

# Corporate Business Servers: An Alternative to Mainframes for Business Computing

With expandable hardware, PA-RISC architecture, symmetric multiprocessing, a new bus structure, and robust error handling, these systems provide a wide range of performance and configurability within a single cabinet. Standard features include one to twelve symmetric PA-RISC 7100 multiprocessors optimized for commercial workloads, main memory configurations from 128M bytes to 2G bytes, and disk storage up to a maximum of 1.9 terabytes.

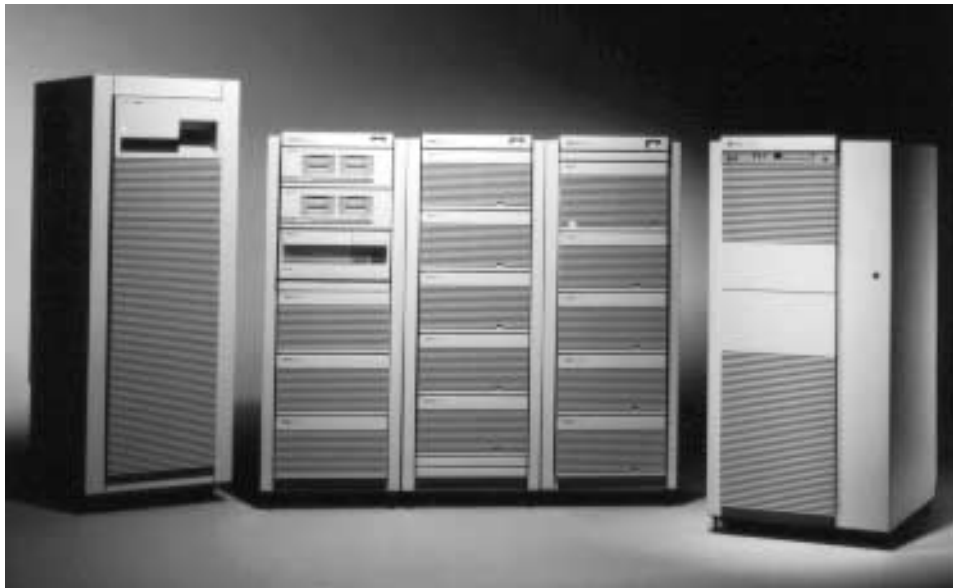
by **Thomas B. Alexander, Kenneth G. Robertson, Dean T. Lindsay, Donald L. Rogers, John R. Obermeyer, John R. Keller, Keith Y. Oka, and Marlin M. Jones, II**

The overall design objective for the HP 9000 Model T500 corporate business server (Fig. 1) was to set new standards for commercial systems performance and affordability. Combining expandable hardware, PA-RISC architecture, symmetric multiprocessing with up to 12 processors, a new bus design, robust error handling, and the HP-UX\* operating system, the Model T500 delivers a cost-effective alternative to mainframe solutions for business computing.

Users of HP's proprietary operating system, MPE/iX, also enjoy the benefits of the Model T500 hardware. These systems are designated the HP 3000 Series 991/995 corporate business systems. They provide high performance by

supporting from one to eight processors with superior value for their class. The MPE/iX system is designed to support business-critical data and offers features such as powerful system management utilities and tools for performance measurement.

In this paper, the hardware platform for both the HP-UX and the MPE/iX systems will be referred to as the Model T500. The Model T500 is an update of the earlier HP 9000 Model 890/100 to 890/400 systems, which supported from one to four PA-RISC processors operating at 60 MHz. For MPE/iX, the Series 991/995 is an update of the earlier Series 990/992 systems.



**Fig. 1.** The HP 9000 Model T500 corporate business server (right) is designed as an alternative to mainframe solutions for online transaction processing and other business computing applications. It runs the HP-UX operating system. The same hardware running the MPE/iX operating system is designated the HP 3000 Series 991/995 corporate business systems. The Model T500 SPU (right) is shown here with various peripherals and expansion modules.

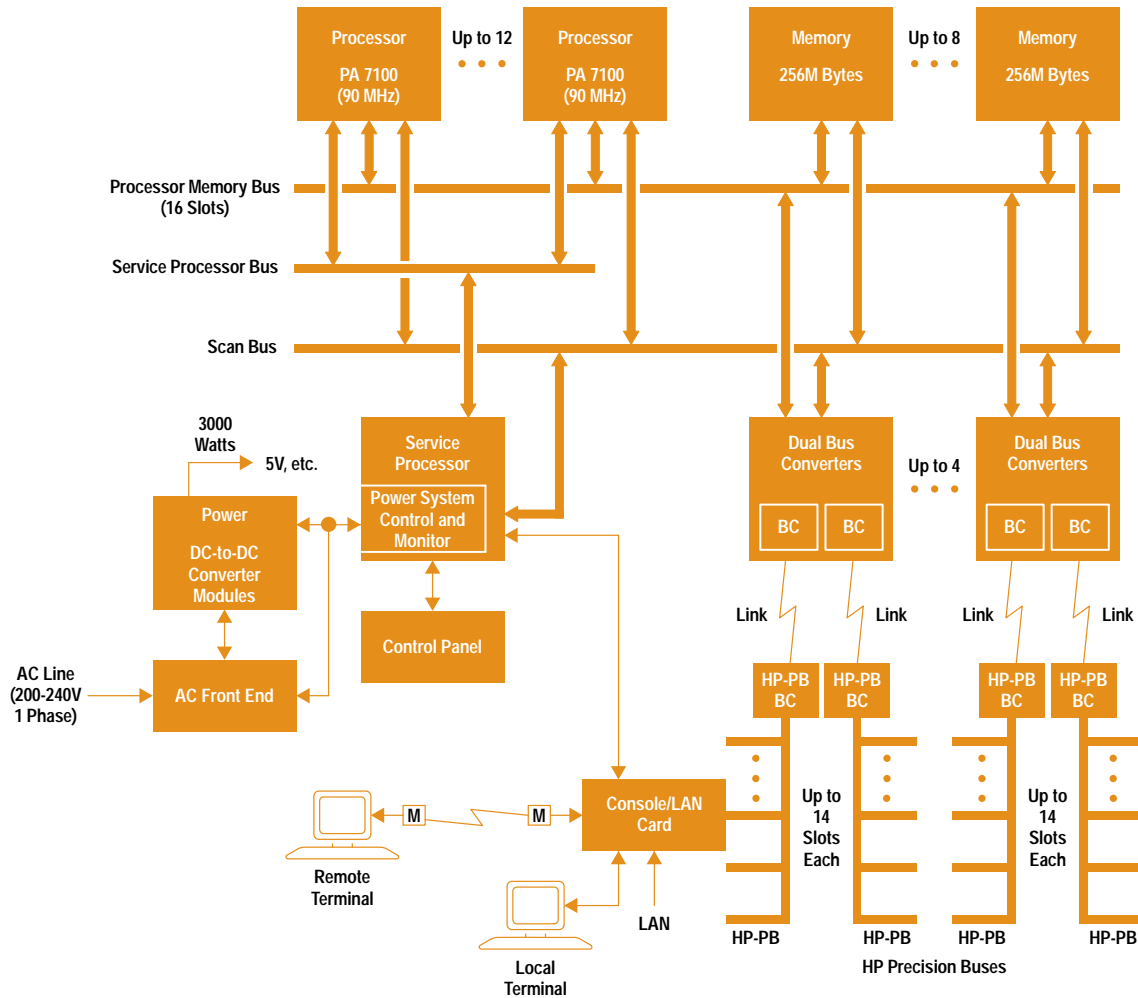


Fig. 2. HP 9000 Model T500 system processing unit block diagram.

Standard features of the Model T500 include one to twelve symmetric PA-RISC multiprocessors optimized for commercial workloads and operating at 90 MHz, main memory configurations from 128M bytes to 2G bytes,† and disk storage up to a maximum of 1.9 Tbytes (1900 Gbytes). This expandability allows the Model T500 to provide a wide range of performance and configurability within a single cabinet. The Model T500's package minimizes the required floor space, while air cooling removes the need for expensive mainframe-type cooling systems.

The Model T500 is designed to provide leading price/performance. The HP 9000 Model T500 with six PA-RISC 7100 processors operating at 90 MHz has achieved 2110.5 transactions per minute on the TPC-C benchmark (U.S.\$2115 per tpmC).†† The SPECrate (SPECrate\_int92 and SPECrate\_fp92) benchmark results show linear scaling with the number of processors, which is expected for CPU-intensive workloads with no mutual data dependencies. The Model

† Hewlett-Packard Journal memory size conventions:  
 1 kbyte = 1,000 bytes                      1K bytes = 1,024 bytes  
 1 Mbyte = 1,000,000 bytes                1M bytes = 1,024<sup>2</sup> bytes = 1,048,576 bytes  
 1 Gbyte = 1,000,000,000 bytes           1G bytes = 1,024<sup>3</sup> bytes = 1,073,741,824 bytes  
 1 Tbyte = 1,000,000,000,000 bytes

†† The Transaction Processing Council requires that the cost per tpm be stated as part of the TPC performance results. Cost per tpm will vary from country to country. The cost stated here is for the U.S.A.

T500/400 reaches 38,780 SPECrate\_fp92 and 23,717 SPECrate\_int92 with twelve processors.

The Model T500 provides this high level of performance by using a balanced bus architecture. The processor memory bus currently provides the main processor-to-memory or processor-to-I/O interconnect with a bandwidth of 500 Mbytes/s and a potential capability up to 1 Gbyte/s. The I/O buses provide a total aggregate I/O bandwidth of 256 Mbytes/s. These bandwidths satisfy the high data sharing requirements of commercial workloads.

### System Overview

The key to the Model T500's expandability and performance is its bus structure. The processor memory bus provides a high-bandwidth coherent framework that ties the tightly coupled symmetrical multiprocessing PA-RISC processors together with I/O and memory. Fig. 2 shows a block diagram of the Model T500.

The processor memory bus is a 60-MHz bus implemented on a 16-slot backplane with eight slots suitable for processors or memory boards and eight slots suitable for I/O adapters or memory boards. Each slot can contain as many as four modules, and can obtain its fair fraction of the bandwidth provided by the system bus. Custom circuit designs allow the bus to operate at a high frequency without sacrificing physical

connectivity. To prevent system bus bandwidth from becoming a bottleneck as the number of processors increases, the bus protocol minimizes bus contention and unproductive traffic without adding undue complexity to the bus modules.

To support such a large number of slots the processor memory bus is physically large and has an electrical length of 13 inches. State-of-the-art VLSI design and mechanical layout allow the processor memory bus to run at 60 MHz—a very high frequency of operation for a bus of this size.

The input/output subsystem links the processors and memory on the processor memory bus to I/O devices, including a variety of networks. The Model T500 supports attachment of up to eight Hewlett-Packard precision buses (HP-PB), each of which connects up to 14 I/O cards. The first HP-PB is internal to the Model T500 and the other HP-PBs would be located in adjacent racks. Each HP-PB connects to the processor memory bus through a linked bus converter consisting of a dual bus converter, a bus converter link, and an HP-PB bus converter. Under normal operating conditions the bus converters are transparent to software.

The service processor consists of a single card whose purpose is to provide hardware control and monitoring functions for the Model T500 and a user interface to these functions. To achieve this purpose, the service processor has connectivity to many parts of the Model T500. The scan bus is controlled by the service processor and provides the service processor with scan access to all of the processor memory bus modules. The scan bus is used for configuration of processor memory bus modules and for manufacturing test. The service processor also provides the clocks used by processor memory bus modules and controls the operation of these clocks. The service processor provides data and instructions for the processors over the service processor bus during system initialization and error recovery. The service processor connects to the control panel and provides the system indications displayed there. The service processor provides its user interface on the console terminals through its connection to the console/LAN card.

The service processor also contains the power system control and monitor, which is responsible for controlling and monitoring the Model T500's power and environmental system. The main power system receives 200-240V single-phase mains ac and converts it to 300Vdc. This 300V supply is then converted by various dc-to-dc converter modules to the needed system voltages (e.g., one module is 300Vdc to 5Vdc at 650W). The power system control and monitor additionally controls the system fans and power-on signals. The power system control and monitor performs its functions under service processor control and reports its results to the service processor.

## Processor Memory Bus

The present implementation of the Model T500 uses 90-MHz PA-RISC central processing units (CPUs)<sup>1,2,3</sup> interconnected with a high-speed processor memory bus to support symmetric twelve-way multiprocessing. This section focuses on the features and design decisions of the processor memory bus, which allows the system to achieve excellent online

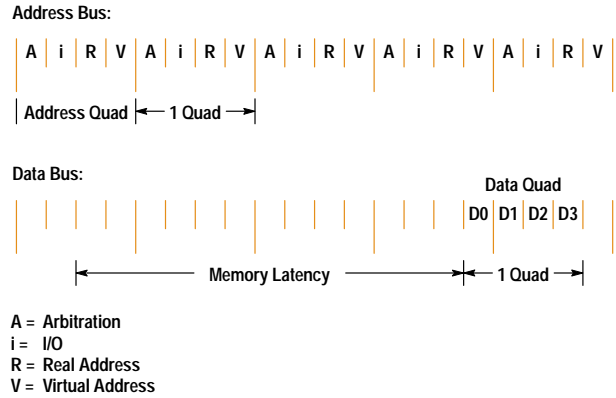


Fig. 3. Processor memory bus pipeline.

transaction processing (OLTP) performance and efficient multiprocessor scaling.

### Bus Protocol

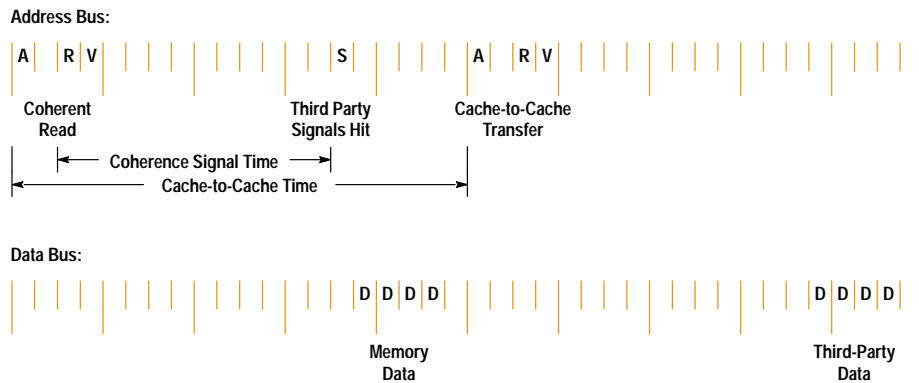
The processor memory bus is a synchronous pipelined bus. The pipelined nature of the bus protocol places it between a split transaction protocol and an atomic transaction protocol. This allows the processor memory bus to have the performance of a split transaction bus with the lower implementation complexity of an atomic transaction bus.

The processor memory bus has separate address and data buses. The address bus is used to transfer address and control information and to initiate transactions. Non-DMA I/O data is also transferred on the address bus. The data bus transfers memory data in blocks of 16, 32, or 64 bytes. The processor data bus in the present implementation is 64 bits wide, although the protocol, backplane, and memory system also support 128-bit-wide accesses. For 32-byte transfers on the data bus, the available bandwidth is 480 Mbytes per second. If processors use all 128 bits of the data bus to perform 64 byte transfers, the bandwidth doubles to 960 Mbytes per second.

Fig. 3 shows the processor memory bus pipeline. Four consecutive processor memory bus states are referred to as a *quad*. A transaction consists of a quad on the address bus, followed at some fixed time by a quad on the data bus.

An address quad consists of an arbitration cycle, an I/O cycle, a real address cycle, and a virtual address cycle. The arbitration cycle is used by bus masters to arbitrate for use of the bus. The I/O cycle is used to transfer data in the I/O address space. The real address cycle is used to transfer the memory or I/O address and to indicate the transaction type. The virtual address cycle is used to transfer the virtual index for cache coherency checks.

A data quad consists of four data transfer cycles. The fixed time between address and data quads is programmed at system initialization. This arrangement allows multiple pipelined transactions to be in progress at the same time. Since data is returned at a fixed time after the address quad, the module returning data automatically gets access to the data bus at that time. The set of supported transactions includes reads and writes to memory address space, reads and writes to I/O address space, and cache and TLB (translation lookaside buffer) control transactions.



**Fig. 4.** Multiprocessor protocol, showing cache-to-cache transfer timing.

If a transaction is initiated on the processor memory bus, but a module (either the slave or a third party) is not prepared to participate in the transaction, that module has the option of *busying* the transaction. When the master sees that its transaction is busied, it must retry the transaction at a later time. Busy is appropriate, for example, when the bus adapter is asked to forward a read transaction to the lower-speed precision bus (see "Arbitration" below for more information).

For cases in which a module requires a brief respite from participating in transactions, it can *wait* the bus, that is, it can freeze all address and data bus activity. It does this by asserting the wait signal on the bus. The wait facility is analogous to a stall in the processor pipeline.

### Multiprocessor Bus Protocol

The processor memory bus provides cache and TLB coherence with a *snoopy*<sup>4</sup> protocol. Whenever a coherent transaction is issued on the bus, each processor (acting as a third party) performs a cache coherency check using the virtual index and real address.

Each third-party processor is responsible for signaling cache coherency status at a fixed time after the address quad. The third party signals that the cache line is in one of four states: shared, private clean, private dirty, or not present. The requesting processor interprets the coherency status to determine how to mark the cache line state (private clean, private dirty, or shared). The third party also updates its cache line state (no change, shared, or not present).

If a third party signals that it has the requested line in the private dirty state, then it initiates a cache-to-cache transaction at a fixed time after the address quad. The requesting

processor discards the data received from main memory for the initial request and instead accepts the data directly from the third party in a cache-to-cache transfer. At this same time the data from the third party is written to main memory. The timing of these events is shown in Fig. 4.

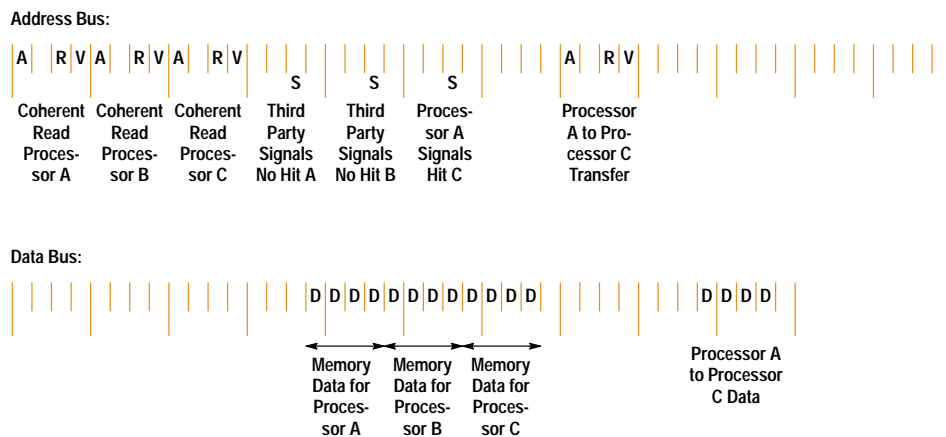
Since the processor memory bus allows multiple outstanding pipelined transactions, it is important that processor modules be able to perform pipelined cache coherency checks to take maximum advantage of the bus bandwidth. Fig. 5 shows an example of pipelined cache coherency checking.

### Programmable Parameters

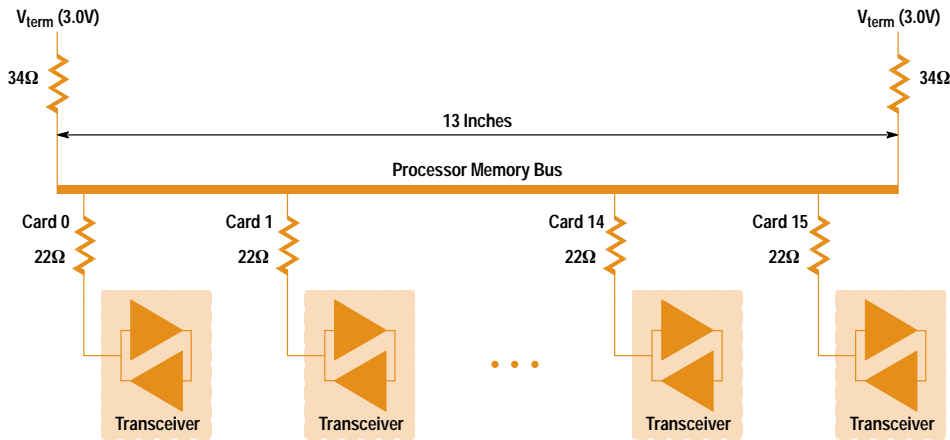
The processor memory bus protocol permits many key bus timing parameters to be programmed by initialization software. Programming allows different implementations to optimize the parameter values to increase system performance and reduce implementation complexity. Initialization software calculates the minimum timing allowed for the given set of installed bus modules. As new modules are designed that can operate with smaller values (higher performance), initialization software simply reassigns the values.

The programmable parameters include:

- Address-to-Data Latency. The time from the real address of the address quad to the first data cycle of the data quad. The present implementation achieves a latency of 217 nanoseconds.
- Coherency Signaling Time. The time required for a processor to perform a cache coherency check and signal the results on the bus.
- Cache-to-Cache Time. The time from the address quad of the coherent read transaction to the address quad of the cache-to-cache transaction. This value is the time required



**Fig. 5.** Pipelined cache coherency checking.



**Fig. 6.** Processor memory bus electrical layout.

for a processor to do a cache coherency check and copy out dirty data.

- Memory Block Recovery Time. The time it takes a memory block to recover from an access and become ready for the next access.
- Memory Interleaving. The memory block identifier assignments. The assignments depend on the size and number of memory blocks installed in the system.

### Arbitration

The processor memory bus uses three different arbitration rules to determine when a module can get access to the bus. The first rule, used for references to memory, states that a master can arbitrate for a memory block only after the block has recovered from the previous access. Bus masters implement this by observing all transactions on the processor memory bus. Since memory references include the block identifier and the recovery times are known, masters refrain from arbitration for a busy block. The benefit of this arbitration rule is that memory modules do not have to queue or busy transactions, and therefore bus bandwidth is conserved because every memory transaction is a useful one.

The second arbitration rule, used for references to I/O address space, requires that a master of a busied I/O transaction not retry the transaction until the slave has indicated that it is ready to accept the transaction. The slave indicates readiness by asserting the original master's arbitration bit on the bus. The master detects that the slave has restarted arbitration and continues to attempt to win arbitration. This rule prevents masters from wasting bus bandwidth by continually retrying the transaction while the slave is not ready to accept it, and avoids most of the complexity of requiring slaves to master a return transaction.

The third mechanism, referred to as *distributed priority list* arbitration, is invoked when multiple masters simultaneously arbitrate for the processor memory bus. Distributed priority list arbitration is a new scheme for general arbitration. It uses a least-recently-used algorithm to determine priority on the bus. A master implements distributed priority list arbitration by maintaining a list of masters that have higher priority than itself and a list of masters that have lower priority. Thus, an arbitrating master can determine it has won by observing that no higher-priority master has arbitrated. The identity of the winning master is driven onto the processor memory bus in the address quad. Masters then update their lists to indicate they now have higher priority than the winner. The

winner becomes the lowest priority on all lists. This arbitration scheme guarantees fair access to the bus by all masters.

### Electrical Design

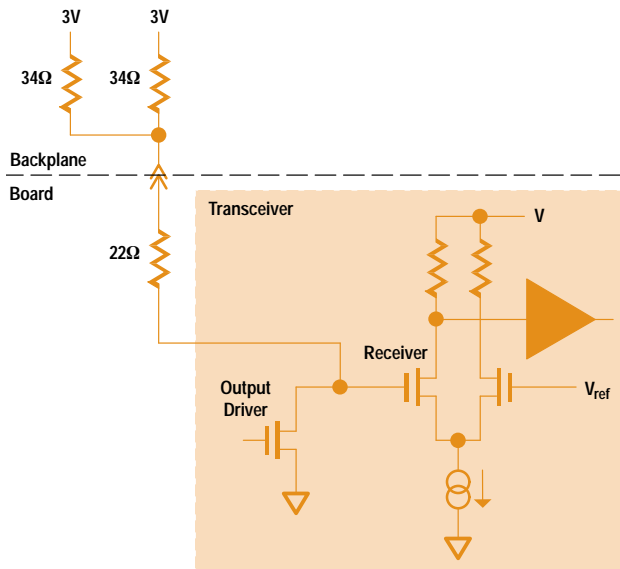
The processor memory bus has the somewhat conflicting goals of high connectivity (which implies a long bus length) and high bandwidth (which implies a high frequency of operation and a correspondingly short bus length). A typical solution to these goals might use custom transceivers and operate at a frequency of 40 MHz. However, by using custom VLSI, incident wave switching, and state-of-the-art design, the Model T500 processor memory bus allows reliable operation at 60 MHz over a 13-inch bus with 16 cards installed.

Each board on the processor memory bus uses two types of custom bus interface transceiver ICs. The first IC type incorporates 10 bits of the processor memory bus (per package), error detection and correction logic, and two input ports (with an internal 2:1 multiplexer) in one 100-pin quad flat package. This IC is referred to as a processor memory bus transceiver in this article. The second IC type performs all of the above duties but adds arbitration control logic and control of 20 bits on the processor memory bus in a 160-pin quad flatpack. This IC is referred to as an arbitration and address buffer in this article. The arbitration and address buffer and the processor memory bus transceivers are implemented in HP's 0.8-micrometer CMOS process.

Fig. 6 shows the basic processor memory bus design. Each processor memory bus signal line has a 34-ohm termination resistor tied to 3V at each end of the bus. Each card installed on the processor memory bus has a series terminating resistor of 22 ohms between the connector and a corresponding bidirectional buffer transceiver for the processor memory bus.

For asserted signals (active low) the output driver transistor in Fig. 7 turns on. This pulls the 22-ohm resistor to approximately ground which (through the resistor divider of 22 ohms and two 34-ohm resistors in parallel) pulls the processor memory bus signal to approximately 1.6 volts. On deasserted signals the output driver is off and the 34-ohm resistors at each end of the bus pull the bus to a high level of approximately 3V.

The receiver (a greatly simplified version is shown in Fig. 7) is a modified differential pair. One input of the differential pair is connected to an external reference voltage of 2.55V. The other input of the differential pair is connected to the



**Fig. 7.** Processor memory bus electrical detail.

processor memory bus. Use of a differential pair receiver allows incident signal switching (i.e., the first transition of the signal is detected by the receiver) and precise level control of the input switch point.

The 22-ohm series resistor performs several important functions. First, when a transceiver asserts a signal, the resistor limits pull-down current. Second, for boards where the transceiver is not driving, the 22-ohm resistor helps isolate the processor memory bus from the capacitive and inductive load presented by the inactive buffers and board traces. Lastly, the 22-ohm resistor helps dampen transient waveforms caused by ringing.

## Processor Board

The processor board used in the Model T500 is a hardware performance upgrade product that replaces the original processor board of the HP 9000 Model 890 corporate business server. With up to two processor modules per processor board, the dual processor board allows the Model T500 system to achieve up to twelve processor systems. Additionally, the use of the PA-RISC 7100 processor improves uniprocessor performance. The key features of the Model T500's processor include:

- Direct replacement of the original processor board (cannot be mixed with original processor boards in the same system).
- Increased multiprocessing performance with support for one to twelve CPUs.
- Processor modules based on 90-MHz PA-RISC 7100 CPU chip<sup>8</sup> with on-chip floating-point coprocessor for higher uniprocessor integer and floating-point performance.
- Processor modules that allow single-processor and dual-processor configurations per processor slot. Easy field upgrade to add a second processor module to a single processor module board.
- Processor clock frequency 90 MHz, processor memory bus clock frequency 60 MHz.
- 1M-byte instruction cache (I cache) and 1M-byte data cache (D cache) per module.

## Performance Improvement

Relative to its predecessor, the Model T500's processor board SPEC integer rate is improved by a factor of 1.9 times and the SPEC floating-point rate is improved by a factor of 3.4 times. The Model T500's processor performance relative to its predecessor is shown in the table below.

	<b>Model T500</b>	<b>Model 890</b>
CPU Clock	90 MHz	60 MHz
Bus Clock	60 MHz	60 MHz
Cache Size (per CPU)	1M-byte I cache, 1M-byte D cache (direct mapped)	2M-byte I cache, 2M-byte D cache (2-way set- associative)
SPECint92	98.3	Not Published
SPECfp92	170.2	Not Published
SPECrate_int92 (1 CPU)	2310	1215
SPECrate_int92 (2 CPUs)	4609	2253
SPECrate_int92 (4 CPUs)	9017	4301
SPECrate_int92 (8 CPUs)	17114	N/A
SPECrate_int92 (12 CPUs)	23717	N/A
SPECrate_fp92 (1 CPU)	4019	1180
SPECrate_fp92 (2 CPUs)	7963	2360
SPECrate_fp92 (4 CPUs)	15341	4685
SPECrate_fp92 (8 CPUs)	28341	N/A
SPECrate_fp92 (12 CPUs)	38780	N/A
tpsA	Not Published	710.43 tpsA at U.S.\$8,258 per tpsA
tpmC	2110.5 tpmC at U.S.\$2,115 per tpmC	Not Published

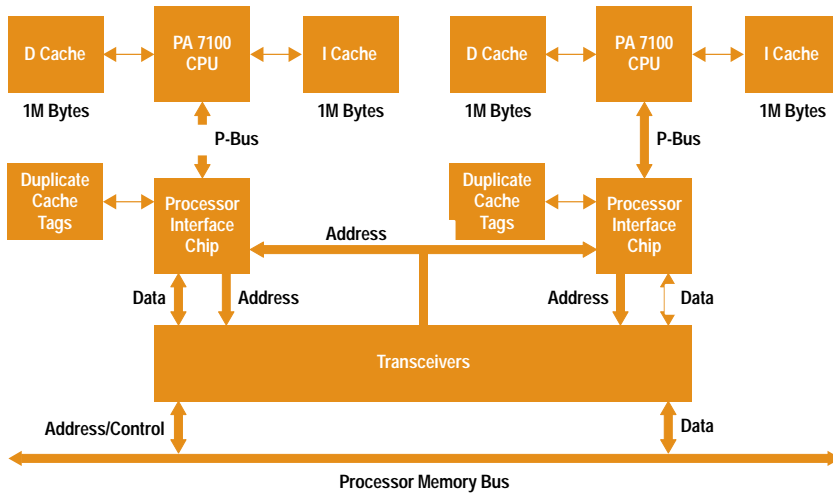
## Hardware Overview

The Model T500's processor board consists of one or two processor modules, a set of 12 processor memory bus transceivers (4 address and 8 data bus transceivers), an arbitration and address buffer, two processor interface chips, two sets of duplicate tag SRAMs, ECL clock generation circuitry, four on-card voltage regulators, scan logic circuitry, connectors, a printed circuit board, and mechanical hardware. Fig. 8 shows the processor board hardware block diagram. Fig. 9 is a photograph of a processor board with two processor modules.

## Processor Modules

The processor board is centered around two identical, removable processor modules based on the HP PA 7100 CPU chip. Each module consists of a CPU chip, 26 SRAMs which make up the 1M-byte instruction cache (I cache) and 1M-byte data cache (D cache), a 4.1-inch-by-4.4-inch 12-layer printed circuit board, and a 100-pin P-bus connector.

Each processor module communicates with its processor interface chip through a 60-MHz, 32-bit multiplexed address/data bus called the P-bus. Each module has a dedicated P-bus. The P-bus has 35 data and address lines and 18 control lines.



**Fig. 8.** Processor board organization.

The I cache and D cache each have the following features:

- 64-bit access (I cache 64-bit double-wide word, D cache two 32-bit words)
- Direct mapped with a hashed address and virtual index
- Bandwidth up to 520 Mbytes/s
- I and D cache bypassing
- Parity error detection in both I and D caches (parity errors in the I cache cause a refetch of the offending instruction)
- 32-byte cache line size.

The CPU chip has the following features:

- Level 1 PA-RISC implementation with 48-bit virtual addressing
- Addresses up to 3.75G bytes of physical memory
- Multiprocessor cache coherency support
- TLB (translation lookaside buffer)
  - 120-entry unified instruction and data TLB
  - Fully associative with NUR (not used recently) replacement
  - 4K page size
- Floating-point coprocessor
  - Located on-chip
  - Superscalar operation
  - Multiply, divide, square root
  - Floating-point arithmetic logic unit (FALU)
- P-bus system interface (to bus interface chip)

- Serial scan path for test and debug
- Operation from dc to 90 MHz
- Performance improvements
  - Load and clear optimizations
  - Hardware TLB miss handler support
  - Hardware static branch prediction
- 504-pin interstitial pin-grid array package.

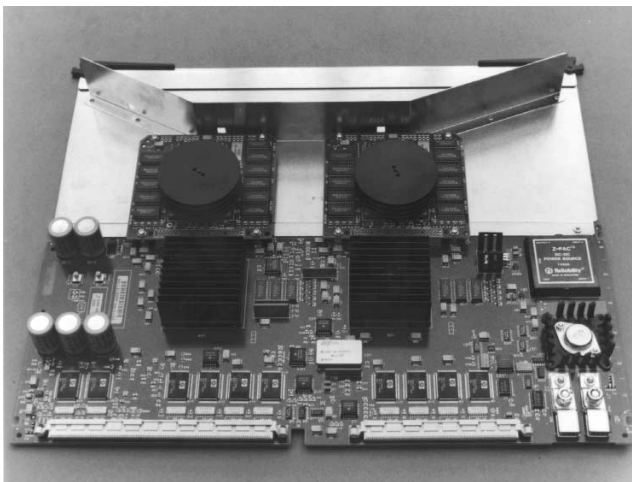
#### Processor Interface Chip

Each processor interface chip transmits transactions between its CPU and the rest of the system (memory, I/O, and other processors) via the processor memory bus. The processor interface chip for each processor module interfaces its CPU (through the P-bus) to the processor memory bus transceivers and the arbitration and address buffer. The CPU's line size is 32 bytes, so the processor interface chip provides a 64-bit data interface to the processor memory bus transceivers. The two processor interface chips communicate through separate ports on the processor memory bus transceivers, which provide the required multiplexing internally.

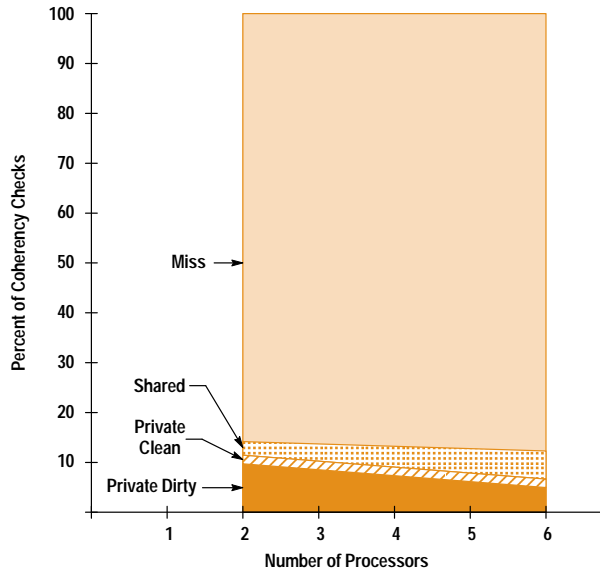
Each processor interface chip also contains an interface that allows it to communicate with self-test, processor dependent code (boot and error code), and processor dependent hardware (time-of-day clock, etc.) on the service processor board. The processor interface chip is implemented in HP's 0.8-micrometer CMOS process and is housed in a 408-pin pin-grid array package.

The processor interface chip has two features to enhance the multiprocessor performance of the system: duplicate data cache tags and coherent write buffers. The coherent buffers support the processor memory bus's multiprocessor implementation of cache coherence protocol.

**Duplicate Data Cache Tags.** The interface chip maintains its own duplicate copy of the CPU's data cache tags in off-chip SRAMs. The tags contain the real address of each cache line and the valid and private bits (but not the dirty bit). The duplicate cache tags are kept consistent with the CPU's data cache tags based only on the transactions through the interface chip. The duplicate tags allow the interface chip to signal the status of a cache line during a coherent transaction without querying the processor (which would require a pair of transactions on the P-bus). Measurements (using the processor interface chip's built-in performance counters) for a



**Fig. 9.** Model T500 processor board with two processor modules.



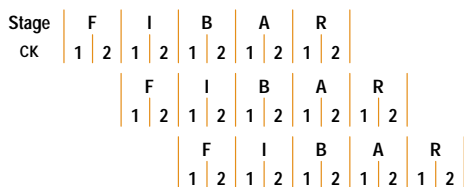
**Fig. 10.** Data sharing in a multiprocessor Model 890 system as measured by results of cache coherency checks.

wide variety of benchmarks show that for 80 to 90 percent of coherent transactions, the cache line is not present in a third-party CPU's data cache, as shown in Fig. 10. The duplicate tags increase system performance to varying degrees for different workloads. Measurements on a four-processor system show duplicate tags increase system throughput by 8% for a CPU-intensive workload and 21% for a multitasking workload.

**Coherent Write Buffers.** To isolate the CPU from traffic on the bus, the interface chip contains a set of five cache line write buffers. The buffers are arranged as a circular FIFO memory with random access. If the CPU writes a line to memory, the interface chip stores the line in one of its buffers until it can win arbitration to write the line to memory. While the line is in a buffer, it is considered part of the CPU's cached data from the system bus point of view and participates in coherence checking on the bus. These buffers are also used for temporary storage of data sent from the cache as a result of a coherency check that hits a dirty cache line. By having many buffers, the interface chip is able to handle multiple outstanding coherency checks.

### Pipeline

The PA 7100 pipeline is a five-stage pipeline. One and a half stages are associated with instruction fetching and three and a half stages are associated with instruction execution. The PA 7100 also has the ability to issue and execute floating-point instructions in parallel with integer instructions. Fig. 11 shows the CPU pipeline.



**Fig. 11.** CPU pipeline.

Instruction fetch starts in CK1 of stage F and ends in CK1 of stage I. For branch prediction, the branch address is calculated in CK1 of I and completes by the end of CK2 of I. This address is issued to the I cache.

From CK2 of I to CK1 of B, the instruction is decoded, operands are fetched, and the ALU and SMU (shift merge unit) produce their results. The data cache address is generated by the ALU by the end of CK1 of B. For branch prediction, the branch address is calculated in CK1 of B and completes by the end of CK2 of B.

Data cache reads start in CK2 of B and end in CK2 of A. Load instructions and subword store instructions read the data portion of the D cache during this stage. For all load and store instructions the tag portion of the D cache is read during this stage. The tag portion of the D cache is addressed independently from the data portion of the D cache so that tag reads can occur concurrently with a data write for the last store instruction. Branch condition evaluation is completed by the end of CK2 of B.

The PA 7100 CPU maintains a store buffer which is set on the cycle after CK2 of A of each store (often CK2 of R). General registers are set in CK2 of R. The store buffer can be written to the D cache starting on CK2 of R and continuing for a total of two cycles. The store buffer is only written on CK2 of R when one of the next instructions is a store instruction. Whenever the next store instruction is encountered, the store buffer will be written out to the cache.

### Clock Generation

The clock generation circuitry provides 60-MHz and 90-MHz differential clock signals to the processor memory bus interface ports and the processor modules, respectively. The Model T500's processor board uses a hybrid phase-locked loop component developed especially for the Model T500. The phase-locked loop generates a synchronized 90-MHz processor clock signal from the 60-MHz processor memory bus clock. Clock distribution is by differential ECL buffers with supplies of +2.0V and -2.5V. The use of offset supplies for the ECL allows optimal termination with the 50-ohm termination resistors tied directly to ground, and allows clock signal levels to be compatible with the CMOS clock receivers.

There is no system support for halting clocks, or for single-stepping or n-stepping clocks. The scan tools do, however, allow halting clocks within each of the scannable VLSI chips.

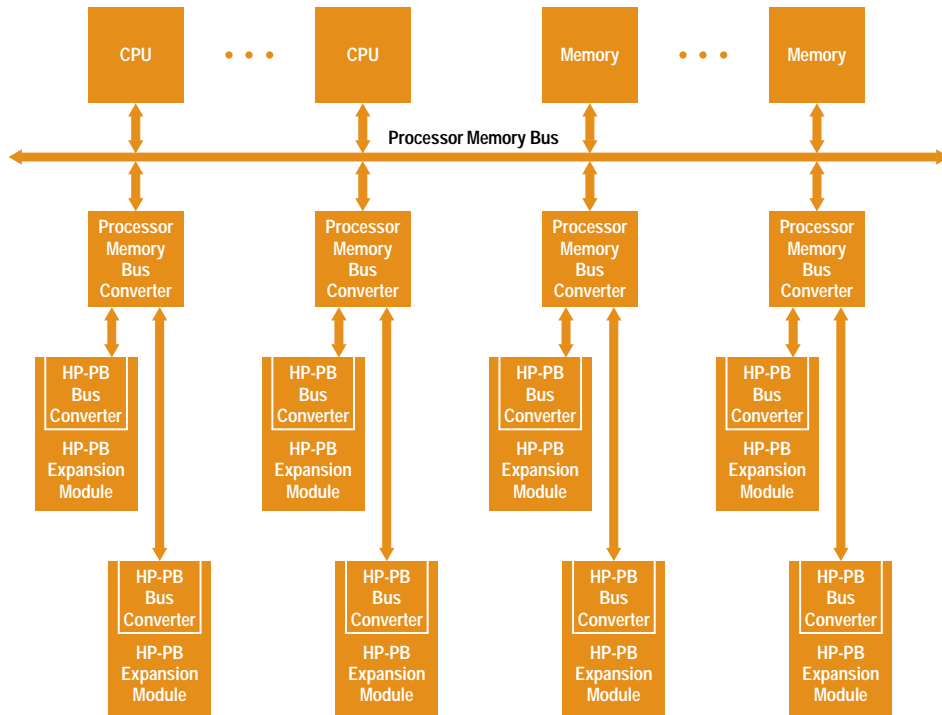
### Scan Circuitry

The processor board's scan circuitry interfaces to the service processor's four-line serial scan port and enables the user, via the service processor, to scan test each of the VLSI chips and transceiver groups selectively. The arbitration and address buffer chip can be scanned independently, whereas the address (4) and data (8) bus transceivers are chained. This scan feature is used as a fault analysis tool in manufacturing.

### Printed Circuit Board and Mechanical

The processor board uses a 12-layer construction and has an approximate overall thickness of 0.075 inch. Among the 12 layers are six signal layers, three ground layers, and three voltage plane layers. Cyanate ester dielectric material is used





**Fig. 12.** Model T500 I/O subsystem.

for its faster signal propagation speed over FR-4 material and its ability to achieve reduced board thickness for a given trace impedance. The nominal signal trace impedance is 51 ohms for all high-speed signal nets.

Every attempt was made to keep high-speed signal traces closely coupled to a neighboring ground layer to minimize signal perturbations and EMI. Bypass capacitors are distributed liberally across the board to suppress high-frequency noise. EMI decoupling techniques consistent with the other Model T500 boards are used to direct common-mode noise to chassis ground.

The dimensions of the processor board are 16.90 inches by 7.35 inches. The two processor modules extend beyond the 7.35-inch dimension by approximately 3.25 inches and are supported by a sheet-metal extender which effectively makes the board assembly 14 inches deep. The modules are mounted parallel to the processor board and the sheet-metal extender and are secured by screws and standoffs. The sheet-metal extender also has a baffle which directs forced air across the modules for increased cooling.

## Input/Output Subsystem

The HP 9000 Model T500 represents a major advance in the areas of high I/O throughput and highly scalable connectivity. The Model T500 system provides large aggregate I/O throughput through the replication of input/output buses with large slot counts. These I/O buses are arranged in a two-level tree. A bus converter subsystem connects the processor memory bus of the Model T500 system with the Hewlett-Packard precision bus (HP-PB) I/O buses, as shown in Fig. 12. The bus converter subsystem consists of a processor memory bus converter, a bus converter link (see Fig. 13), and an HP-PB bus converter. It translates the logical protocol

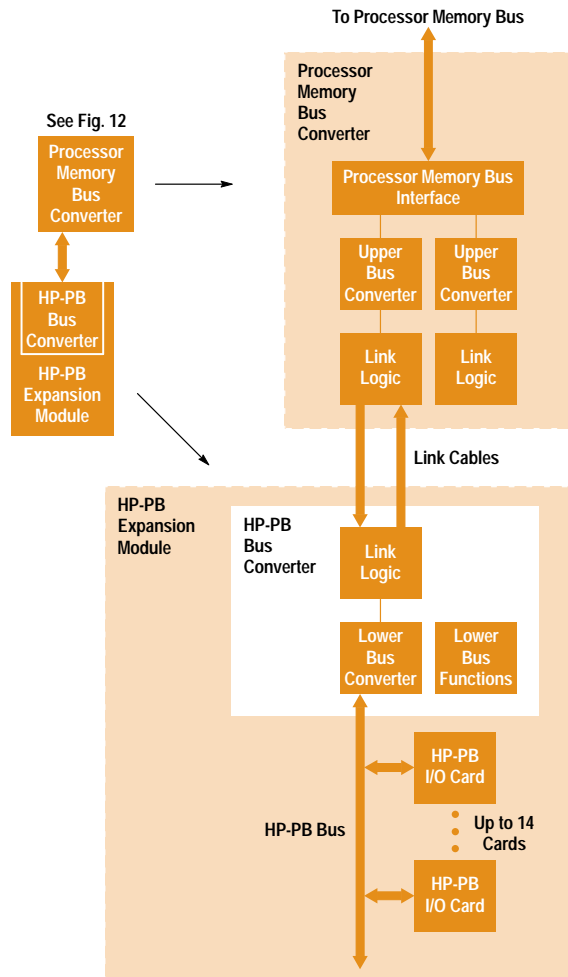
and electrical signaling of data transfers between the processor memory bus and the I/O cards on the HP-PB bus.

The I/O subsystem guarantees data integrity and provides high reliability through parity protection of all data and transactions and through the hardware capability of online replaceable cards.

The bus converter subsystem is transparent to software under normal operating conditions. Each I/O module on an HP-PB bus in the system is assigned a range of physical memory addresses. I/O modules appear to software as sets of registers.

All modules can be DMA capable and generally implement scatter/gather DMA controllers. These scatter/gather DMA controllers allow virtually contiguous data located in physically noncontiguous pages to be transferred with minimal CPU assistance. A chain of DMA commands is written into memory by the processor. The I/O card is notified of the location of the chain and that it is ready for use. The I/O card then uses the scatter/gather DMA controller to follow the chain and execute the commands. In this manner the I/O card can write data (scatter) to different physical pages during the same DMA operation. The I/O card can also read data (gather) from different physical pages during the same DMA operation. When the I/O card finishes all of the commands in the chain, it notifies the processor, usually through an interrupt.

The processor memory bus converter is a dual bus converter that connects to two HP-PB buses through a pair of cables and the HP-PB bus converter. The HP-PB bus converter is plugged into a slot in an HP-PB expansion module and provides the central HP-PB bus resources of arbitration, clock generation, and online replacement signals in addition to the connection to the processor memory bus.



**Fig. 13.** Detail of Model T500 I/O subsystem.

Each HP-PB expansion module is a 19-inch rack-mountable assembly that connects any combination of up to 14 single-height or 7 double-height cards to the HP-PB bus. A Model T500 supports connection of 112 single-height HP-PB cards.

Each HP-PB bus is a 32-bit multiplexed address and data bus with byte-wise parity protection and additional parity protection across the control signals. The frequency of operation is fixed at 8 MHz, leading to a peak bandwidth of 32 Mbytes/s. The aggregate I/O rate for the Model T500 system is thus 256 Mbytes/s.

The HP-PB I/O function cards include SCSI, fast/wide SCSI, FDDI (doubly connected), Ethernet LAN, token ring LAN, HP-FL fiber-link disk connect, IEEE 488 (IEC 625), X.25 and other WAN connects, terminal multiplexer cards, and other I/O functions. Using HP-FL cards and HP C2250A disk arrays, the corporate business server hardware can support over 1.9 terabytes of disk storage on over 1000 disk spindles.

### Processor Memory Bus Converter

The Model T500 accepts up to four processor memory bus converters plugged into the processor memory bus backplane. Each processor memory bus converter consists of two logically separate upper bus converter modules sharing a single bus interface (see Fig. 13). This reduces the electrical loading on the processor memory bus while providing the necessary fanout for a high-connectivity I/O subsystem.

The processor memory bus converter provides resource-driven arbitration and transaction steering on the processor memory bus for transactions involving the I/O subsystem. The processor memory bus converter provides a maximum bandwidth of 96 Mbytes/s. Transactions through the processor memory bus converter are parity protected, and error correcting code is generated and checked at the processor memory bus interface to guarantee data and transaction integrity.

The upper bus converter modules are implemented in custom CMOS26 VLSI chips in 408-pin pin-grid array packages. They arbitrate with each other for the processor memory bus interface chips on the processor memory bus side and implement the bus converter link protocol on the link side.

The processor memory bus interface consists of 12 bus transceiver chips (eight data and four address) and an arbitration and address buffer chip. These chips are used in a two-module mode. The data bus transceivers drive independent bidirectional data buses to the two upper bus converter module chips. The address bus transceivers drive a single unidirectional address to both bus converter chips, but receive independent address buses from the two upper bus converter chips.

The processor memory bus converter also provides discrete industry-standard logic to translate the bus converter link signals between the bus converter link signals between the CMOS levels of the upper bus converter chip and the +5V ECL levels of the link cable.

### Bus Converter Link

Each of the two upper bus converter modules connects through two cables to a lower bus converter module, the

HP-PB bus converter (see Fig. 13). Each cable is a high-performance 80-conductor flat ribbon insulation displacement connector cable which allows the lower bus converter module and the HP-PB expansion module to be located up to 10 meters away. These cables and the protocol that is used on them make up the bus converter link.

The bus converter link protocol is a proprietary protocol allowing pipelining of two transactions with positive acknowledgment. The signals are point-to-point +5V ECL differential signals, two bytes wide and parity protected. The status information from the opposite bus converter module is embedded in the link protocol. The signaling rate across the bus converter link is one-half the processor memory bus frequency or 30 MHz in the Model T500 system. The peak bus converter link bandwidth is therefore 60 Mbytes/s with an average protocol overhead of 10%. The address overhead is on the order of 20% leaving an average data transfer rate of 42 Mbytes/s.

### HP-PB Bus Converter

The HP-PB bus converter connects the bus converter link to the HP-PB bus in the HP-PB expansion module. In addition to the bus converter functions, the HP-PB bus converter provides the central resources for the HP-PB bus to which it connects, including bus clock generation, arbitration logic and online replacement power-on signals. The bus clock generation and arbitration are performed by discrete industry-standard components on the board. The HP-PB bus converter functions are implemented in a custom CMOS26 chip in a

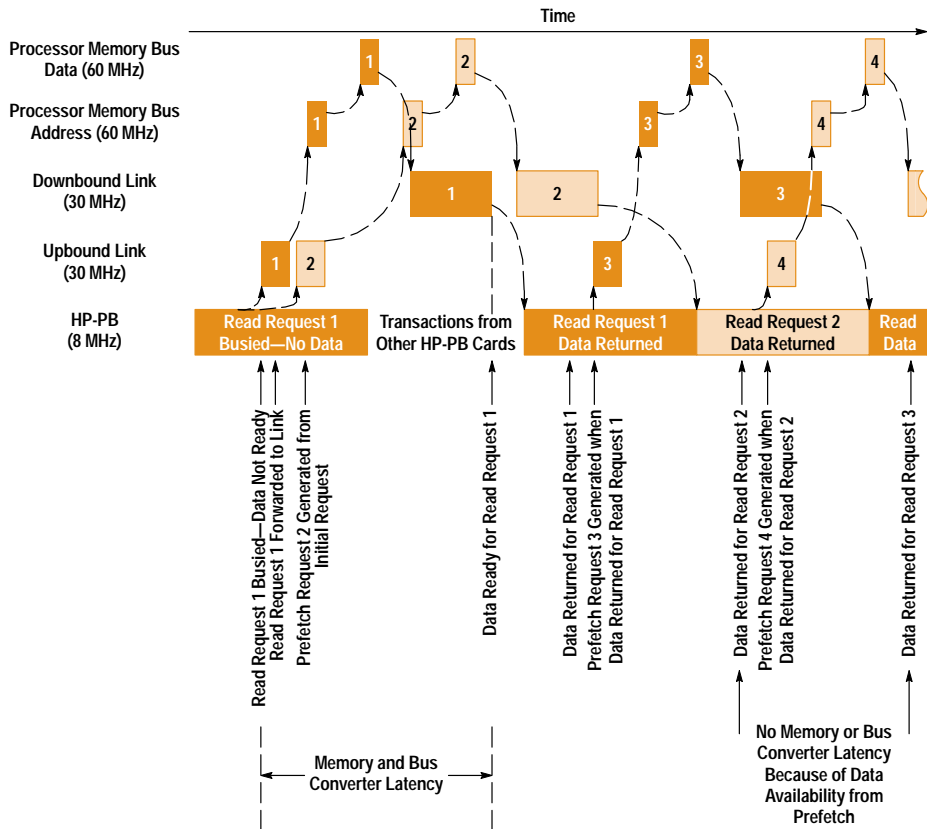


Fig. 14. Speculative prefetch.

272-pin pin-grid array package. Electrical signal level translation between the CMOS of the lower bus converter chip and the +5V ECL of the link cable is performed using the same discrete industry-standard components as are used on the processor memory bus converter. The HP-PB bus converter acts as a concentrator for the I/O traffic from the HP-PB cards bound for the system memory or the processors.

The HP-PB bus converter implements a speculative prefetch for DMA reads of memory by HP-PB cards (data transferred from memory to an I/O device under the I/O card's control). This provides greater performance by offsetting the transaction and memory latency. The prefetch algorithm always has two read requests in the transaction pipeline to memory (see Fig. 14). When a read transaction to memory is accepted for forwarding by the HP-PB bus converter, it forwards the first read and then issues a second read request with the address incremented by the length of the original read transaction. As the data is returned to the requester, a new read transaction with the address incremented by twice the length of the transaction is issued on the bus converter link. The prefetching stops when the I/O card does not request the next read in the next transaction interval on the HP-PB bus or when the address generated would cross a 4K page boundary. Speculative prefetch increases the possible read data bandwidth from 3 Mbytes/s to over 18 Mbytes/s.

The HP-PB bus converter supports DMA writes at the full HP-PB data bandwidth of 18 Mbytes/s for 16-byte writes and 23 Mbytes/s for 32-byte writes. The difference between the

peak bandwidth and the data bandwidth represents the effects of the address overhead and bus turnaround cycles.

The HP-PB bus converter carries parity through the entire data path and checks the parity before forwarding any transaction onto the link or the HP-PB bus to guarantee data and transaction integrity.

The HP-PB bus converter and HP-PB backplane in the HP-PB expansion module together provide the hardware and mechanisms to allow online replacement of HP-PB I/O cards. The HP-PB bus converter provides a read/write register through which the power-on signal to each HP-PB card can be controlled independently. When this signal is deasserted to an HP-PB card, the card's bus drivers are tristated (set to a high-impedance state) and the card is prepared for withdrawal from the HP-PB expansion module. The HP-PB backplane provides the proper inductance and capacitance for each slot so that a card can be withdrawn while the system is powered up without disturbing the power to the adjacent cards. The hardware online replacement capability makes possible future enhancements to the Model T500 for even higher availability.

Logic in the HP-PB expansion module monitors the ac power into the module and indicates to the HP-PB bus converter via a backplane signal when power is about to fail or when the dc voltages are going out of specification. The powerfail warning signal is passed up through the bus converter modules to allow the Model T500 system to prevent corruption of the machine state.

## HP Precision Bus

The HP-PB is a multiplexed 32-bit address and data bus with a fixed clock rate of 8 MHz. The HP-PBs in the Model T500 system are completely independent of the processor memory bus clocks. The HP-PB bus converter synchronizes the data between the HP-PB and the bus converter link.

The HP-PB provides for global 32-bit addressing of the I/O cards and for flexibility in address assignment. Each HP-PB is allocated a minimum of 256K bytes during configuration. This address space is evenly divided between 16 possible slots. Each slot on the HP-PB supports up to four I/O modules, each of which is allocated a 4K-byte address space. This 4K-byte space is called the hard physical address space. Any module that requires additional address space is assigned address space at the next available bus address. This additional address space is called the soft physical address space. Soft physical address space assigned to all I/O modules on a single HP-PB is contiguous. The processor memory bus converter determines if a transaction is bound for a given HP-PB by checking for inclusion in the range determined by the hard physical address and soft physical address space of the HP-PB.

The hard physical address of an I/O card contains the control and status registers defined by the PA-RISC architecture through which software can access the I/O card. Each HP-PB card has a boot ROM called the I/O dependent code ROM, which is accessed by indirection through a hard physical address. This ROM contains the card identification, configuration parameters, test code, and possibly boot code. The I/O dependent code ROM allows I/O cards to be configured into a system before the operating system is running and allows the operating system to link to the correct driver for each card.

The HP-PB transaction set is sufficiently rich to support efficient I/O. There are three classes of transactions: write, read, and clear or semaphore. Each transaction is atomic but the HP-PB bus protocol provides buffered writes for high performance and provides a busy-retry capability to allow reads of memory to be split, providing parallelism and higher bandwidth. Each HP-PB transaction specifies the data payload. The transaction set supports transactions of 1, 2, 4, 16, and 32 bytes. DMA is performed using 16-byte or 32-byte transactions initiated under the control of the I/O card. Each HP-PB transaction contains information about the master of the transaction so that errors can be reported and data easily returned for reads.

The HP-PB and I/O subsystem provides an efficient, flexible, and reliable means to achieve high I/O throughput and highly scalable connectivity.

## Memory System

The memory subsystem for the HP 9000 Model T500 corporate business server uses 4M-bit DRAMs for a 256M-byte capacity on each board. It is expandable up to 2G bytes of error-correcting memory. To minimize access latency in a multiprocessor environment, the memory subsystem is highly interleaved to support concurrent access from multiple processors and I/O modules. A single memory board can

contain 1, 2, or 4 interleaved banks of 64M bytes. The combination of interleaving and low latency for the board provide a bandwidth of 960 Mbytes/s. Furthermore, different-sized memory boards using different generations of DRAMs can coexist in the system, allowing future memory expansion while preserving customer memory investments.

From the standpoint of complexity, the memory board is the most sophisticated board in the Model T500 system. To meet its performance requirements, the design uses leading-edge printed circuit technologies and new board materials. These are described under "Manufacturing" later in this article. The memory board includes 4273 nets (or signals), 2183 components, and over 28,850 solder joints. Double-sided surface mount assembly provides high component density. The 2183 components are mounted in an area of only 235 square inches.

The processor memory bus electrical design limits the length of the bus for 60-MHz operation to 13 inches. Consequently, the memory board design is considerably constrained. The limited number of slots requires the capacity of each memory board to be high. The short bus length makes each of the slots narrow, forcing a low profile for each memory board. Bus transceivers are located close to the connector on each daughter card to keep stub lengths to a minimum.

## Memory Interleaving

Memory boards are manufactured in 64M-byte, 128M-byte, and 256M-byte capacities. The 64M-byte and 128M-byte memory capacities are achieved by partially loading the 256M-byte board. Memory interleaving tends to distribute memory references evenly among all blocks in the system. In the event that two processors desire to access memory in consecutive quads, interleaving provides that the second access will likely be to an idle bank. The memory design for the Model T500 allows the benefits of interleaving to be based on the total number of memory banks installed in the system, regardless of the number of boards that the banks are spread across.<sup>9</sup> The processor memory bus protocol maximizes performance by interleaving all the banks evenly across the entire physical address space, regardless of the number of banks. This is superior to interleaving schemes that limit the effect of interleaving to numbers of banks that are powers of two.

## Memory Board Partitioning

Partitioning of the memory board into VLSI chips follows the requirements of the DRAMs and the bank organization. This partitioning is illustrated in the memory board block diagram, Fig. 15. 256M-byte capacity with single-bit error correction requires 576 4M-bit DRAMs, each of which is organized as 1M by 4 bits. 64-byte data transfers and minimized latency require a 576-bit bidirectional data bus for each bank's DRAMs. The effort to minimize latency and the restriction of the processor memory bus to narrow slots prevented the use of SIMM modules similar to those used in PCs and workstations. The fixed timing relationships on the processor memory bus required that there be four of these 576-bit data buses for the four banks on the 256M-byte memory board to prevent contention between writes to one bank and reads from another bank. A multiplexing function is provided

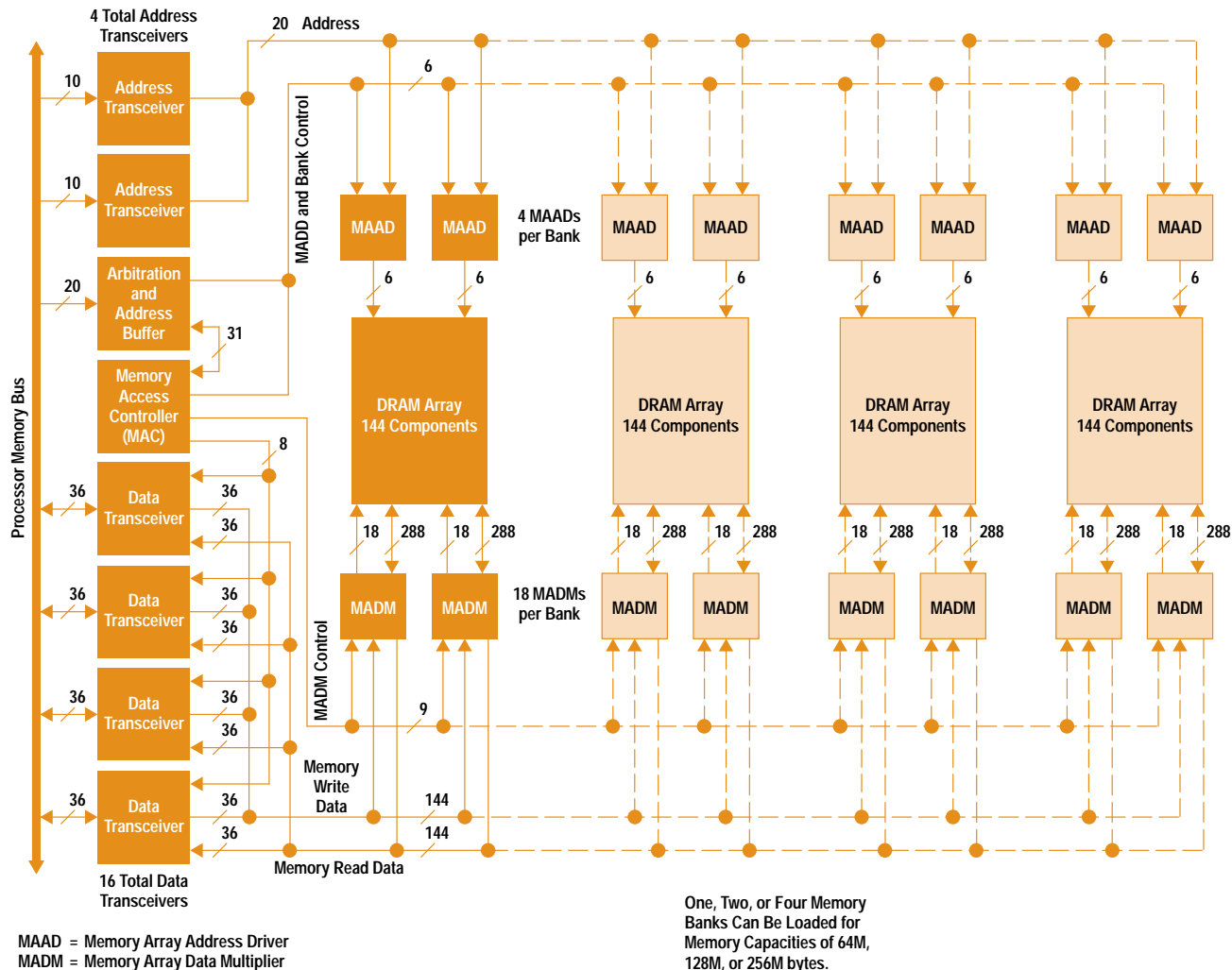


Fig. 15. Memory board block diagram.

between the four slow 576-bit DRAM data buses and the 60-MHz 128-bit data bus of the processor memory bus.

To implement these requirements, a set of five VLSI chips is used. As identified on the block diagram, these are:

- Bus transceivers. This design is also used on the processor and bus converter boards.
- Arbitration and address buffer. This chip provides for arbitration and acts as an additional pair of address transceivers. This design is also used on the processor and bus converter boards.
- Memory array data multiplexer (MADM). This chip multiplexes the slow DRAM data signals to a pair of unidirectional 60-MHz, 128-bit buses to and from the data transceivers.
- Memory array address driver (MAAD). This chip drives address and RAS and CAS to the DRAMs. It is a modified version of a standard commercial part.
- Memory access controller (MAC). This chip provides the overall control function for the memory board. In particular, the MAC implements the required architectural features of the memory system and controls DRAM refresh.

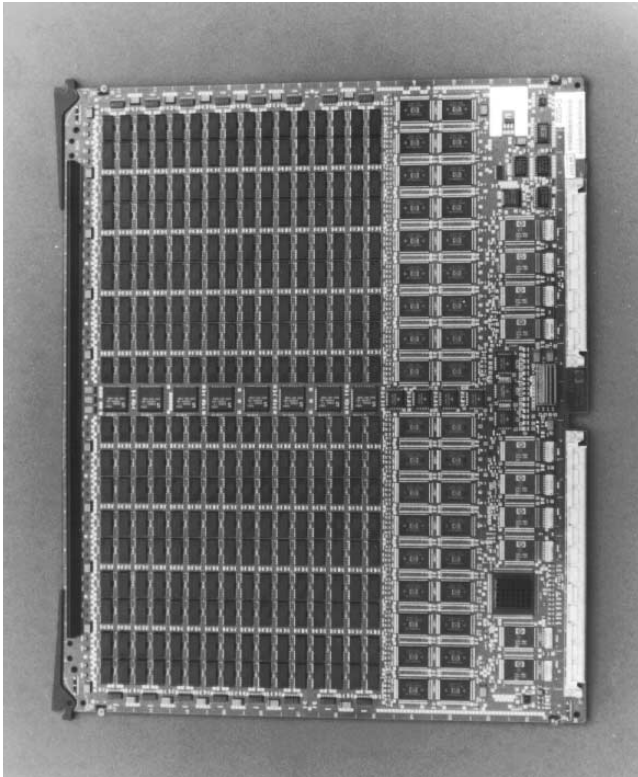
Except for the MAAD, which is in a 44-pin PLCC (plastic leaded chip carrier), each of these ICs is a fine-pitch, quad flatpack (QFP) component, with leads spaced 0.025 inch apart. The bus transceiver and MADM are packaged in

100-pin QFPs and the arbitration and address buffer and MAC are in 160-pin QFPs. The full 256M-byte board includes 20 bus transceivers, one arbitration and address buffer, 72 MADMs, 16 MAADs, and one MAC as well as the 576 4M-bit DRAM chips.

Fig. 16 is a photograph of the 256M-byte memory board.

### Printed Circuit Board Design

In addition to restrictions on the memory board caused by the processor memory bus design, there were a significant number of other electrical design and manufacturing requirements on the board. The onboard version of the processor memory bus address bus is a 31.70-inch, 60-MHz unidirectional bus with 16 loads on each line. There are two 128-bit, 60-MHz, 9.15-inch buses with five loads on each line. With the large number of components already required for the board, it would not have been feasible to terminate these buses. The clock tree for the VLSI on the board feeds a total of 94 bidirectional shifted ECL-level inputs and 16 single-ended inputs, with a goal of less than 250 ps of skew across all 110 inputs. The size chosen for the memory board is 14.00 by 16.90 inches, the maximum size allowed by surface mount equipment for efficient volume production. Restriction

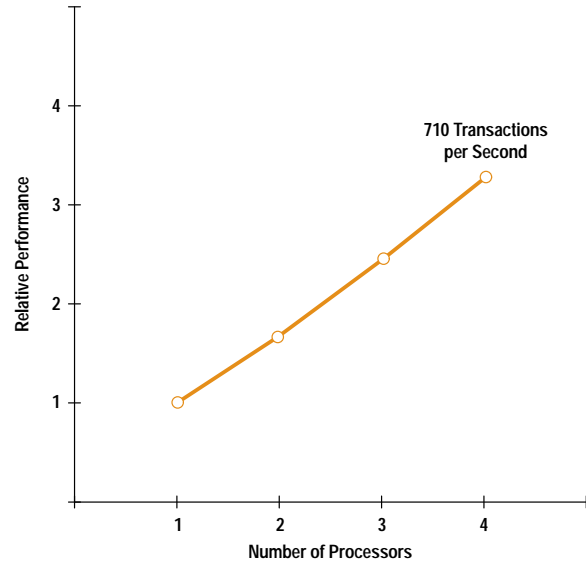


**Fig. 16.** 256M-byte memory board.

to this size was an important factor in almost every design decision made for the board.

Preliminary designs of critical areas of the board showed that the densest feasible routing would be required. Leading-edge HP printed circuit production technology allows a minimum of 0.005-inch lines and 0.005-inch spaces. Vias can be either 0.008-inch finished hole size with 0.021-inch pads, or 0.012-inch finished hole size with 0.025-inch pads. Both of these alternatives are currently limited to a maximum aspect ratio of 10:1 (board thickness divided by finished hole size). The aspect ratio also influences the production cost of the board significantly because of plating yields, as well as the achievable drill stack height.

With the given layout conditions, several trade-off studies were done to find the best alternative in terms of electrical performance, manufacturing cost for the loaded assembly, reliability, and risk for procurement and process availability at both fabrication and assembly. The best alternative finally uses the leading-edge layout geometries, eight full signal layers, and two partial signal layers. Since the initial projections of the number of layers required to route the board led to an anticipated board thickness greater than 0.080 inch, the aspect ratio requirements caused the 0.008-inch finished hole size via option to be rejected. Even with 0.025-inch pads and 0.012-inch finished hole size vias, the aspect ratio approaches 10. Therefore, a sophisticated board material is required to prevent thermal cycling from stressing vias and generating distortions on the board by expansion of the thickness of the board. Cyanate ester material (HT-2) was chosen over other substrate alternatives because of its superior electrical and mechanical performance.<sup>10</sup>



**Fig. 17.** Scaling of online transaction processing (OLTP) performance with number of processors.

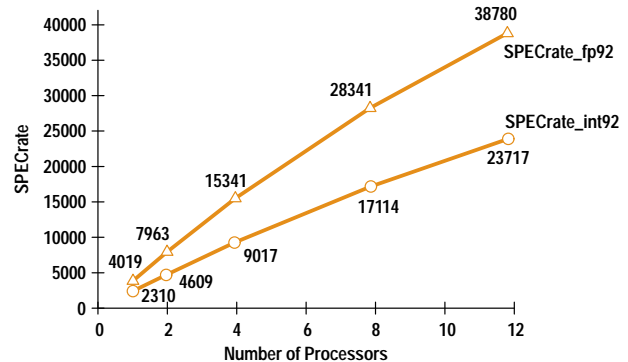
## Multiprocessor Performance

### Performance

An HP 9000 Model T500 corporate business server, a six-processor, 90-MHz PA-RISC 7100 CPU with a 60-MHz bus, achieved 2110.5 transactions per minute (U.S.\$2,115 per tpmC) on the TPC-C benchmark.<sup>5</sup> In the following discussions, the available multiprocessor performance data is a mixture of data from both the Model T500 and the older Model 890 systems.

Data for the HP 9000 Model 890 (the precursor of the Model T500, which uses one to four 60-MHz PA-RISC processors and the same memory, bus, and I/O subsystems as the Model T500) is available for the TPC-A benchmark and one to four processors. Fig. 17 shows how multiprocessing performance scales on a benchmark indicative of OLTP performance.<sup>6</sup>

The SPECrate performance for the Model T500 is shown in Fig. 18.<sup>7</sup> The SPEC results show linear scaling with the number of processors, which is expected for CPU-intensive workloads with no mutual data dependencies. The OLTP benchmarks are more typical for real commercial applications. The



**Fig. 18.** Model T500 SPECrate performance.

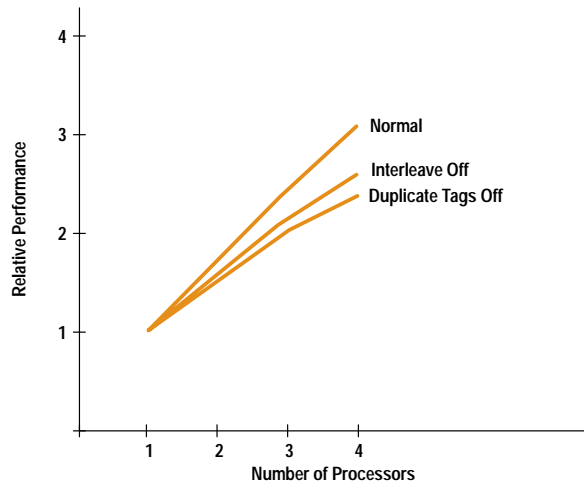


Fig. 19. Model 890 program development performance.

losses in efficiency are caused by factors such as serialization in the I/O subsystem and contention for operating system resources.

Fig. 19 shows the performance of the Model 890 on an HP-internal benchmark representative of 24 interactive users executing tasks typical of a program development environment.

The benchmark results confirm the value of key design decisions. For example, nearly all transactions were useful—only 6% of all transactions were busied and only 1.5% of all bus quads were waited. Disabling the interleaving or disabling the duplicate cache tags did not affect bus utilization.

The efficiency of the bus was reflected in system throughput. Normal operation showed near-linear multiprocessor scaling through four processors. Changing the interleaving algorithm from the normal case of four blocks interleaved four ways to four blocks not interleaved caused a significant performance impact. As expected, the penalty was greater at higher degrees of multiprocessing, peaking at a penalty of 15% in a four-processor system. Disabling the duplicate cache tags incurred an even greater cost: the decrease in system performance was as much as 22%, with the four-processor system again being the worst case.

These tests showed that the high-speed pipelined processor memory bus, fast CPUs with large caches, duplicate cache tags in the processor interfaces, and highly interleaved large physical memory allow the Model T500 system to scale efficiently up to twelve-way multiprocessing.

## Service Processor

As part of the challenge of producing the HP 9000 Model T500 corporate business server, targeted at demanding business applications, it was decided to try to make a significant improvement in system hardware availability. Hardware availability has two components: mean time between failures (MTBF), which measures how often the computer hardware fails, and mean time to repair (MTTR), which measures how long it takes to repair a hardware failure once one has occurred. The service processor makes a significant improvement in the MTTR portion of the availability equation by

reducing the time required to repair the system when hardware failures do occur.

HP's computer systems are typically supported from our response centers, where HP has concentrated some of the most knowledgeable and experienced support staff. These support engineers generally provide the first response to a customer problem. They make the initial problem diagnosis and determine which of HP's resources will be applied to fixing the customer's system. The greatest opportunity to improve the system's MTTR existed in improving the ability of the support engineers at the response centers to access failure information and control the system hardware. The following specific goals were set:

- All of the troubleshooting information that is available locally (at the failed system) should be available remotely (at the response center).
- Information should be collected about hardware failures that prevent the normal operating system code from starting or running.
- Information about power and environmental anomalies should be collected.
- Information about operating system state changes should be collected.
- Error information should be available to error analysis software running under the operating system if the operating system is able to recover after an anomaly occurs.
- A means should exist to allow support personnel to determine the system hardware configuration and alter it without being present at the site to allow problems to be worked around and to aid in problem determination.
- The support hardware should be as independent of the remainder of the computer system as possible, so that failures in the main hardware will not cause support access to become unavailable.
- Error reporting paths should be designed to maximize the probability that failure symptoms will be observable even in the presence of hardware failures.
- Failure in the support hardware should not cause failure of the main computer system.
- Failure of the support hardware should not go unnoticed until a failure of the main system occurs.
- The hardware support functions should be easily upgradable without requiring a visit by support personnel and without replacing hardware.

## Hardware Implementation

The above goals are achieved by providing a single-board service processor for the Model T500 system. The service processor is a microprocessor-controlled board that is located in the main cardcage. This board has control and observation connections into all of the hardware in the main cardcage. This board also contains the power system control and monitor which controls the power system. The service processor has a command-oriented user interface which is accessible through the same console mechanism as the operating system console connections on previous systems (through the system's access port). The logical location of the service processor is shown in Fig. 20.

The service processor and power system control and monitor are powered by special bias power which is available whenever ac power is applied to the system cabinet. The service

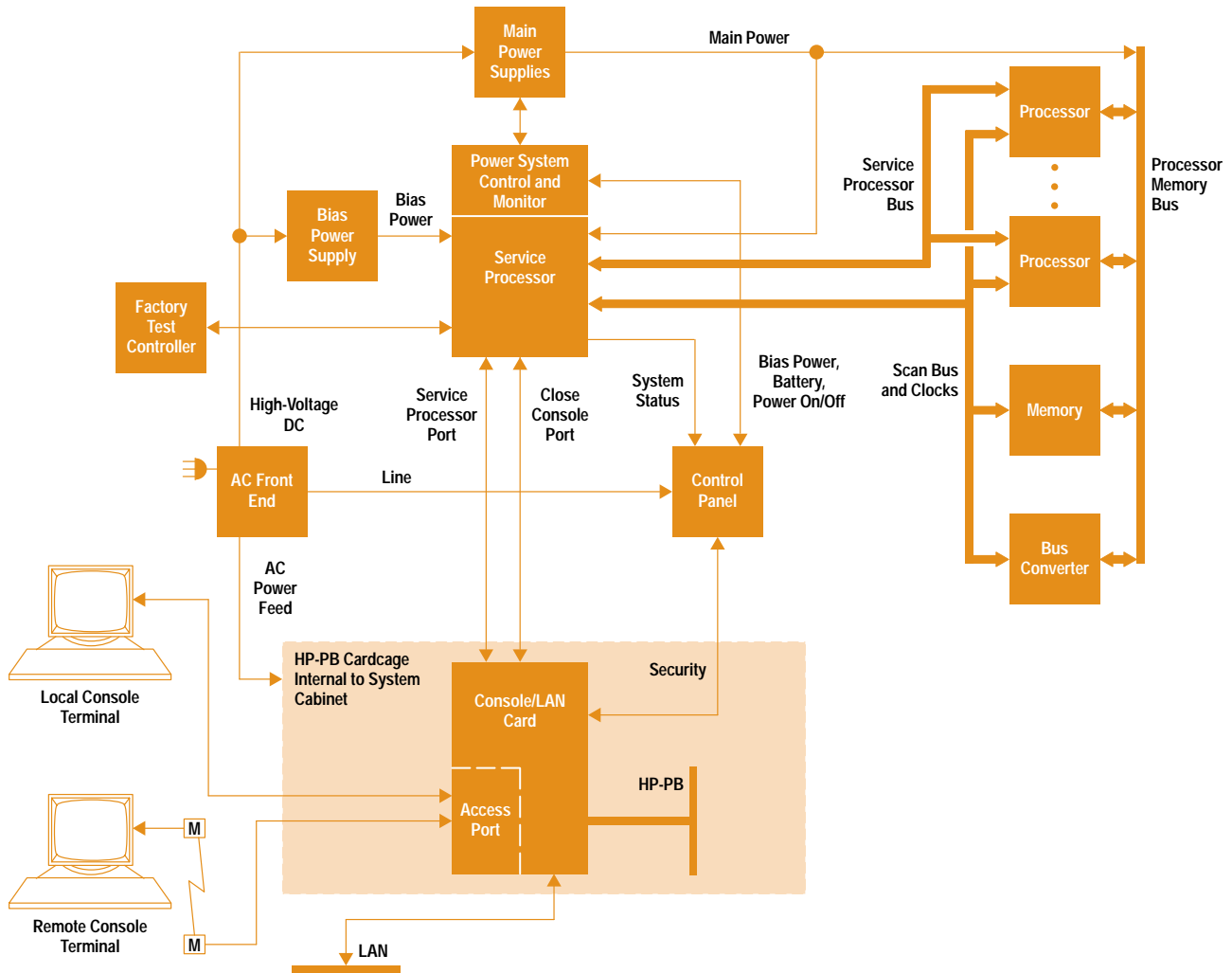


Fig. 20. Service processor block diagram.

processor is thus independent of the main system power supplies, and can be accessed under almost all system fault conditions.

The service processor has a communications channel to the power system control and monitor which allows it to provide the operating code for the power system control and monitor microprocessor, and then to issue commands to the power system control and monitor and monitor its progress. The power system control and monitor controls the power system under service processor supervision and notifies the service processor of power and environmental problems. The service processor provides a user interface to the power system which is used by support personnel when troubleshooting power and environmental problems.

The service processor is connected to each card on the processor memory bus by both the system clocks and the scan bus. Through the system clocks, the service processor provides clocking for the entire Model T500 system. The scan bus allows the service processor to set and read the state of the cards without using the main system bus. This mechanism is used to determine and alter system configuration and for factory testing.

The service processor is connected to the processors in the system by the service processor bus. The service processor bus allows the processors to access instructions and data stored on the service processor. The instructions include processor self-test code and processor dependent code, which performs architected system functions. The data stored on the service processor includes configuration information and logs of system activity and problems. The service processor bus also allows the processors to access common system hardware that is part of the service processor, such as system stable storage which is required by the PA-RISC architecture, and provides access to the console terminals through the close console port. Because service processor bus access is independent of the condition of the processor memory bus, the processors can access error handling code, make error logs, and communicate with the console terminals even if the processor memory bus has totally failed.

The service processor drives the system status displays on the control panel. These include the large status lights, the number of processors display, and the activity display. The service processor also mirrors this information onto the console terminals on the status line.



The connections between the service processor and the console/LAN card provide several functions. The service processor's user interface is made available on the local and remote console terminals by the access port firmware which is part of the console/LAN card. The user interface data is carried through the service processor port connection. Because the internal HP-PB cardage which houses the console/LAN card is powered by the same source of ac power as the service processor, the access port and its path to the console terminals are functional whenever the service processor is powered. The system processors access the console terminals through the close console port connection to the access port firmware during the early stages of the boot process and during machine check processing when the I/O subsystem is not necessarily functional. The service processor also sends control information and communicates its status to the console/LAN card through the service processor port. Console terminal access to the system and service processor functions is controlled by the access port firmware on the console/LAN card.

A connection exists between the service processor and a test controller used for system testing in the factory. This connection allows the system internal state to be controlled and observed during testing.

Because the service processor and power system control and monitor do not operate from the same power supplies as the processor memory bus, the service processor's control features and error logs are available even when the remainder of the system is inoperable. Because logging, error handling, and console communications paths exist that are independent of the system buses, these functions can operate even when system buses are unusable. The service processor is architected so that its failure does not cause the operating system or power system to fail, so that failure of the service processor does not cause the system to stop. The access port is independent of the service processor and detects service processor failure. It notifies the user of service processor failure on the console terminals, providing time for the service processor to be repaired before it is needed for system-critical functions.

## Features

The hardware implementation described above is extremely flexible because of its large connectivity into all of the main system areas. As a result, the service processor's features can be tailored and changed to ensure that the customer's service needs are adequately met. The service processor in its current implementation includes the service features described in the following paragraphs.

**Configuration Control.** The service processor keeps a record of the processor memory bus configuration including slot number, board type, revision, and serial number. The service processor reconciles and updates this information each time the system is booted by scanning the processor memory bus and identifying the modules it finds. Various error conditions cause defective processor memory bus modules to be automatically removed from the configuration. The user is alerted to such changes and boot can be optionally paused on configuration changes. The service processor's user interface contains commands to display and alter the configuration,

including removing modules from the configuration or adding them back into the configuration. Modules that are removed no longer electrically affect the system, making configuration an effective means of remotely troubleshooting problems on the processor memory bus.

**Logs.** The service processor has a large log area that contains logs of all service-processor-visible events of support significance. Each log contains the times of event occurrences. Logs that warn of critical problems cause control panel and console terminal indications until they have been read by the system operator. The service processor user interface contains commands to read and manage the service processor logs. Information in the service processor logs can be accessed by diagnostic software running under the operating system. The service processor logs include:

- Power system anomalies
- Environmental anomalies
- Ac power failure information
- Automatic processor memory bus module deconfigurations that occur because of failures
- Operating system major state changes (such as boot, testing, initialization, running, warning, shutdown)
- High-priority machine check information
- Problems that occur during system startup before the processors begin execution
- Processor self-test failure information.

**Operating System Watchdog.** The service processor can be configured to observe operating system activity and to make log entries and control panel and console indications in the event of apparent operating system failure.

**Electronic Firmware Updates.** The service processor and processor dependent code work together to update system firmware without the need for hardware replacement. The service processor contains the system processor dependent code (boot and error firmware), the firmware for the power system control and monitor to control the power system, and its own firmware. Two copies of each exist in electrically erasable storage so that one copy can be updated while the other copy is unchanged. The service processor can switch between the two copies in case a problem occurs in one copy.

**Remote Access.** The user gains access to the service processor user interface through the access port. The access port is the single point of connection for the system console terminals, both local and remote. As a result, all troubleshooting information that is available on local console terminals is available remotely.

**Factory Test Support.** The service processor serves as a scan controller, providing full access to the internal state of the custom VLSI chips contained on processor memory bus cards. This access is provided through the programmable clock system and the scan bus. Using the scan controller features of the service processor, a factory test controller can test the logic in the processor memory bus portion of the system under automatic control.

**System Status Control.** Because the service processor controls the system status indicators, it is able to display an accurate summary of the complete hardware and software state of the

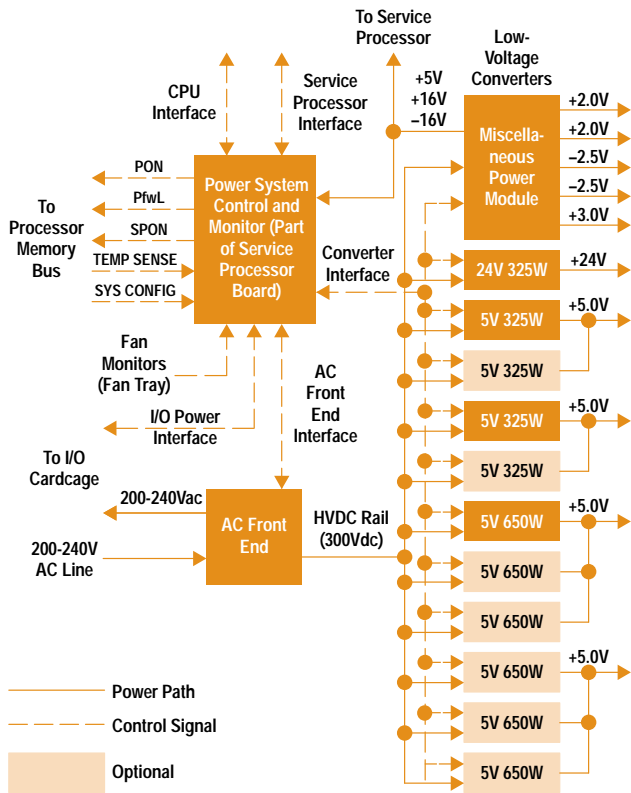


Fig. 21. Power system block diagram.

system. The service processor can do this even when the system processors or main power system are unable to operate.

## Power System

The power system provides regulated low-voltage dc to all logic assemblies in the processor memory bus cardcage, and to the array of fans located just below the cardcage assembly. The power system is designed to grow and reliably support the need for ever-increasing processor, memory, and I/O performance. It has the capacity to deliver almost 4,000 watts of dc load power continuously. The block diagram of the power system is shown in Fig. 21. The modular design consists of an ac front-end assembly, several low-voltage dc-to-dc converters, and a power system control and monitor built within the service processor.

The ac front end, shown in Fig. 22, contains one to three power-factor-correcting upconverter modules, each providing regulated 300Vdc, and has an output capacity of 2.2 kilowatts. The upconverter modules run on single-phase or dual-phase, 208Vac input. They have output ORing diodes, implement output current sharing, and are capable of providing true N+1 redundancy for higher system capacity and availability. (N+1 redundancy means that a system is configured with one more module than is necessary for normal operation. If there is a subsequent single module failure the extra module will take over the failed module's operation.)

The active power-factor-correcting design allows the product to draw near unity power factor, eliminating the harmonic currents typically generated by the switching power supplies in a computer system. The design also has a very wide ac input operating range, is relatively insensitive to line voltage transients and variations, and allows a common design to be used worldwide. It also provides a well-regulated 300Vdc output to the low-voltage dc-to-dc converters.

The low-voltage dc-to-dc converters are fed from a single 300V rail and deliver regulated dc voltage throughout the main processor cardcage. The single-output converters, of which there are two types, have capacities of 325 and 650 watts and a power density of about 3 watts per cubic inch. They have current sharing capability for increased output capacity, and are designed to recover quickly in the event of a module failure in a redundant configuration. The converters have output on/off control and a low-power mode to minimize power drain on the 300V rail when shut down. Their output voltage can be adjusted by the power system control and monitor.

The power system control and monitor provides control for power sequencing, fan speed control, and temperature measurement. It ensures that the modular converters and the system load are consistent with each other. The controller also monitors status and system voltages. This information is communicated to the service processor and saved in a log to aid in the support and maintenance of the system.

Together, the power system control and monitor, power-factor-correcting upconverters, and low-voltage dc-to-dc converters form a scalable, high-capacity, highly available, modular power system. The system is easily updated and can be upgraded to support higher-performance processor,

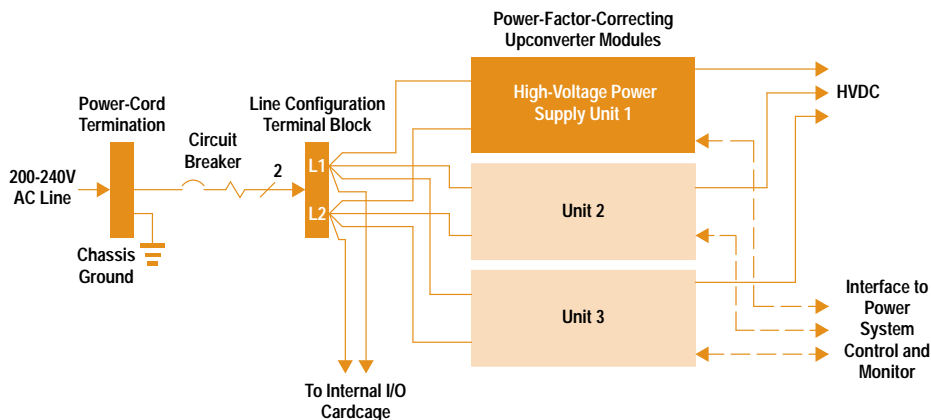


Fig. 22. Ac front end block diagram.

memory, and I/O technologies as they are developed for the Model T500 platform.

## Product Design

The Model T500 package is a single-bay cabinet. Overall, it is 750 mm wide by 905 mm deep by 1620 mm tall. A fully loaded cabinet can weigh as much as 360 kg (800 lb). A skeletal frame provides the cabinet's structure, supporting the card cages, fan tray, and ac front end rack. External enclosures with vents and a control panel attach to the frame.

The processor memory bus boards and the low-voltage dc-to-dc converters reside in cardcages in the upper half of the Model T500 cabinet. They plug into both sides of a vertically oriented, centered backplane to meet bus length restrictions. A bus bar assembly attaches to the upper half of the backplane to distribute power from the larger 650-watt converters to the extended-power slots that the processor boards use.

There are 16 processor memory bus slots in the Model T500: six in the front cardcage and ten in the rear cardcage. Eight of the 16 slots are extended-power slots, which have a board-to-board pitch of 2.4 inches, twice the 1.2-inch pitch of the other eight standard slots. These wider slots allow increased cooling capability for the processor board heat sinks. The standard slots are used for bus converters. Memory boards can go in either standard slots or extended-power slots.

Looking at the front view of the cabinet in Fig. 23, six extended-power processor memory bus slots are to the left of the low-voltage dc-to-dc converter cardcages in which

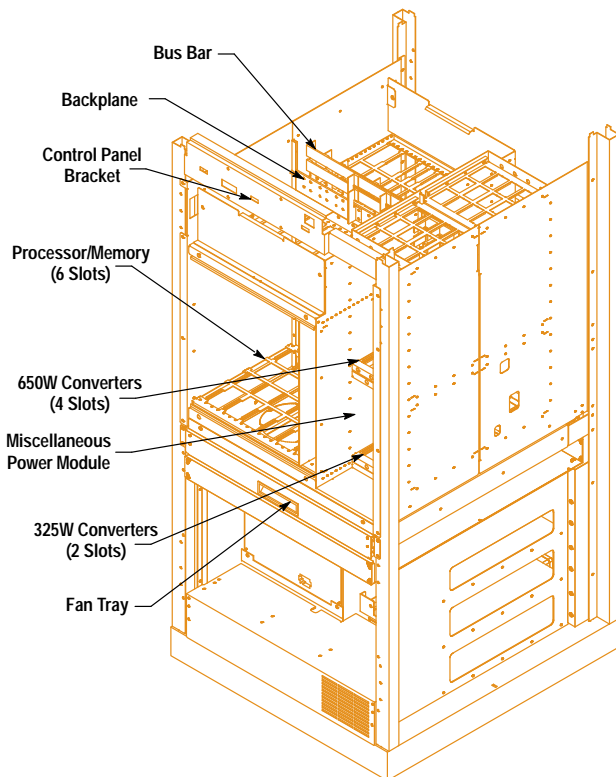


Fig. 23. Front view of Model T500 cabinet.

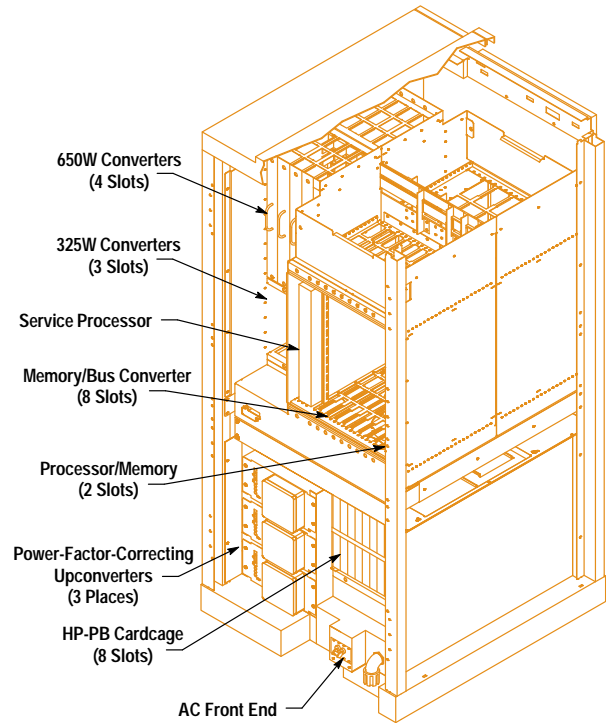


Fig. 24. Rear view of Model T500 cabinet.

four 650-watt converters reside above two 325-watt converters and the miscellaneous power module. When viewing the rear of the cabinet in Fig. 24, ten processor memory bus slots, two of which are extended-power, reside to the right of the converter cardcages in which four 650-watt converters are above three 325-watt converters. The service processor is located in a dedicated slot between the rear processor memory bus and the converter cardcages.

The fan tray is located beneath the cardcages. Air enters through the top vents of the cabinet and is pulled through air filters and then through the processor memory bus and dc-to-dc converter cardcages to the fan tray. Half of the air is exhausted through the lower cabinet vents while the other half is directed to cool the HP-PB cardcage boards located in the ac front end rack. The fan tray is mounted on chassis slides to allow quick access to the fans.

The ac front end rack is mounted on the base of the Model T500 cabinet. This rack holds up to three power-factor-correcting power supply modules, an internal HP-PB cardcage, and the ac input unit. The HP-PB power supply has its own integral cooling fan. The ac front end power-factor-correcting modules have their own fans and air filters and take in cool air from the rear lower portion of the cabinet and exhaust air out at the front lower portion of the cabinet.

The rear of the internal HP-PB cardcage has an HP-PB bus converter and seven double-high or 14 single-high HP-PB slots as well as the battery for battery backup. The front of the HP-PB cardcage has a power supply and the power system control and monitor module. HP-PB backplane insertion is from the top of the cardcage by way of a sheet-metal carrier.

Additional rackmount HP-PB expansion modules and system peripherals are housed in peripheral racks. Both HP-PB card cages (internal and rackmount) leverage the same power supply and backplane assemblies, but have different overall package designs. The rackmount version has a cooling fan that directs air in a front-to-back direction. The HP-PB boards mount in a horizontal orientation and the cables exit towards the rear of the peripheral rack. The rackmount unit is 305 mm high (7 EIA standard increments) by 425 mm wide by 470 mm deep. The peripheral racks are 600 mm wide by 905 mm deep by 1620 mm tall and have mountings to hold products conforming to the EIA 19-inch standard.

The Model T500 industrial design team drove the system packaging design to come up with a unified appearance for HP's high-end and midrange multiuser systems. The result is an industrial design standard for a peripheral rack system that fits well with the Model T500 design. This cooperative effort ensured consistency in appearance and functionality.

### Electromagnetic Compatibility

EMC shielding takes place at the printed circuit board level, the power supply level, and the cardcage level and does not rely on external enclosures for containment. This keeps noise contained close to the source. A hexagonally perforated metal screen is used above and below the processor memory bus cardcage to minimize resistance to airflow while providing the required EMI protection. Nickel plating is used on the steel cardcage pieces to ensure low electrical resistance between mating metal parts. The backplane has plated pads in the areas that contact the cardcage pieces. Conductive gaskets are used to ensure good contact between the backplane, the cardcages, and the cover plates. ESD (electrostatic discharge) grounding wrist straps are provided in both the front and rear of the cabinet.

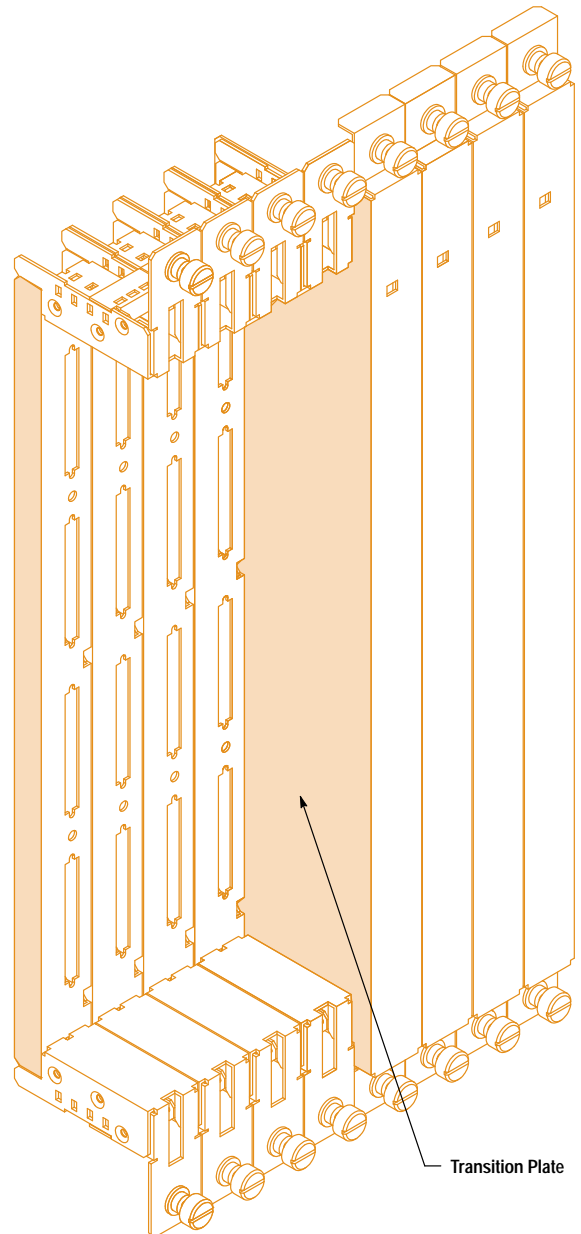
Surface mount filtering is used on the backplane to control noise on signal lines exiting the high-frequency processor memory bus cardcages and to prevent noise from coupling into the low-voltage dc-to-dc converters.

All processor memory bus boards have a routed detail in four corner locations along the board perimeter to allow for grounding contact. A small custom spring fits into via holes and resides in the routed-out space. This spring protrudes past the edge of the board and contacts the card guides in the cardcage. Surface mount resistance elements lie between the vias and the board ground planes. This method of grounding the processor memory bus boards helps reduce EMI emissions.

### System Printed Circuit Boards

The processor memory bus cardcage is designed to accept 16.9-inch-high-by-14-inch-deep boards. This large size was required for the 256M-byte memory board. Since the processor and processor memory bus converter did not require large boards, it was important to have a cardcage and cover plate design that allows boards of various depths to be plugged into the same cardcage, thereby optimizing board panel use.

The processor plugs into this deep cardcage by means of a sheet-metal extender. The bus converter was more difficult to accommodate since this shallow board requires cables to



**Fig. 25.** Bus converter sheet-metal design, showing transition plate.

attach to its frontplane. Therefore, a transition plate was developed to transition from the shallower board bulkheads to the full-depth cardcage cover plates as shown in Fig. 25. This transition plate locks into the adjacent bulkhead to maintain the EMI enclosure. However, either the transition plate or the adjacent bulkhead to which it latches can be removed without disturbing the other.

The Model T500 backplane is 25.3 inches wide by 21.7 inches high by 0.140 inch thick and has 14 layers. This backplane has many passive components on both sides, including press-fit and solder-tail connectors, surface mount resistors and capacitors, processor bus bars, and filters. The backplane connectors are designed to allow at least 200 insertions and withdrawals.

A controlled-impedance connector is used on the processor memory bus boards to mate to the backplane. The codevelopment that took place with the connector supplier was a major undertaking to ensure that the connector would work in our surface mount processes repeatably and meet our reliability and serviceability requirements.

### Cooling

The Model T500 cooling system is designed to deliver high system availability. This is achieved by incorporating redundant fans, fan-speed tachometers, air temperature sensors on the hottest parts of the boards, and multiple-speed fans. The Model T500 meets the HP environmental Class C2 specification for altitudes up to 10,000 feet with an extension in temperature range up to 40°C.

Computational fluid dynamics software and thermal analysis spreadsheets were used to evaluate various components, heat sinks, and board placements. These tools helped the team make quick design decisions during the prototype stages. All high-powered components that were calculated to operate close to their maximum allowable junction temperature in the worst-case environment were packaged with thermal test dies to record chip junction temperatures accurately. Small wind tunnels were used to determine package and heat sink thermal performance for various airflows. Larger wind tunnels were used to evaluate board airflow to give the board designers feedback on component placement by monitoring preheat conditions and flow obstructions. On printed circuit boards, external plane thermal dissipation pads were used where possible in lieu of adding heat sinks to some surface mount parts.

A full-scale system mockup was built. Various board models, air filters, EMI screens, and vents were tried to gather system airflow resistance data to determine the size and number of fans required. Various cooling schemes were evaluated by altering airflow direction and fan location. Pulling air down through the cabinet was found to provide uniform airflow across the cardcages while keeping the air filters clean by their high location. Having the fans low in the product and away from the vents kept noise sources farther away from operators and made servicing the fans easier.

The eleven dc fans in the fan tray have the ability to run at three different speeds: high, normal, or low. Seven fans run at low speed during startup and battery backup to keep the power use at a minimum while supplying sufficient cooling. All eleven fans run at normal speed while the system is up and running with the inlet air at or below 30°C. In this case the system meets the acoustic noise limit of 7.5 bels (A-weighted) sound power. The fans run at high speed while the system is up and running with the inlet air above 30°C or when the temperature rise through the processor memory bus cardcage exceeds 15°C.

At high speed, the fan tray has a volumetric airflow of approximately 1200 ft<sup>3</sup>/min, which is designed to handle over six kilowatts of heat dissipation. This amount of power was considered early in the project when alternate chip technologies were being investigated. Therefore, the Model T500 has a cooling capacity of approximately one watt per square centimeter of floor space, a threefold increase over the high-end platform that the Model T500 is replacing, yet it is still

air-cooled. The minimum air velocity is two meters per second in all of the processor memory bus slots and the typical air velocity is 3 m/s.

Because the processor memory bus cardcage contains high pressure drops and airflows, the board loading sequence is important, especially for the processor boards. Since the heat sinks are on the right side of the vertical processor boards, they are loaded sequentially from right to left. This ensures that air is channeled through the processor heat sinks instead of bypassing them in large unfilled portions of the cardcage.

### Manufacturing

The fundamental strategy for manufacturing the HP 9000 Model T500 corporate business server was concurrent engineering, that is, development of both the computer and the technologies and processes to manufacture it at the same time. This resulted in a set of extensions to existing high-volume, cost-optimized production lines that allow sophisticated, performance enhancing features to be added to the corporate business server.

#### Cyanate Ester Board Material

Printed circuit boards based on cyanate ester chemistry (referred to as HT-2) have much better thermal, mechanical, and electrical performance than typical FR-4 substrates. These properties make HT-2 ideally suited for large printed circuit assemblies with intensive use of components with finely spaced leads, high-reliability applications, high-frequency applications, and applications with tight electrical tolerances.

More advanced printed circuit board designs tend to increase the aspect ratio of the board, or the ratio of the thickness of the board to the width of the vias for layer-to-layer connections. This is hazardous for FR-4 substrates because higher-aspect-ratio vias tend to be damaged in the thermal cycles of printed circuit assembly processes because of the expansion of the thickness of the boards in these cycles. The reliability of vias and through-hole connections (where the processor memory bus connector or VLSI pin-grid arrays are soldered to the board) is essential to the overall reliability, manufacturability, and repairability of the Model T500 memory board.

Because of their high glass transition temperature, HT-2 substrates are ideally suited to survive the stressful assembly and repair processes and to increase the yields within these processes. The glass transition temperature is the temperature at which the laminated fiberglass printed circuit board transitions from a solid to a pliable material. This is exceeded for FR-4 in the printed circuit assembly process, resulting in distortions of the boards. If no fine-pitch or extra-fine-pitch parts are used, the distortion for FR-4 is acceptable in the surface mount process. For large boards that use fine-pitch components, the surface mount processes tolerate less distortion. HT-2 has the advantage that it remains stable because it doesn't reach its glass transition temperature in the manufacturing process.

## Printed Circuit Assembly and Test

Model T500 system requirements, through their impact on the memory board design, required development of significant new printed circuit assembly process capability. This process development effort began two years before volume shipments and identified the important areas for engineering effort. New technology introduced in printed circuit assembly included advanced reflow techniques. This is important because the total thermal mass of the components reflowed on the memory board is large, and because of the nature of the connector used for the processor memory bus. Special solder paste application methods were developed for the processor memory bus connector. This provides the assembly process with wide latitude for the connectors, pin-grid arrays, standard surface mount parts, and fine-pitch parts.

A key benefit of the cyanate ester choice for double-sided assemblies is reduced runout of the board, resulting in improved registration of the solder-paste stencil and components for higher yields of solder joints. In the double-sided surface mount process, the B side components are placed and reflowed before placement of the A side components. Since reflow for the B side is conducted at a temperature far above the glass transition temperature for standard FR-4 material, the boards would have been distorted in this step if FR-4 had been used. Thus FR-4 boards would have a higher failure rate on solder connections for the A side components.

Printed circuit board test was another area identified by the early concurrent engineering effort. Model T500 printed circuit assemblies are tested using a strategy extended from the HP 3070 board test system. Much of the test is conducted using leading-edge scan techniques. For example, because of trace length and capacitance, it was impossible to add test points between the processor memory bus drivers of the bus transceivers and the arbitration and address buffer and the resistor pack that connects to the bus without impacting performance. A scheme was devised using the scan port to activate each chip's drivers. The HP 3070 is set to apply a known current to each resistor and measure the voltage drop, from which the resistor's value and connectivity can be determined. There is no loss of test coverage for these fine-pitch parts and the scheme has the added benefit of verifying much of the chip's functionality. Because of the chip design lead time, HP's own scan port architecture (designed several years in advance of the IEEE 1149 standard for this type of test approach) is used and custom software tools were developed. Current chip designs contain the IEEE 1149 scan port which is directly supported by the HP 3070.

A major manufacturing challenge was the total number of nets and the board layout density found in the memory board. With 4273 nets, if normal HP 3070 design rules, which require one test point per net, were followed as much as 20% of the surface of the board would have been dedicated to test points. To solve this problem a scan-based approach is used on the nets where VLSI parts have scan ports. By using the scan ports and exercising some of the MSI part functionality, the number of nets that need test points is reduced to 2026.

This approach freed board space and allowed the needed density to be achieved. If this density had not been achieved, the alternative would have been to lower the capacity of

---

---

## Package Design Using 3D Solid Modeling

The industrial design and product design groups designed the HP 9000 Model T500 corporate business server package using the HP ME 30 solid modeling system. In the past, designs were drawn as 2D orthographic layouts. These layouts were then dimensioned and paper copies were given to the vendor for fabrication. Now, 3D bodies are sent directly to vendors via modem without having to dimension them. A 2D drawing is also sent to the vendor to provide a view of the part, usually isometric, and to call out notes and necessary secondary operations (plating, tolerances, cosmetic requirements, press-in fastener installations, etc.).

Using 3D solid modeling allowed the product design group to reduce design time, reduce part documentation time, and reduce design errors caused by using 2D layouts (with orthographic views, all three 2D views must be updated for a design change instead of a single 3D body). Additional benefits are faster turnaround on prototypes and an improved process for creating assembly documentation (isometric views of assembly positions are easily created by manipulating 3D bodies).

Eight engineers created approximately 150 sheet-metal parts, ten plastic parts, 25 cables, 15 miscellaneous parts, and many board components. Managing such a large mechanical assembly was initially thought to be too difficult. But having an organized file structure and 3D body placement strategy allowed the design team to work together efficiently. All engineers worked on their own assemblies, stored in separate write-protected directories, and were able to view adjoining assemblies for interface design.

---

---

each memory board, thereby lowering the overall system memory capacity.

The service processor presented two major challenges to make it fit both electrically and mechanically onto the HP 3070 test fixture. The total of 2312 nets on this board made it important to make all possible electrical pins of the test fixture available, which was difficult considering the large number of components. This problem was alleviated by careful layout of the service processor with the test fixture in mind. A custom fixture was designed to accommodate the board with its 2.5-inch bulkhead.

All of the boards and fixtures are designed to accommodate the transition to a no-clean process, which allows manufacturing of printed circuit assemblies without a chlorofluorocarbon (CFC) wash. This advanced work was driven by Hewlett-Packard's commitment to the total elimination of CFCs, which have been shown to destroy the ozone layer. The elimination of CFC use at HP was accomplished by May 15, 1993, more than two years ahead of the Montreal Protocol goal for an international ban on the use of these chemicals.

## Mechanical and Final Assembly and Test

A key focus of concurrent design for manufacturability was the frame and cardcage design. Early effort by the design team and manufacturing went into detecting areas to improve the design for ease of assembly, to minimize the number and variety of fasteners, and to reduce the number of stocked items. This resulted in a set of features that include:

- Extensive use of captive fasteners, that is, fasteners that are preplaced in mechanical subassemblies. This reduces the number of individual mechanical parts to handle during assembly.
- A minimal set of unique fasteners with extensive use of Torx fasteners.

- One-direction fastening. Assemblers are not required to reach around or use awkward movements during assembly.
- A simplified assembly procedure. Only one piece to pick up and handle during any operation.
- Modularity. It is very easy to install or replace many components in the chassis without interference.
- Extensive use of high-density connectors for wiring harnesses. This reduces wiring time and errors. Point-to-point wiring is minimal.
- A robust cabinet and a very strong frame. The frame can survive shipping on its casters alone, and does not require a special pallet for most shipments.
- A refrigerator-sized cabinet that when fully loaded (approximately 360 kg) can still be moved easily by any operator or technician.

The Model T500 is designed with many inherent testability features, most of which are accessible using the system console. The system console is one of the most fundamental functions of the Model T500. It can be used in the earliest steps in bringing up and testing a newly assembled system. This permits extensive control and monitoring capability from a single communication point for manufacturing's automated test control host, and eliminates the need for many additional custom devices traditionally used for testing large computer systems. Many of the testability features benefit both manufacturing and customer support. The capabilities used for manufacturing test include the following:

- Monitor and change system parameters (such as secondary voltages or power system status) from the system console.
- Review from the console the system activity logs which track events that may indicate incorrect operation.
- Change self-test configuration. Select only the tests desired, or repeat tests to aid defect analysis.
- Access diagnostics through a LAN connection standard on all configurations of the system.
- Diagnose potential failure sources down to a specific integrated circuit.
- Use scan tools designed closely to manufacturing test specifications.

### Acknowledgments

The authors would like to thank the following for their contributions to this article: Manfred Buchwald, Ken Chan, Ira Feldman, Dick Fowles, Chuen Hu, Ed Jacobs, Dong Nguyen,

Phil Nielson, Nazeem Noordeen, Tom Wylegala, Syrus Ziai, and Steve Thomas for his photographic support. The authors also wish to thank the many individuals involved in making the HP corporate business server possible. This includes individuals from HP's mainline systems lab in Cupertino, California, the engineering systems lab in Fort Collins, Colorado, the IC design lab in Fort Collins, Colorado, the Böblingen Printed Circuit Manufacturing Operation in Böblingen, Germany, the Networked Computer Manufacturing Operation in Roseville, California, and the Cupertino open systems lab in Cupertino, California.

### References

1. R.B. Lee, "Precision Architecture," *IEEE Computer*, Vol. 22, January 1989.
2. D. Tanksalvala, et al, "A 90-MHz CMOS RISC CPU Designed for Sustained Performance," *ISSCC Digest of Technical Papers*, February 1990, pp. 52-53.
3. J. Lotz, B. Miller, E. DeLano, J. Lamb, M. Forsyth, and T. Hotchkiss, "A CMOS RISC CPU Designed for Sustained High Performance on Large Applications," *IEEE Journal of Solid-State Circuits*, October 1990, pp. 1190-1198.
4. P. Stenstrom, "A Survey of Cache Coherence Schemes for Multiprocessors," *IEEE Computer*, Vol. 23, no. 6, June 1990, pp. 12-25.
5. *TPC Benchmark C Full Disclosure Report: HP 9000 Corporate Business Server T500, Using HP-UX 10.0 and INFORMIX OnLine 5.02*, Hewlett-Packard Company, November 1993.
6. *TPC Benchmark A Full Disclosure Report: HP 9000 Corporate Business Server 890, Using HP-UX and ORACLE7 Client/Server Configuration*, Hewlett-Packard Company, publication no. 5091-6844E, February 1993.
7. Standard Performance Evaluation Corporation, Results Summaries, *SPEC Newsletter*, Volume 5, Issue 4, December 1993.
8. E. Delano, W. Walker, J. Yetter, and M. Forsyth, "A High-Speed Superscalar PA-RISC Processor," *COMPCON Spring '92 Digest of Technical Papers*, February 1992.
9. R.C. Brockmann, W.S. Jaffe, and W.R. Bryg, *Flexible N-Way Memory Interleaving*, U.S. Patent Application, February 1991.
10. K. Chan, et al, "Multiprocessor Features of the HP Corporate Business Servers," *COMPCON Spring '93 Digest of Technical Papers*, February 1993.

HP-UX is based on and is compatible with UNIX System Laboratories' UNIX\* operating system. It also complies with X/Open's\* XPG3, POSIX 1003.1 and SVID2 interface specifications.

UNIX is a registered trademark of UNIX System Laboratories Inc. in the U.S.A. and other countries.

X/Open is a trademark of X/Open Company Limited in the UK and other countries.