventional

CMOS IC and printed circuit board processes, NEC Rambus technology achieves high performance at low cost.

Dense slave packaging is essential. Slaves use a vertical-surface-mount package, the details of which appear later in this document. This package allows NEC RDRAMs to be spaced 100 mils apart, densely packing the memory subsystem while keeping the underlying wires short.

This package has a number of electrical advantages. Dense packing means the printed circuit board traces are uniformly loaded without discontinuities. Also, the vertical package aligns the die pads with the PCB traces. The alignment keeps the internal package leads equally short. This equalizes and minimizes the parasitic capacitance and stub inductance that loads the traces. As a result, the wires of an NEC Rambus Channel are matched, high-quality transmission lines.

Signaling

An NEC Rambus Channel contains 13 such controlled impedance, matched transmission lines:

- ClkToMaster
- ClkFromMaster
- BusData [8:0]
- BusEnable
- BusCtrl

These high-speed signals are terminated in their characteristic impedance. As shown in Figure 3, the NEC Rambus Channel has a bus topology with the master at one end, terminators at the other end, and the slaves in between.

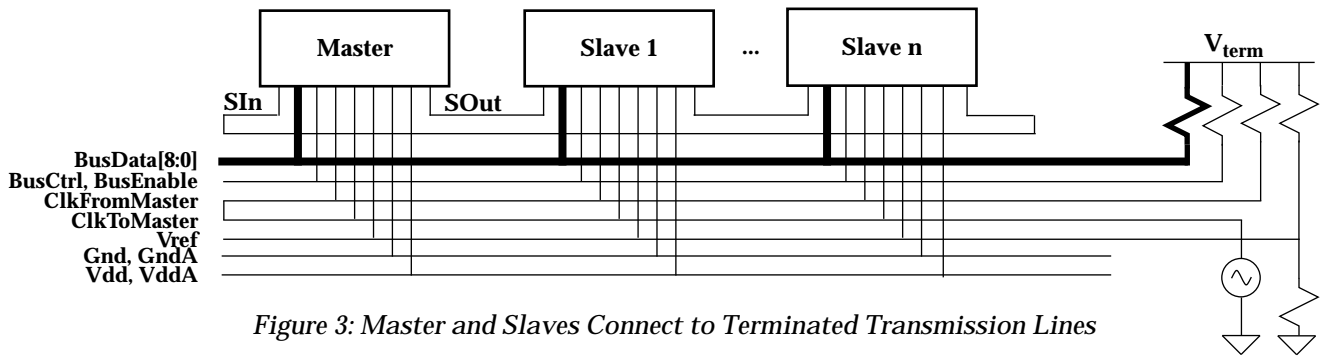


Figure 3: Master and Slaves Connect to Terminated Transmission Lines

The terminators pull the signals up to the system-supplied V_{term} voltage, which corresponds to logic 0. A master or slave asserts a logic 1 by sinking current from the wire, typically with an open-drain NMOS transistor.

All high-speed signals on the NEC Rambus Channel use low voltage swings of about 600 mV. Figure 4 shows the nominal voltages of V_{term} , the DC reference, V_{ref} , and the logic 1 level, V_{OL} . Within limits, all these voltage levels can be set by the system designer to control power consumption and noise margin. Figure 3 shows that V_{ref} may be generated with a resistive divider.

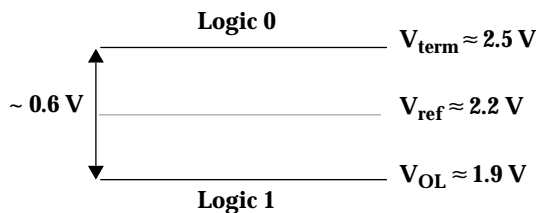


Figure 4: High-Speed Signals Use Low - Voltage Swings About V_{ref}

V_{ref} sets the logic threshold for the high-speed, low-swing signals. This provides immunity from common mode noise on the channel. As shown in Figure 3, V_{ref} connects to each device. All devices receive the low-swing signals with differential input circuits and use V_{ref} to set the logic threshold.

This differential sensing allows the channel to use a low-voltage swing. Low-voltage-swing signals minimize dv/dt and thus di/dt to provide the following advantages:

- Reduced ground bounce
- Reduced power consumption
- Reduced electromagnetic interference
- Compatibility with 3.3V devices

Clocking

The NEC Rambus Channel is synchronous, meaning that all data transfers are referenced to clock edges. At NEC Rambus frequencies, special care must be taken to minimize clock-to-data skew. At the physical level, data means BusData, BusCtrl, or BusEnable and clock means ClkFromMaster or ClkToMaster.

Figure 3 shows the clock distribution. The clock source can be a separate oscillator as shown or can be generated in the master. The clock begins at the slave end of the channel and propagates to the master end as ClkToMaster, where it loops back as ClkFromMaster to the slave end and terminates.

Clock and data travel in parallel to minimize skew. So, a slave sends data to the master synchronously with ClkToMaster, and the master sends data to the slaves synchronously with ClkFromMaster. Because the transmission lines are matched, the clock and data signals remain synchronized as they continue to their destination.

Data Transfer

Data transfers occur only between master and slave, and never directly between slaves. Thus, signals may be terminated at one end of the channel (Figure 3). Data driven by the master propagates past all slaves with the desired voltage swing. So, all slaves correctly sense data driven by the master. The matched terminator prevents any reflections.

Data driven by a slave moves in both directions at one-half the desired voltage swing. The NEC Rambus protocol guarantees that none of the other slaves expect to see valid data on the channel during this time. At the master end, the half-swing pulse reflects off the open end of the wire, doubling in amplitude. Superposition of waveforms and matched termination at the slave end mean that slave-to-master data transfers take place at full speed and full amplitude.

The quality of these signals is shown, as actual oscilloscope waveforms, in Figure 5.

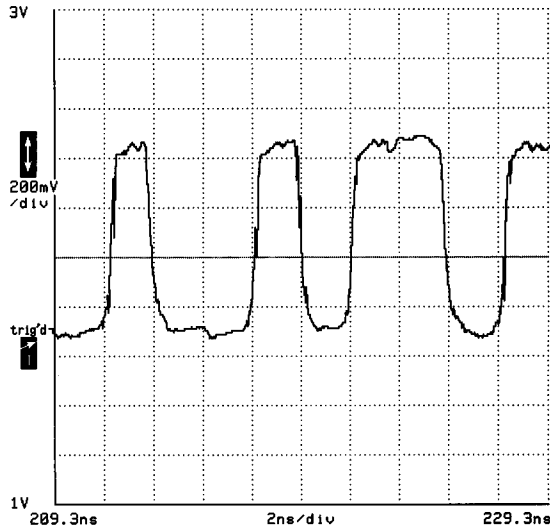


Figure 5: The Use of Terminated, Controlled Impedance Signal Lines Ensures High Signal Quality

Data is effectively transferred on both edges of a 250MHz clock, resulting in a 500-Mbits-per-second, per-wire transfer rate. Each data transfer uses a 2 nanosecond interval, with two of these intervals per clock period. The two clock edges provide a natural way to label intervals as even and odd (Figure 6). Even intervals occur during clock falling edges and odd intervals during rising edges, with each clock edge at the midpoint of data. Request, acknowledge, and data packets begin on even intervals.

Presently, the bus length is limited to slightly less than one interval time. At 500-Mbits-per-second, per-wire, this corresponds to a maximum of 32 NEC RDRAMs.

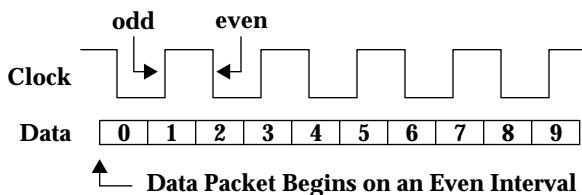


Figure 6: Data Cycle Definitions and Their Relationship to Clock

Signals

In each device, the NEC Rambus Interface has 15 active pins. Including power supplies, the total is 28 pins:

- BusData [8:0], BusCtrl, BusEnable, ClkToMaster, and ClkFromMaster correspond to the 13 low-voltage-swing pins

- SIn and SOut are daisy-chained TTL signals used during device initialization
- Vref, Gnd, GndA, Vdd and VddA are a reference voltage, six ground and three power pins
- Two isolation pins
- One reserved pin

GndA and VddA are separate supply voltages for the phase-locked-loop circuits. The isolation pins may optionally be used for power, and additional pins may be added beyond the main 28. For example, the NEC RDRAM leaves the isolation pins unconnected and adds four more pins to power the DRAM core, for a total of 32 pins.

The Logical Layer

Data on the NEC Rambus Channel moves in blocks. The size of these block-oriented transfers can be optimized to match the needs of each particular system. To implement this, the NEC Rambus Channel defines a protocol with the following three types of packets:

- Request
- Acknowledge
- Data

The combination of a request, an acknowledge, and a data packet constitute a transaction. The following types of transactions are defined:

- Read memory space
- Write memory space
- Read register space
- Write register space
- Broadcast write register space

To ensure proper synchronization of all devices connected to the NEC Rambus Channel, request, acknowledge and data packets all begin during even intervals (falling clock edge).

Request Packet

A request is issued by the master. Each request is, as depicted in Figure 7, six intervals long. There are ten bits per interval. A request contains:

- A start bit (Start)
- An opcode (Op[3:0])
- An address (Adr[35:0])
- A count (Count[7:0])

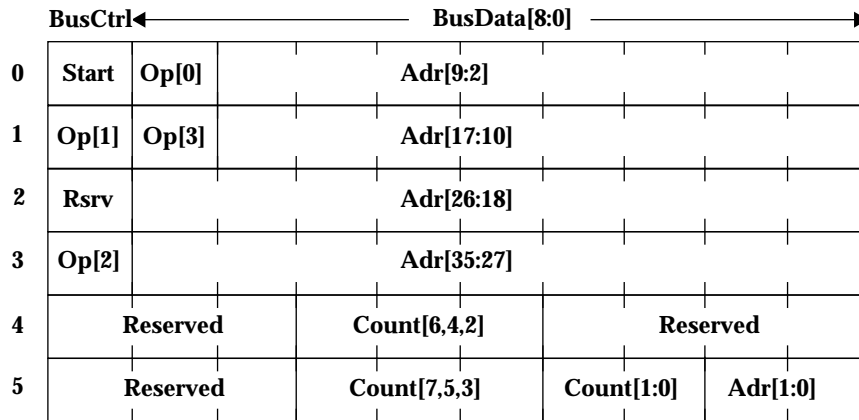


Figure 7: Request Packet Format

The opcode specifies the type of data transfer to take place. Read and write operations are defined for both memory and register spaces. Each slave device contains configuration registers (in the register space) in addition to memory space.

The 36-bit address specifies the first byte that is transferred. From 1 to 256 bytes can be moved with a single transaction, as specified by the 8-bit count field.

Acknowledge Packet

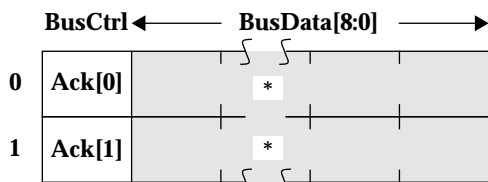
Upon receipt of a request, the addressed slave responds with an acknowledge. As shown in Figure 8, an acknowledge is sent across the BusCtrl wire to the master. This may be concurrent with a data packet.

Coding of the acknowledge packet is given in Table 1.

Table 1: Acknowledge Packet Encoding

Ack[1:0]	Definition
00	Addressed slave does not exist
01	Okay; slave will respond to request
10	Nack; slave busy, try request later
11	Reserved

With the exception of broadcast writes, a transaction addresses only a single slave. Thus, slaves never arbitrate for use of the NEC Rambus Channel.



* May contain a portion of a Data Packet

Figure 8: Acknowledge Packet Format

Data Packet

Data packets contain 1 to 256 nine-bit data bytes. Figure 9 depicts the format of a data packet.

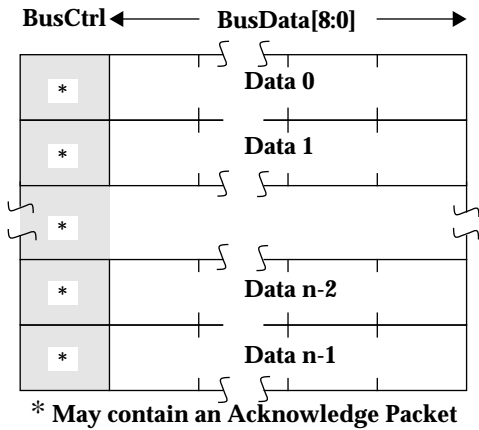


Figure 9: Data Packet Format

Read Transaction

The format of a read transaction is shown in Figure 10. After a request packet is issued, an acknowledge packet is returned a time **AckDelay** later. If the acknowledge is Okay, the read data packet is returned a time **ReadDelay** after the request packet. Both of these delay values are programmed into the configuration registers of all devices during system initialization.

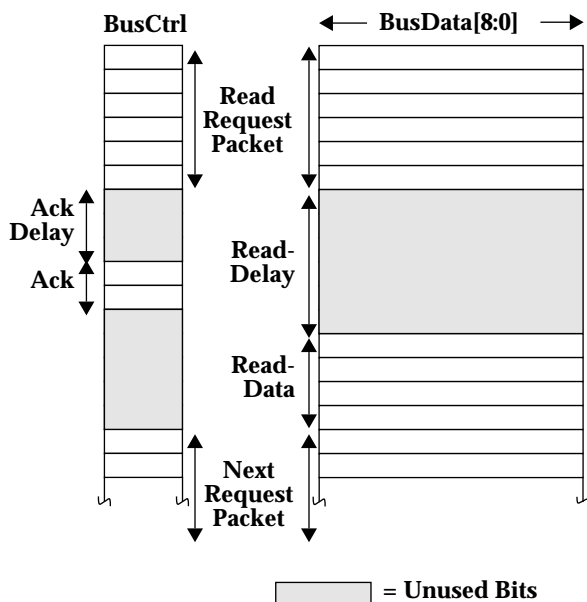


Figure 10: Read Transaction Format

Write Transaction

A write transaction is shown in Figure 11. As in the read transaction, the acknowledge packet is returned a time **AckDelay** after the request packet. If the acknowledge is Okay, the write data packet is transferred a time **WriteDelay** after the request packet.

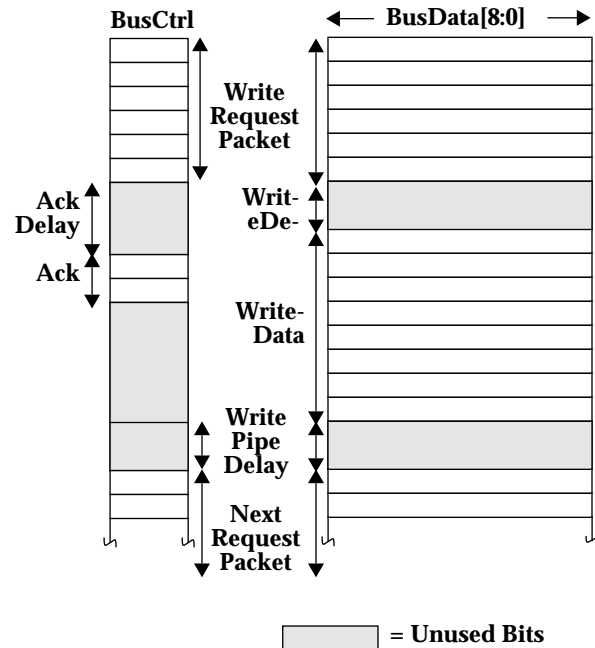


Figure 11: Write Transaction Format

Depending on the address and count values, a small delay may be required between the data packet of the current transaction and the request packet of the next transaction for write transactions. This delay is the **WritePipeDelay**. It allows the Rambus device enough time to finish one operation before beginning the next.

Inside an NEC Rambus DRAM (NEC RDRAM)

NEC RDRAMs provide dramatic performance gains because they:

- Rely on a block-oriented synchronous protocol and transfer nine bits every two nanoseconds
- Use existing sense amplifiers as a high-speed cache, offering better price/performance than a traditional second-level cache
- Integrate address mapping registers, permitting creation of sense-amplified caches with hit rates above 90%

DRAM Arrays and Sense Amplifiers

Internally, NEC RDRAMs use traditional sense amplifiers and standard DRAM arrays. The 16/18-Mbit RDRAM, depicted in Figure 12, is internally arranged as two arrays of 1M x 8/9 bits each.

Figure 13 shows the logical content of an 16/18-Mbit RDRAM. The 16,384/18,432 (2K x 8/2K x 9) sense amplifiers associated with each array are a cache that can be read or written at 500 MBytes per second across the Rambus Channel when a cache hit occurs (the requested location is in the sense-amplified latches). Only a single NEC RDRAM is required on a Rambus Channel in order to transfer at this rate.

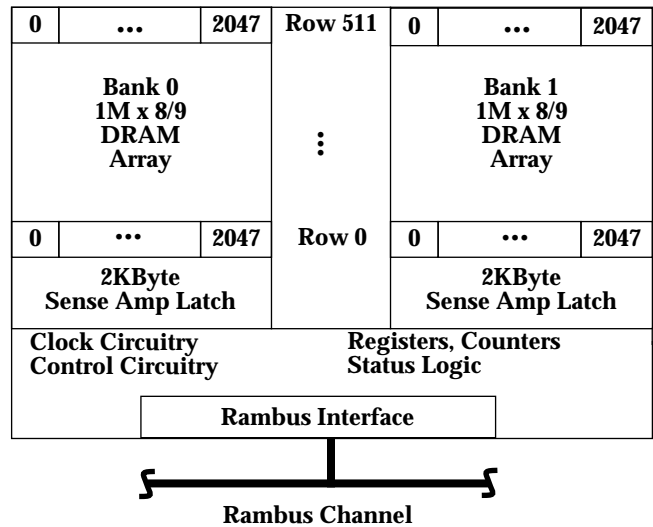


Figure 13: 16/18-Mbit NEC RDRAM Architecture

Hit and Miss Delays for NEC RDRAMs

NEC RDRAM banks normally operate in an activated state, this is, each bank has an active row. When an access is made to an active row, a hit occurs. The time required for an access in this case corresponds to the CAS access time, t_{CAC} , of a conventional page-mode DRAM.

When an access is made to a row that is not active, a miss occurs. The time required for an access in this case corresponds to the RAS cycle time, t_{RC} , of a conventional page-mode DRAM. While one NEC

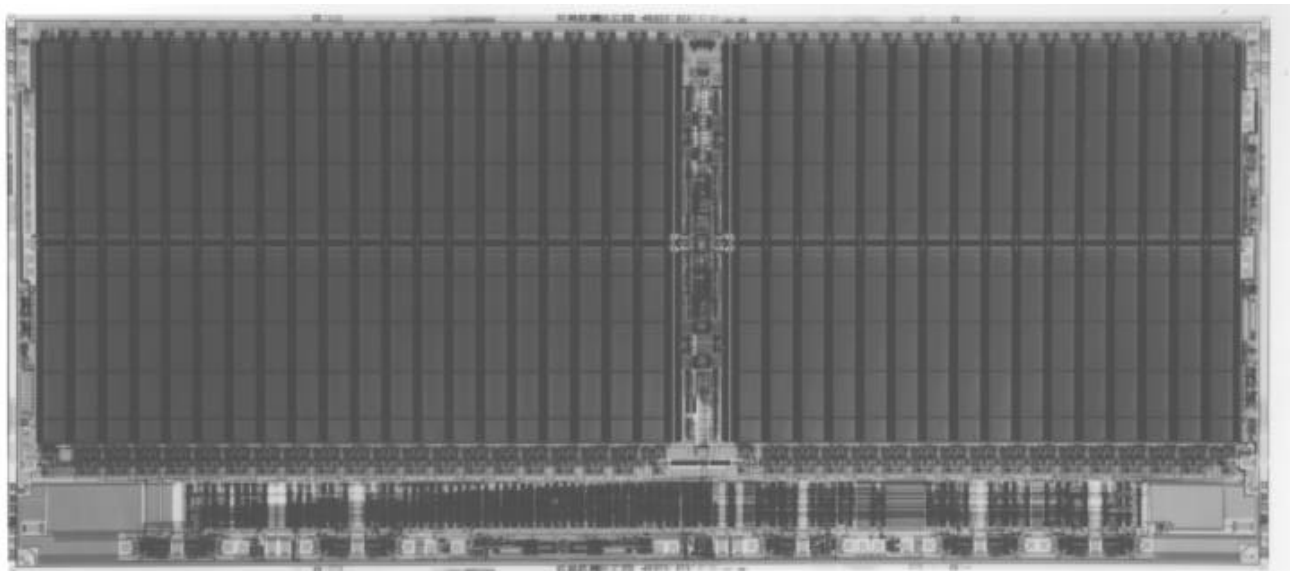


Figure 12: 16/18-Mbit NEC RDRAM Die

RDRAM is processing a miss, it is possible to access a different NEC RDRAM. The first NEC RDRAM may then be accessed again, resulting in a hit.

When a cache miss occurs, the NEC RDRAM tells the master it cannot fulfill the request (a Nack) and then automatically fetches the required line, and stores it in the sense-amplified latches. Specific timing of the delays, as defined in the previous section, are configurable via registers located in the NEC RDRAM. The operation of the Rambus Channel remains constant between manufacturers.

The Transaction Protocol

The request packets issued by master devices specify an initial starting address and a count of up to 256 bytes to be written or read during one transaction. The NEC RDRAM decodes the address in the packet, and if the data is contained within that NEC RDRAM, an acknowledge is sent back to the master. If the NEC RDRAM can complete the transaction (for example, when the requested address corresponds to an address in the sense amplifiers), then it responds with an Okay. Otherwise, it returns a Nack indicating that the master should try the request again. The transaction protocol minimizes the access time latency that results when a Nack is returned by permitting transactions to be overlapped.

The use of a synchronous block-oriented protocol and a consistent interface means that architectures incorporating the Rambus Channel are both modular and scalable. In other words, hardware designs need not change when higher-capacity memory devices become available because they will be pin-compatible with current-generation parts.

16/18-Mbit NEC RDRAMs

The 16/18-Mbit NEC RDRAM is organized as either 2M x 8 or 2M x 9 and operates at 3.3 volts. Enhancements make the 3.3-volt NEC RDRAM ideal for portable and handheld applications because the power consumed per byte transferred is much lower than any alternative device.

A number of sophisticated features enhance the performance of and simplify accesses to the 16/18-Mbit NEC RDRAM. One of the most important of these features is a set of commands that permits random access transfers. This allows the master device to access any byte location within the currently active

row. This greatly accelerates the speed of line drawing algorithms in graphics applications.

Two other performance-enhancing features are write-per-bit (WPB) and mask-per-bit (MPB). WPB permits individual bits within a data packet to be selectively written according to either a mask register (persistent WPB) or a field in the data packet (nonpersistent WPB). MPB uses the data packet to mask data provided from an internal register. All of these operations further increase the NEC RDRAM advantage in 2D and 3D graphics applications.

The 16/18-Mbit NEC RDRAM also provides the ability to prematurely terminate burst transactions through the use of a Stop command. This command, for example, minimizes memory latency for CPU accesses while allowing long and thus efficient transfers to be used for operations such as graphics screen refresh. This feature is particularly valuable in unified memory implementations, where main memory, graphics memory, and video frame buffers are located in the same physical devices.

A final feature is support for interleaved transactions. This can be used to greatly increase bus utilization by pipelining accesses to multiple RDRAMs. Such interleaved transactions allow request packets and data packets to share the Rambus Channel with no delays between them. The result is optimized use of bandwidth and enhanced performance.

Self-Refresh Capability

Self-refresh capability is built in. If a refresh cycle is in progress when the master initiates a request, the slave responds on the BusCtrl line with a Nack. This indicates that the slave is busy, and the request should be sent again later by the master. Self-refresh is an optional configuration mode that can be set during device initialization. Otherwise, the system can ensure that refreshing is accomplished by accessing all rows within the specified maximum refresh time.

The Rambus Interface

NEC RDRAMs incorporate a Rambus Interface that converts on-chip CMOS levels to the low-voltage-swing Rambus Channel levels used off-chip. This circuitry, the Rambus Interface, also contains logic that implements the acknowledge protocol, decodes the request packets generated by masters, and mul-

tiplexes one of the sense-amplified latch arrays of the NEC RDRAM to the Rambus Channel. The latch array accessed is determined by a bit field in the address of the request packet. The Rambus Interface circuitry uses phase-locked loops to synchronize on- and off-chip clocks and data, ensuring that data remains centered about the clock.

Operating Modes

NEC RDRAMs have three operating modes distinguished from one another by their power consumption. Active mode has the fastest access to data. Standby and Power Down modes use progressively less power. Standby mode limits power consumption by turning off the data bus receivers while keeping the clock circuitry active. Less current is used in this state, at the expense of a short latency time before Active mode can be entered and transactions processed.

In Power Down mode, the internal clock generator is shut down, thereby increasing the latency time required to enter Active mode. As a result, an external refresh clock must be supplied via the SIn input to preserve DRAM data. The rest of the NEC RDRAM will be powered down in this mode, allowing the lowest possible operating current.

Internal Registers

NEC RDRAMs use on-chip registers for the following:

- Device type and size encoding
- Manufacturer's ID
- Control of DRAM refresh
- Control of request/acknowledge transaction timing
- Mode control
- Current control
- Address mapping

The Mode register is used to place the NEC RDRAM in Power Down mode, to enable self-refresh, and to set other parameters necessary for proper operation.

Initialization

Each NEC RDRAM holds a unique base memory address. This unique base address is programmed into each device via SIn and SOut during initialization. This is accomplished by first resetting all NEC RDRAMs and then taking the SIn pin high on the first NEC RDRAM on the channel. Only that NEC RDRAM will respond to initialization requests sent as packets on the channel.

After the first NEC RDRAM is initialized with its unique address, it then sends a high-level signal on its SOut line to the SIn line of the next NEC RDRAM. This continues in succession until all NEC RDRAMs are programmed and properly initialized. Thereafter, the SIn and SOut lines remain unused for normal operation of the channel. The SIn and SOut pins are also used during Power Down mode, when the Rambus Channel is inactive, in order to pass along refresh requests from the system.

Other Slave Devices

Slave devices are not limited to just DRAMs. The Rambus Interface can be added to EPROMs, flash memories, SRAMs, RAMDACs, nonvolatile RAMs or even to other types of devices in general. Since the silicon overhead is small for a Rambus Interface, it allows incorporation of this high-speed channel onto cost-sensitive, high-volume devices.

Attaching Masters and Slaves to the Rambus Channel

Rambus Interface circuitry can be added to existing devices as shown in Figure 14. The application designer need not be concerned with the details of the physical layer interface. Instead, the low-level, high-speed circuitry is provided, either as an ASIC I/O cell or as a full custom cell obtained from Rambus Inc. This Rambus Interface circuitry can be placed within the pad ring of an embedded array or standard cell design as shown in Figure 15. In either case, the design can be done using standard sub-micron CMOS processes.

The Rambus Interface circuitry consists of serial-to-parallel and parallel-to-serial conversion plus clock recovery. This conversion reduces the Rambus Channel data rate of 1 byte every 2 nanoseconds to a data rate seen by the application circuitry of 8 bytes every 16 nanoseconds. The application designer must implement the logical layer of the Rambus Channel protocol. This is simplified by the availability of sample implementations from Rambus Inc.

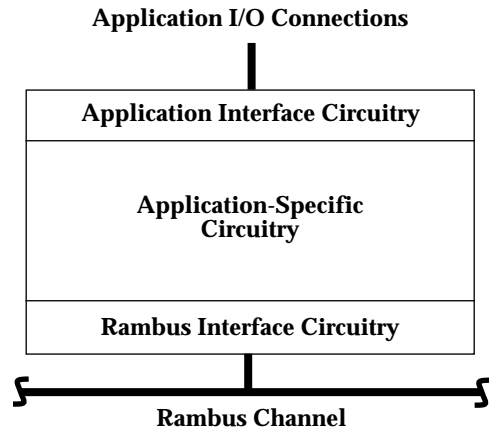


Figure 14: Block Diagram of Rambus Channel Master Implementation

Initialization

Registers contained within the devices attached to a Rambus Channel must be initialized prior to the normal operation of the Rambus Channel. Although other techniques are possible, this is most easily accomplished by including a processor somewhere in the system that can execute from a boot ROM prior to utilizing the Rambus Channel.

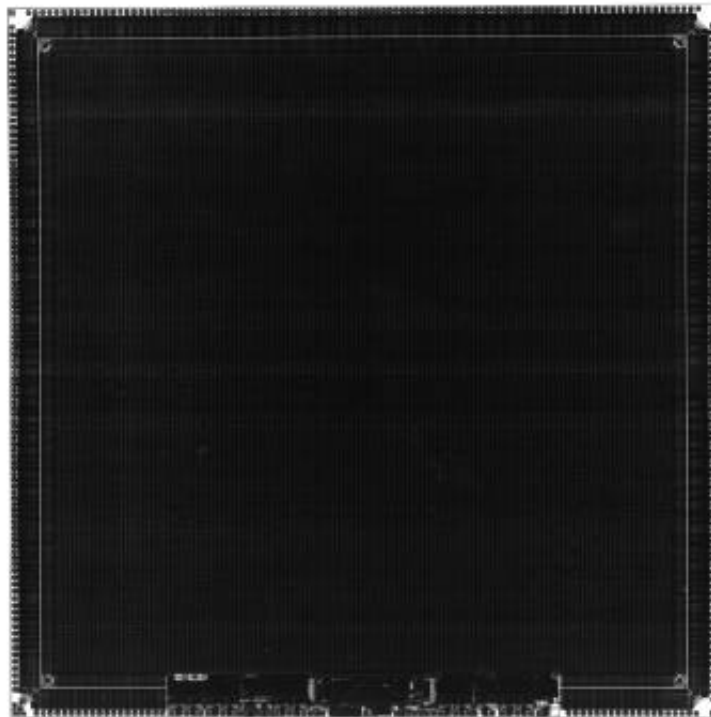


Figure 15: NEC Rambus ASIC Cell Plot

Increasing Maximum Capacity

The Primary Rambus Channel supports up to 32 slaves, such as NEC RDRAMs, soldered to a PC board. The maximum number of slaves connected to a Rambus Channel can be increased with a special device called an RTransceiver™. The RTransceiver, as shown in Figure 16, acts as a slave on the primary Rambus Channel. It repeats on the secondary Rambus Channel requests that originated with the primary channel master. The RTransceiver also ensures that data is properly forwarded between the primary and secondary channels.

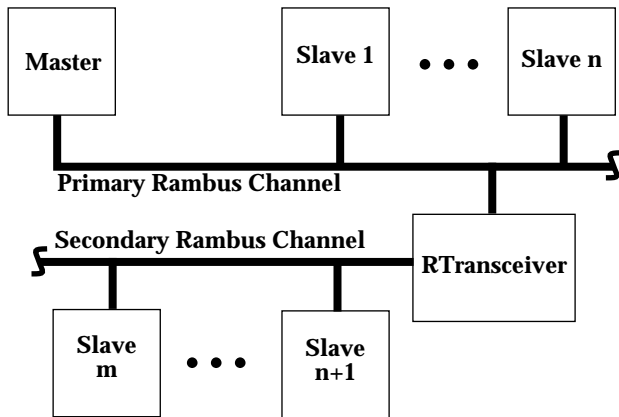


Figure 16: Location of Masters and Slaves on Primary and Secondary Rambus Channels

Field Expansion

The Rambus primary channel can support up to 10 secondary channels. Each secondary channel can support 32 slave devices. Thus a single Rambus primary channel can be expanded to accommodate a total of 320 RDRAMs.

To facilitate field expansion of channels, Rambus has defined the Rambus Socket (RSocket) and Rambus Module (RModule™). An RSocket, which is soldered to the PC board, allows connection of a single Rambus Module, which holds the secondary channel.

In some applications, graphics is one example, it is desirable to save space and keep upgrade options to a minimum. In such cases, a single RModule would typically be used. With this type of configuration, the RModule may be treated as an extension of the primary channel and an RTransceiver is not required. An example of such a configuration is provided in Figure 17.

Attaching Processors

Integrating the Rambus Channel with a traditional CPU bus can be done in either of two ways:

1. By placing Rambus Interface circuitry directly on a processor die
2. By using a separate device that contains Rambus Interface circuitry

Some possible configurations for machines incorporating a Rambus Channel and (optionally) a traditional bus are shown in Figure 18. Configuration A incorporates first-generation Rambus components into a typical system designed today. Configuration B has both a traditional bus and a Rambus Interface on the CPU die. Configuration C adds traditional bus capability to a system with a Rambus Interface on the CPU die. Configuration D uses only a Rambus Channel in the system. These diagrams represent a progression toward lower cost (fewer pins and smaller die size) and higher performance (less intermediate logic delay and greater bandwidth).

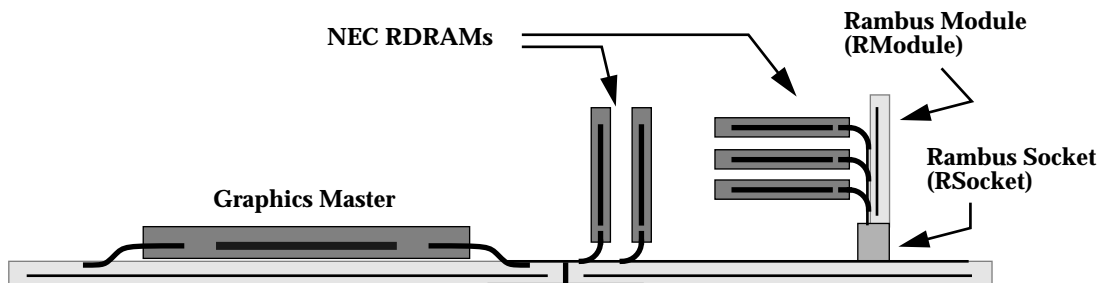


Figure 17: Easy Expandability of An NEC RDRAM-Based Graphics System

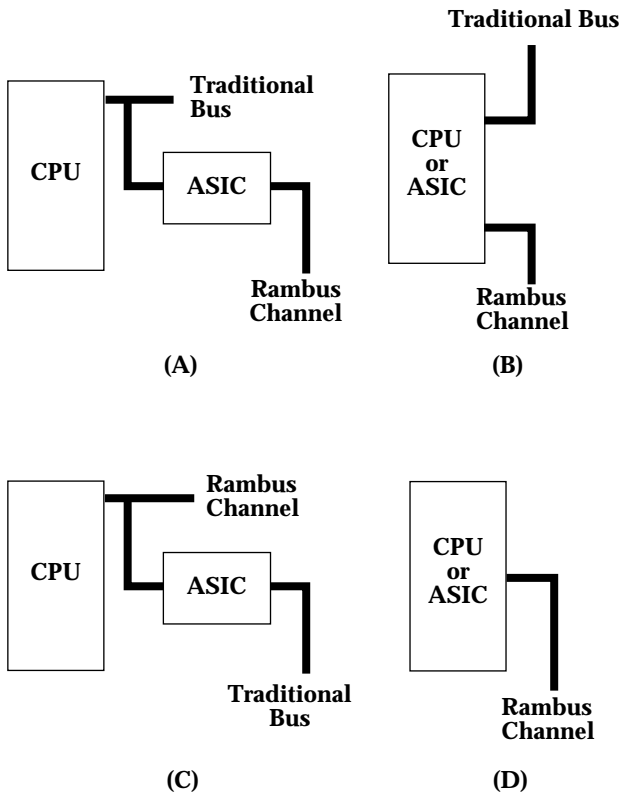


Figure 18: Integrating a Rambus Channel With a Traditional Bus

The traditional bus shown in the Figure 18 diagrams could be an i486™ bus or any of a number of other buses. It operates at CMOS- or TTL-voltage levels, uses a wide data bus of 16 to 64 bits, and operates at 20 to 70 MHz. By contrast, the Rambus Channel transfers data at 500 MHz on a 9-bit-wide data bus.

Attaching Graphics Controllers

Many different channel configurations are possible when implementing graphics controllers. A fundamental choice is illustrated in Figure 19. Configuration A represents the traditional approach of separating the graphics frame buffer from main memory.

Configuration B combines the main memory and frame buffer on a single Rambus Channel. This configuration minimizes the latency between the CPU and graphics memory, thereby permitting extremely fast main memory to screen BitBlts.

In configuration C, the graphics controller is a slave on the primary Rambus Channel but is the master on the secondary Rambus Channel. Basically, it is a smart RTransceiver. This configuration removes the refresh bandwidth from the primary Rambus Channel. The update bandwidth is mostly removed from the primary Rambus Channel, presuming that the graphics controller receives high-level imaging commands rather than low-level frame buffer updates.

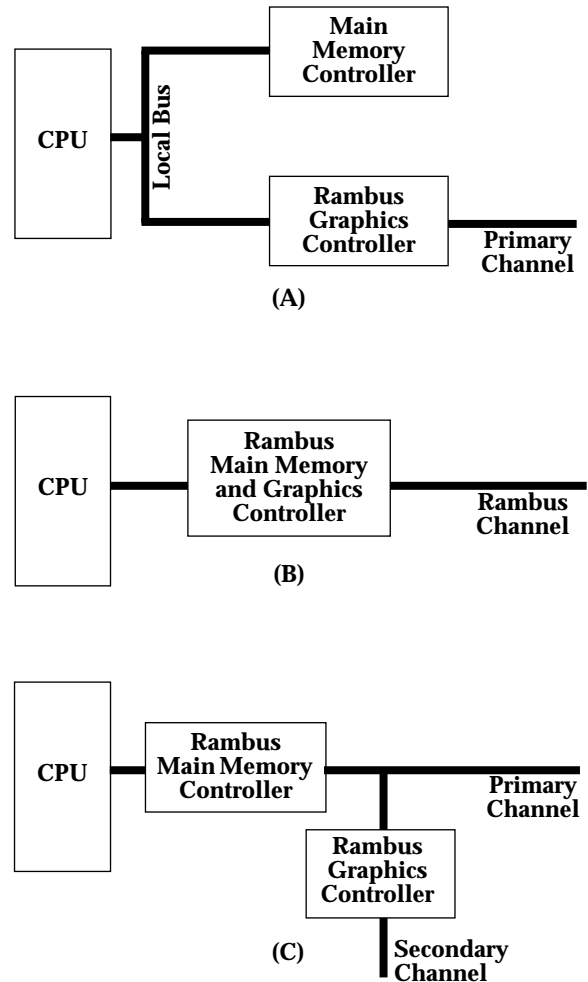


Figure 19: Sample Graphics Channel Configurations

Packaging for Speed

When transferring data at 500-Mbits-per second per wire, the impedance of on- and off-chip interconnect, as well as packaging structures such as leads and bond wires becomes critical. Die layout is also a concern.

NEC and Rambus Inc. have met these challenges with innovative packaging technology that uses proven materials and techniques. All Rambus Interface pins are located on one edge of the package. This dramatically improves the electrical environment as follows:

- Die pads, package leads, and printed circuit board traces are aligned. Thus, the electrical stubs they form on the Rambus Interface are uniform in length. As a result, all signals are delayed by an equivalent amount and skew is nearly nonexistent. Bus delays are predictable.
- The length of package leads and bond wires is minimized (less than 2.5 mm) and so is lead inductance. The resulting loads are mostly capacitive.

High-speed signal lines are separated by either a power or a ground pin, which are both AC grounds. This minimizes ground bounce and signal coupling.

In addition, devices have their leads formed to allow vertically oriented, edge surface mounting. The result is greater packaging density.

As an example, consider the packaged NEC Rambus DRAM in Figure 20. Gull wing leads on each end provide support and conduct heat to the printed circuit board. The same package is used for both 4-

Mbit and 16-Mbit RDRAMs and has been standardized by EIAJ.

The physical length of any one Rambus Channel is limited by the clock interval time of 2 nanoseconds. As a result, no more than 32 NEC RDRAMs can be located on a single Rambus Channel. Recall that additional capacity can be added by defining a single primary Rambus Channel and one or more secondary Rambus Channels. A packaged RTransceiver connects a primary Channel to a secondary Channel.

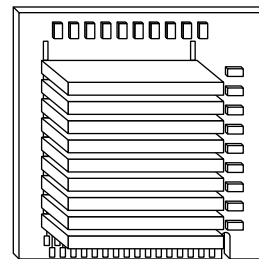


Figure 21: 16/18 MByte RModule Using 16/18-Mbit RDRAMs

Rambus Inc., in consultation with major interconnect vendors, has developed the RSocket for memory and other modules. As depicted in Figures 21 and 22, a memory RModule incorporates termination resistors, decoupling capacitors, vertically mounted NEC RDRAMs and, optionally, an RTransceiver.

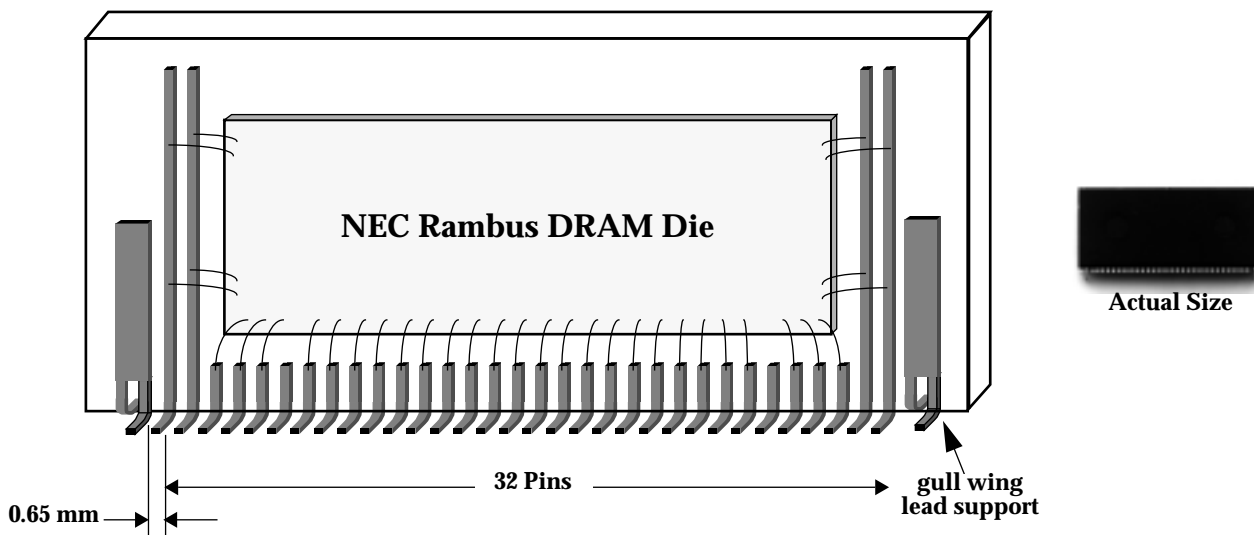


Figure 20: NEC Rambus DRAM Package

System Packaging

Printed circuit (PC) boards carrying Rambus Channel signals use standard FR-4 construction. Dielectric thickness is 5 mils (surface trace to ground layer) with 8-mil copper traces, resulting in a nominal 55Ω trace impedance. This impedance must be controlled to within a $\pm 20\%$ tolerance during bare-board manufacturing.

Separate power and ground planes are required for noise immunity. Two or more signal layers may be used. Design rules call for 8 ± 1 -mil-wide signal traces on 0.65-mm (about 25 mil) centers. This spacing matches that of pins emanating from packages incorporating a Rambus Channel. Rambus supplies detailed information for use by PC board layout personnel; this "cookbook" streamlines the layout process. In addition, due to its smaller number of signals, a Rambus Channel design generally uses fewer PC board layers than traditional designs.

Power supply decoupling capacitors are used for NEC RDRAMs and other devices on the Rambus Channel. For example NEC RDRAMs require a $0.01\mu\text{F}$ bypass capacitor per package.

As mentioned above, the physical length of any one Rambus Channel is presently limited (to approximately 10 cm) by the 2-nanosecond propagation time of signals from one end to the other. This length can accommodate up to 32 NEC RDRAMs, or up to ten memory RModules, or combinations of the two. Ten memory RModules hold up to 320 DRAMs, giving a total of 640 MBytes of memory capacity using 16-Mbit NEC RDRAMs.

Figure 22 shows a high-end system requiring RTransceivers due to its size. A smaller system (used in a graphics subsystem or desktop computer) would use a single RSocket for expansion and no RTransceiver.

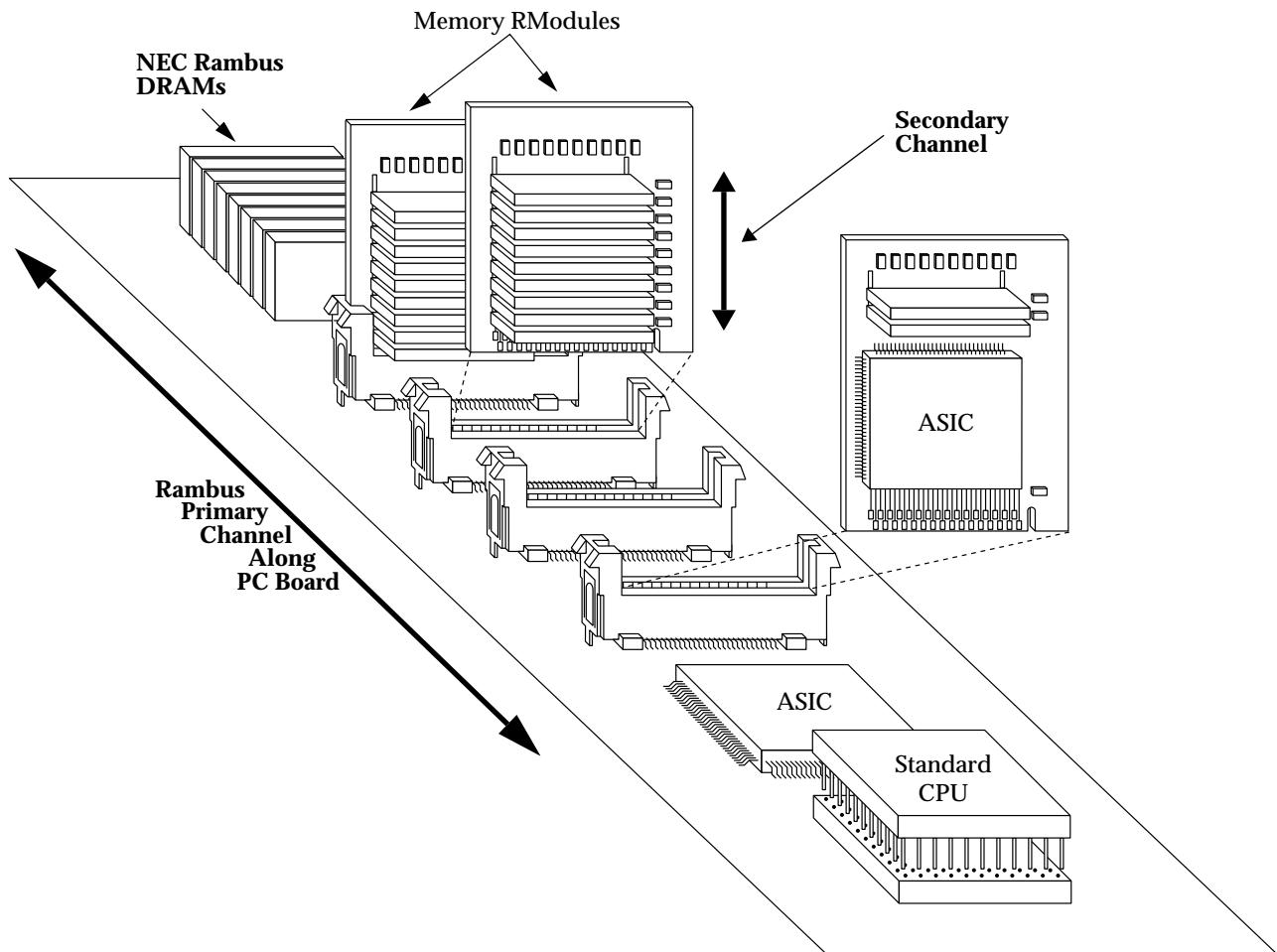


Figure 22: Physical Organization of a Rambus-Based System

Main Memory Application

A direct comparison of an NEC RDRAM memory system to a traditional one using page mode DRAMs shows the substantial advantages of an NEC RDRAM system.

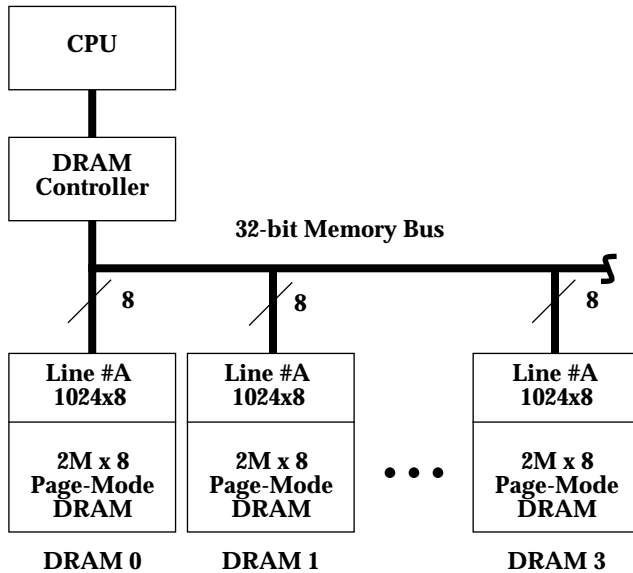


Figure 23: Typical System Using Page-Mode DRAMs

A page-mode DRAM provides two-stage access to memory data. It accepts a row address and transfers the requested row data into its sense-amplified latches. Each successive access to the same row (a page hit) can provide data with a column-access cycle. An access to a different row (a page miss) requires a slower combination of row- and column-access cycles.

A typical 2 Mbit x 8 (16 Mbit) page-mode DRAM has 1024x8 sense-amplified latches and eight data pins. A typical memory system uses four page-mode DRAMs in parallel to form a 32-bit word. Figure 23 shows the organization of such a system. "Line #A" represents the row currently latched in the sense amplifiers. Together, the sense-amplified latches of the four DRAMs form one line, 4 KBytes in size.

With just one line, the hit rate is only 30 to 40%. When hits do occur, data is transferred every 50 to 80 ns for a peak data rate of 80 MBytes per second.

Figure 24 shows a system with four NEC RDRAMs. Each NEC RDRAM is independent, that is, each

supplies full-size memory blocks (16 bytes, for example) at up to 500 MBytes per second. This system could operate with a single NEC RDRAM for a minimum granularity of 2 MBytes. By contrast, the page-mode DRAM system requires four parallel DRAMs with a minimum granularity of 8 MBytes to provide wide words and an adequate data rate. This is not necessary with NEC RDRAMs.

The NEC RDRAM's sense-amplified latches operate as a cache. A 16-Mbit NEC RDRAM has two fully independent banks, each with its own cache line. With four RDRAMs as shown in Figure 24, the sense amplifiers form eight lines of direct-mapped cache, each 2 KBytes in size. With eight cache lines, the hit rate rises above 85%. When a hit occurs, one byte of data is transferred every 2 nanoseconds for a peak data rate of 500 MBytes per second.

The systems shown in Figures 23 and 24 have the same number of DRAMs. The NEC RDRAM-based system has four times the total cache size, 8 times the number of lines, less than one-sixth the miss rate, and over six times the peak data rate of the system based on page-mode DRAMs. Similar ratios apply to larger memory systems. Consider also that the finer granularity will be even more valuable as DRAM densities increase in the future.

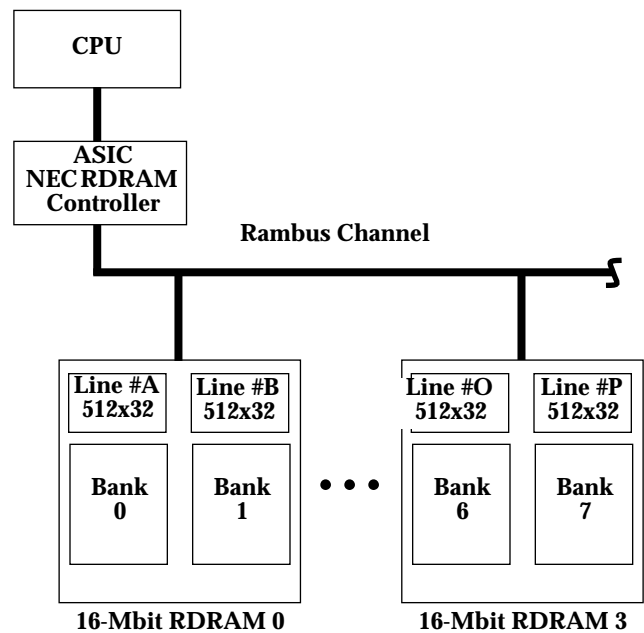


Figure 24: NEC RDRAM-Based System With Independent Sense Amp Latches

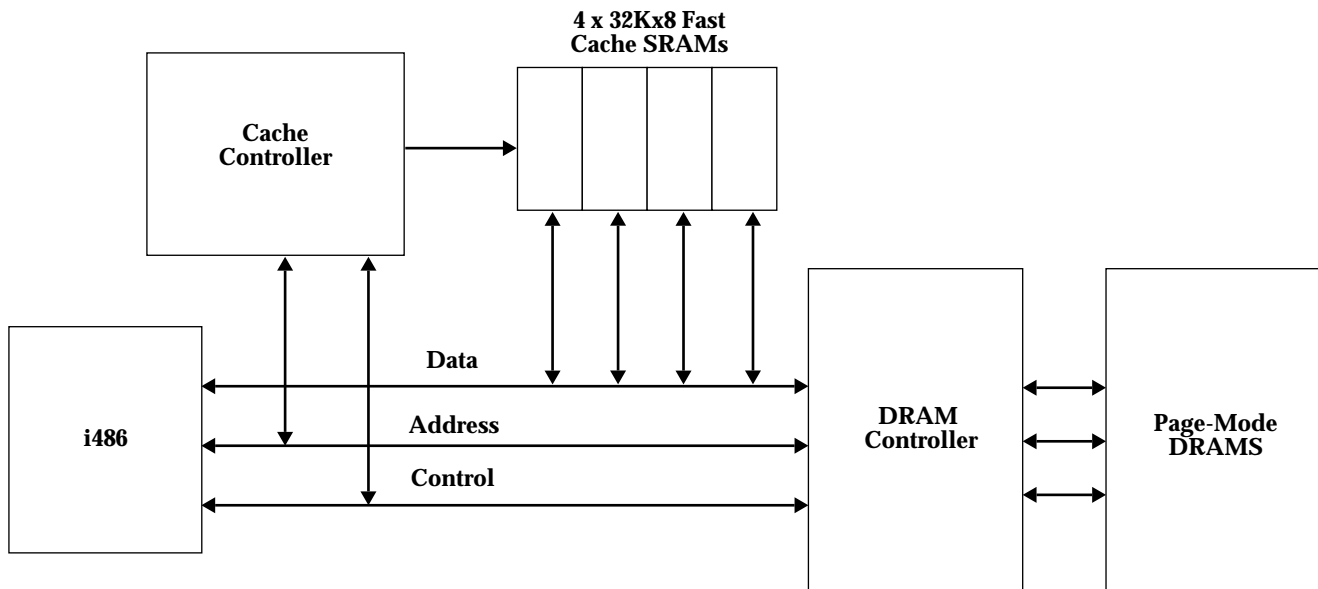


Figure 25: Typical i486 System Memory Architecture

Level Two Cache Replacement

High-performance systems typically include an L2 (level two) cache between the processor and main memory. Consider the i486 architecture depicted in Figure 25. It uses four 32-KByte SRAMs to implement a 128-KByte L2 cache (the i486 contains an 8-KByte L1 cache). This design also requires a cache controller with 20 KBytes of tag storage.

A Rambus-based solution offers equivalent performance with a greatly reduced chip count. This solution, shown in Figure 26, uses an ASIC as an NEC RDRAM controller. The ASIC integrates:

- Separate code and data prefetch buffers (128 bytes total) used for reads from main memory
- A prefetch assist cache (0.5 to 2 KBytes)
- A write buffer (16 bytes)
- An L2 cache (1 to 8 KBytes) plus tag RAM

Prefetching boosts performance by loading data before it is requested by the CPU. When sufficient main memory bandwidth exists, as with a Rambus-based system, a prefetching 8-KByte L2 cache can equal the performance of a 128-KByte L2 cache.

This ASIC system is less costly and uses less power and board space because a separate cache controller and cache SRAMs are not required. The ASIC uses a conventional sub-micron CMOS process. NEC Rambus DRAMs, with a price-per-bit comparable to conventional DRAMs, are used for main memory. The speed of the Rambus Channel, at 500 MBytes per second, permits prefetching algorithms to boost performance without limiting non-prefetch accesses to main memory.

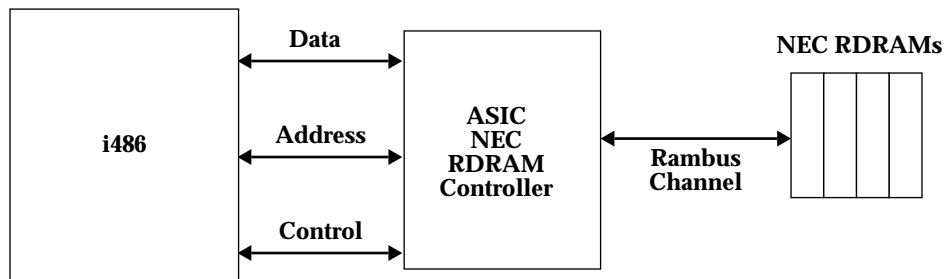


Figure 26: i486 System Using NEC Rambus DRAM Main Memory

Bit-Mapped Graphics in a Rambus World

The clear trend for PC and workstation graphics is toward higher-resolution, flicker-free displays able to display increasing numbers of colors. In addition, the images displayed today are moving from static two dimensional to animated two dimensional. In the near future, real-time video will appear on the desktop, followed eventually by animated three dimensional with photorealism. The technology to make these transitions is available now; the constraint to its widespread use is affordability. Current graphics memory approaches are fast becoming obsolete, and technology available today from NEC and Rambus provides the capability to meet today's needs and those well into the future.

Consider the standard page-mode DRAMs commonly used for computer main memory. These are also the most economical form of video memory. However, high-performance graphics can only be achieved with specialized, higher-bandwidth memory chips. These consist of either video RAMs (VRAMs) or 16-bit-wide versions of page-mode DRAMs. (VRAMs are dual-port DRAMs; one port is a standard DRAM interface used for CPU and graphics controller access. The serial port is used for screen refreshing.) Such specialized chips tend to be one generation behind standard DRAMs with respect to density, that is, video designs using them typically have higher parts costs and require more PC board space. In addition, a VRAM die is larger than a DRAM die of the corresponding density. These factors drive up cost.

Graphics System Requirements

A modern graphics system must meet two major performance objectives. First, it must provide flicker-free screen refresh. Second, it must update the image maintained in the frame buffer fast enough to meet the requirements of the application currently being executed.

At a minimum, the frame buffer must contain the following number of bits:

$$(\text{horizontal resolution}) \times (\text{vertical resolution}) \times (\text{color bits/pixel})$$

The standard non-interlaced refresh rate is currently 72 Hz. That is, every pixel on the screen is repainted 72 times per second. Table 2 shows the performance requirements for this refresh rate as screen resolution and color capacity increase.

Graphics applications have an almost insatiable desire for bandwidth. Scrolling windows, repainting complex drawings, animated frames, video windows, and teleconferencing all have enormous bandwidth requirements. In order to satisfy these requirements, future generations of VRAMs will have to make a quantum leap in performance. This will not be easy to do without making the device more expensive. NEC RDRAMs have 10 times the bandwidth of current 256K by 8 VRAMs, enabling high-bandwidth applications while providing enough bandwidth for refreshing of the display.

Rambus Advantages for Graphics Systems

The Rambus Channel and NEC RDRAMs provide a far superior solution for the high-performance and low-cost requirements of large-screen, true-color graphics. This is the case for a number of reasons.

First, a Rambus Channel uses just 15 active signals and requires just 28 pins on the graphics controller, but can transfer data at up to 500 MBytes per second. This is sufficient for all but the most advanced requirements today. If more performance is needed, additional channels can be added at 28 pins per channel. Also, NEC Rambus technology will perform frame buffer updates 3 to 5 times faster than is possible with VRAMs.

Second, NEC RDRAMs are less costly than VRAMs. This is the case since the Rambus Interface circuitry that is added to a DRAM core to make an NEC RDRAM uses less die area than that required to create a VRAM from a DRAM. Also, NEC RDRAMs are multipurpose devices and are optimal for high-volume main memory applications. Consequently, they will be produced in high volumes and priced much closer to standard DRAMs than to VRAMs. And

Table 2: Geometric Growth of Frame Buffer Size and Bandwidth Needed at a 72-Hz Refresh Rate

Resolution	Bits per Pixel	Frame Buffer Size (MB)	Relative Frame Buffer Size	Required Bandwidth (MB/sec)
640 x 480	4	0.15	1x	10
1024 x 768	8	0.75	5x	54
1280 x 1024	16	2.5	17x	180
1600 x 1200	32	7.3	50x	527

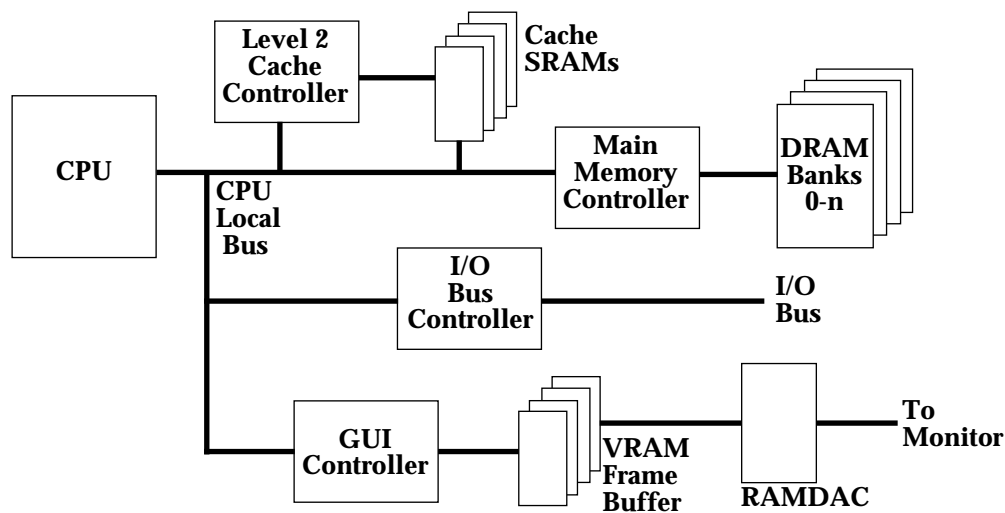


Figure 27: Architecture of a Traditional System with Local-Bus-Resident Graphics

because NEC RDRAMs are a generation ahead of VRAMs, chip counts for NEC Rambus-based systems will be significantly lower than those for VRAM-based systems.

Third, the same interface will be used on future generations of NEC RDRAMs (64 Mbits). This will simplify the task of designing a graphics subsystem and will help keep NEC RDRAM prices low.

Fourth, NEC Rambus-based graphics system designs are scalable. A graphics controller can be designed to handle a wide range of resolutions and bits per pixel.

For example, a designer might use an NEC Rambus-based graphics controller and one 16-Mbit NEC RDRAM for the frame buffer to handle 1024 x 768 x 16 graphics. This configuration outperforms one using a complex, interleaved, 4-MByte VRAM frame buffer. An RSocket on the board would allow field upgrading by the user to expand the number of colors or the resolution.

Graphics System Architectures

To increase graphics performance, more and more manufacturers are placing the graphics subsystem directly on the CPU local bus (Figure 27) instead of on the much slower I/O bus. Putting graphics hardware on the local bus speeds up data transfers between main memory and the frame buffer but adds complexity and cost.

Since NEC RDRAMs are superior for both main memory and graphics applications, the ultimate integration will be the use of NEC RDRAMs for both

the frame buffer and main memory. The architecture of a system using this type of unified and simplified memory configuration is depicted in Figure 28.

There are two ways to accomplish this integration. The frame buffer can be kept separate from the main memory, with a 500-MByte-per second Rambus Channel between them for BitBlt transfers. Alternatively, the frame buffer could be integrated into the main memory. Within such integrated systems, the Rambus Channel has ample bandwidth for both graphics and main memory traffic by using one or more Rambus Channels to meet the price/performance requirements of the system.

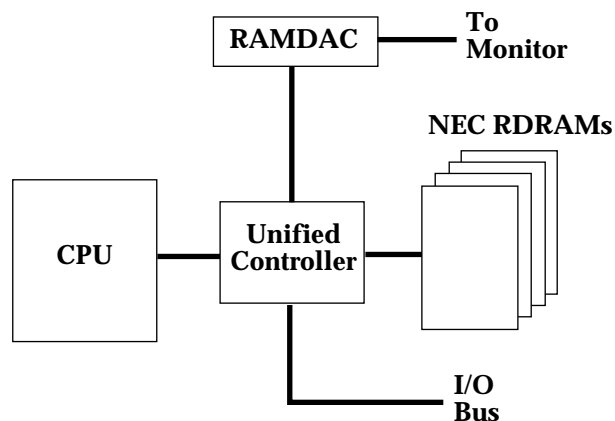


Figure 28: An NEC Rambus-Based System with Unified Graphics and Main Memory

The Advantages of NEC Rambus Technology

As explained in the preceding sections of this *Architectural Overview*, the concepts and techniques pioneered by NEC and Rambus offer a number of significant advantages. These advantages include high performance, reduced system size and complexity, lower cost and shorter time to market, lower power consumption and improved expandability. These points are summarized in the table below.

High Performance	<ul style="list-style-type: none"> • 500 MBytes per second of bandwidth per Rambus Channel • Multiple channels can be used for even higher performance and bandwidth • Ten times faster than traditional DRAM or VRAM devices • Provides 3X to 5X VRAM system performance • Performance is equivalent to cached memory systems • Enables CPU and graphics memory to be unified for improved system price/performance
Cost Effective	<ul style="list-style-type: none"> • Eliminates cache controller, cache SRAMs and glue devices • Lower cost basis than VRAM devices • Uses mature base DRAM and PCB technologies • Eliminates the need for expensive MCM technology
Low Power	<ul style="list-style-type: none"> • Inherently use less energy per byte transferred • 16 Mbit NEC RDRAMs include low-power modes that reduce overall power consumption
Expandable / Granular	<ul style="list-style-type: none"> • 4-, 16-, 64-Mbit generations of RDRAMs have identical pinouts • A single NEC RDRAM delivers 500 MBytes per second of bandwidth • Memory can be incremented by a single NEC RDRAM • Hundreds of NEC RDRAMs can be supported by a single channel • The master system clock is decoupled from the Rambus Channel clock
Reduces Board Space	<ul style="list-style-type: none"> • The Rambus Channel is narrow and short – only 15 active signals • Eliminates components • Permits dense packaging
Quick Time to Market	<ul style="list-style-type: none"> • A pre-engineered solution is provided to the system designer/architect • Modularity allows one board design to cover multiple options
Standard	<ul style="list-style-type: none"> • All Rambus-based ICs are 100% compatible

Applying the Advantages of NEC Rambus Technology

NEC/Rambus technology represents a superior memory solution for a broad range of applications. These applications include graphics subsystems and consumer digital video products as well as desktop computers, workstations and supercomputers. A few of these applications are highlighted below.

Desktop Computer Main Memory	<ul style="list-style-type: none">• Eliminates expensive second-level caches• Ideal for high-performance, high-bandwidth main memory systems• Easily expandable to support hundreds of megabytes• Simple PC board design requires minimal space• Unifying main memory with the frame buffer leads to maximum performance at reduced cost
Graphics Frame Buffers	<ul style="list-style-type: none">• Makes workstation-quality graphics accessible at mass market cost• Additional NEC RDRAMs can be added easily for greater resolution and more colors• A single 16-Mbit RDRAM supports a 1280 x 1024 x 8 or a 1024 x 768 x 16 display with workstation-level performance• Use up to four channels per graphics controller for 2 gigabytes per second of bandwidth• Provides ample bandwidth to merge the video and graphics frame buffers
Portable and Handheld Computers	<ul style="list-style-type: none">• Built-in power management modes reduce current consumption• Gives the performance of a cached desktop computer without the physical size or power consumption• Has the lowest power per byte transferred of any DRAM alternative• Can drive internal LCD displays and an external, high-resolution monitor• Assures high performance from a very small subsystem• At most only one NEC RDRAM is active on a channel at one time
Other Applications	<ul style="list-style-type: none">• Embedded control• Digital television• Large computers• Communications systems

The adoption of NEC Rambus technology will dramatically simplify computer architectures, since hardware traditionally used to increase the speed of the processor to memory interface will not be necessary. In addition, the use by of standard CMOS processes, low-cost IC packaging and standard PC board technologies will reduce the expense of building systems based on NEC/Rambus technology. These factors will result in faster machines that consume less power, cost less and occupy less physical space.

