MICROPROCESSOR ⬡ REPORT

THE INSIDERS' GUIDE TO MICROPROCESSOR HARDWARE

Vitesse Unveils MP Chip Set for P54C

Connects to New V-Bus for High-Bandwidth Multiprocessor Systems

by Linley Gwennap

Aiming to set a standard for high-speed multiprocessor systems, Vitesse has revealed a two-chip set that connects 100-MHz Pentium processors to its new 500-Mbyte/s V-Bus. The high-bandwidth bus can connect four or more Pentium processors in a PC server. Once operating systems that support Intel's MultiProcessor Specification *(see 080603.PDF)* become available, such a server can be built entirely from off-the-shelf components, allowing small vendors to challenge the current leaders in the high-margin server market.

Vitesse is the leading maker of gallium-arsenide (GaAs) gate arrays and ASICs, and it has long sought to move this high-end IC technology from exotic applications into the mainstream PC market. The company's initial foray was a GaAs cache controller for Pentium *(see 070703.PDF)*, a $100 uniprocessor chip set that is being used by AST but few other vendors. The new 947/8 chip set is similar to the original Pentium controller but enables multiprocessing and includes the new V-Bus interface.

The chip set consists of the VSP947, a GaAs gate array that controls the second-level cache and the V-Bus interface, and the VSP948, a CMOS ASIC that routes data between the CPU, cache, and V-Bus. No glue logic is needed to connect the chip set to either a P54C processor or to the V-Bus, and SRAMs for cache data and tags complete a processor module design. Unfortunately, no standard memory or I/O controllers are yet available for V-Bus.

V-Bus is a 64-bit bus that operates at up to 66 MHz using GTL *(see 070301.PDF)* signal levels. By overlapping transactions (using a split-transaction protocol) and providing separate arbitration and address lines, the bus can sustain a peak data rate of 500 Mbytes/s, more than twice the throughput of Sun's multiprocessor XDBus *(see 070301.PDF)*. Both in-order and out-of-order transaction models are supported. Vitesse is offering the V-Bus specification openly, with neither license fees nor patent restrictions, to all interested parties.

Chip Set Builds on Earlier Design

As with Vitesse's original chip set, the new design consists of a cache control chip and a datapath chip, as shown in Figure 1 (see below). This partitioning allows the GaAs chip to handle complex control algorithms at high speeds while the CMOS device provides an inexpensive set of buffers for incoming and outgoing data.

The most important differences between the old and new chip sets are that the 947/8 supports multiprocessing and attaches directly to V-Bus, whereas the older design provides a 486-like bus interface. V-Bus has enough bandwidth to satisfy the fast P54C processors, even in multiprocessor configurations with up to five CPUs.

The new design also adds a second set of cache tags to allow bus snooping without stalling the CPU. The chips use the same 3.3-V logic levels as the P54C.

Vitesse has increased the maximum cache size to 8M, plenty of room for servers. The cache can be either direct-mapped or two-way set-associative with a line size of 32 or 64 bytes. The chip set supports single-cycle accesses (2-1-1-1, or 1-1-1-1 for pipelined transactions), but in a two-way cache, accessing the set that has been less recently used takes an extra cycle.

The current version (revision A) of the chip set supports V-Bus at exactly one-half of the CPU frequency, up to 50 MHz with a TTL interface. It also supports only the in-order transaction model. Vitesse plans to release revision B by the end of this year, adding support for the out-of-order model and a GTL-level V-Bus at 66 MHz, with no fixed relationship between the CPU and bus clocks.

V-Bus Delivers High Performance

V-Bus is similar to the high-bandwidth proprietary buses used by many vendors of multiprocessor systems. The combination of several high-performance processors, each with a voracious appetite for data, demands a large bandwidth to avoid starving the system and restricting its performance. An MP system bus must also

support the coherency protocols that ensure that data is kept consistent among multiple caches.

To address the bandwidth issue, V-Bus begins with a 64-bit data path. Clock frequency is always a tricky issue in MP systems; unlike a processor local bus, which is contained within a few inches on a single board, an MP bus typically runs across a backplane with several stubs. V-Bus achieves 66-MHz operation across a 10" backplane with eight slots. It uses GTL signal levels, which swing only 800 mV, to improve signal transition times on a heavily loaded bus. The bus can also be used with TTL signal levels, but this reduces the maximum frequency slightly, to 50 MHz.

Combining the data width with the frequency gives V-Bus a peak data rate of 528 Mbytes/s. Many buses, however, cannot sustain a data rate close to their peak bandwidth. Vitesse has designed V-Bus to make the best possible use of the available bandwidth. The key factor is a split-transaction design similar to the packet-switched model used by XDBus. Instead of waiting for each transaction to complete before starting the next one, V-Bus overlaps up to eight transactions at once, eliminating the dead cycles caused by memory latency.

The Vitesse design supports two split-transaction models. For smaller systems (2–4 processors), an in-order model reduces latency by simplifying the arbitration process. For larger systems, a more complicated token (out-of-order) protocol is available. Although this mode can extend the latency of a given transaction, it can more efficiently use the bus by preventing a long transaction from stalling the entire bus.

To further increase bus utilization, V-Bus uses separate signal lines for arbitration, address (40 bits), and data. Although combining these functions on a single bus can reduce pin count, it also requires that bus cycles be wasted on arbitration and address transmission. XDBus, for example, allocates no more than 73% of all cycles for data transfers. Vitesse minimizes the cost impact of the larger pin count by using BGA packages *(see **071203.PDF***)*; presumably, other V-Bus interface chips will do the same.

Table 1 lists the signals used by V-Bus in the 947/8 implementation. The V-Bus specification defines a total of 157 signals, plus 23 additional system-specific signals. V-Bus uses a dual-sided edge connector from AMP with 90 signal pins on each side (180 total). Four bits are provided for the different IDs, allowing a maximum of 16 devices. This configuration is allowed when the bus operates at 50 MHz, but electrical issues limit the bus to eight devices at 66 MHz.

The specification defines an extended V-Bus with a 128-bit data bus using a 260-pin connector. Although this change doubles the bandwidth of the bus, it also significantly increases the implementation cost. The 947/8 chip set supports only the 64-bit V-Bus.

Split Transactions Improve Utilization

In a conventional bus, transactions cannot be overlapped. If a processor, for example, requests data from memory, no other device may use the bus during the memory latency period. While this protocol works fine in a uniprocessor system, the model breaks down when several devices may need the bus at once.

The V-Bus in-order transaction model is fairly simple, yet it provides much greater bus utilization than the conventional protocol. Once a device finishes requesting data, the bus is available for the next device to make a request. Because read requests use the separate address bus, a new request can be issued even as a previous request is being fulfilled.

Figure 2 shows some sample V-Bus transactions. Data is requested from memory by issuing an address on the address bus (A). During the memory latency, an I/O write is issued (B). As the first read's data is received, a second memory read is requested (C). In this way, the bus can be used even when waiting for memory (or I/O) to respond. Using the in-order model, however, the CPU must wait to issue the data for the I/O write until after the data from the first memory read is returned.

The token protocol breaks this ordering rule, as Figure 2(b) shows. In this case, the ID signals provide a token number to match data with addresses. This model allows the I/O write to issue its data before the memory read completes, eliminating the stall in the previous example. The drawback to this method is that an arbitration cycle is required for data transmission as well as address transmission, slightly increasing the latency. The out-of-order protocol also requires more complexity in the V-Bus devices.

V-Bus supports a number of transaction types. Memory reads and writes can consist of a full cache line (either four or eight cycles) or a single 64-bit transfer. In the latter case, eight byte-enable signals are available
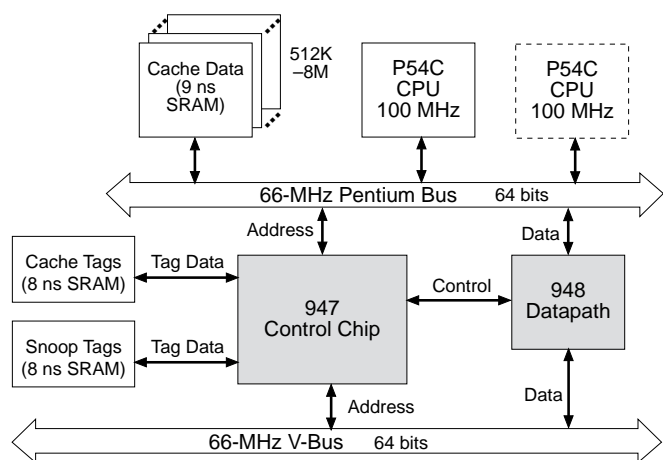

Figure 1. Vitesse's 947/8 V-Bus chip set consists of a GaAs control chip and a CMOS datapath chip.

for byte and word transfers, and the data is not cached. Transactions to I/O space are always a single 64-bit transfer and, of course, are not cached. Finally, special transactions are provided for shutdown, halt, flush acknowledge, and branch tracing.

As a multiprocessor bus, V-Bus implements a snooping protocol to ensure cache coherency. All devices snoop all cachable transactions and check the addresses against their own caches. The caches use a MESI protocol to maintain cache coherency. If one device attempts to write to an address that is contained in another device's cache, the latter device flushes its data back to main memory; the first device can snarf the data for its own cache as it is being written to memory, eliminating the need for a second read transaction.

Vitesse has changed the typical MESI protocol to include a fifth state called write protected (WP). This state can be used to cache ROM code, for example. If the CPU attempts to write to a line in the WP state, that line is invalidated in the cache and the write is forwarded to the system where, for ROM data, it would be ignored. V-Bus also allows the modified (M) state to be set independently for the two 32-byte sectors of a 64-byte cache line; 32-byte cache lines are not sectored and use only a single M-bit.

Although the initial chip set is designed for Pentium, Vitesse says that V-Bus itself is processor-independent. The current specification, however, requires Intel burst ordering on the bus, which makes Pentium systems simple but other implementations more difficult. The specification allows either 32-byte or 64-byte cache lines, but all external caches in the system must use the same cache-line length.

Arbitration Requires Central Arbiter

The V-Bus specification leaves arbitration as an implementation-dependent function, requiring only that the arbitration be fair: all transactions must eventually be serviced. The arbiter may assign higher priority to certain transfers, such as DMA block transfers, while ensuring that no single device hogs the bus.

To minimize latency in a lightly loaded system, bus ownership should default to the previous winner whenever there are no outstanding requests. This "parking" scheme allows a single device to continue to use the bus without arbitrating. Parking is particularly important in a uniprocessor system; although V-Bus is designed for MP systems, some systems may begin with a single processor while allowing expansion to multiple CPUs.

A locking mechanism is provided for implementing semaphores and other atomic transactions. When a bus owner asserts VLOCK#, it retains ownership of the bus for a series of uninterruptible transactions. In addition to semaphores, this mechanism could also be used to implement a fast DMA block transfer, for example.

The V-Bus specification includes the 947/8 arbitra-

Signal Name	No.	Description
VA[39:3], VAP	38	Address bus, parity
VD[63:0], VDP[7:0]	72	64-bit data bus, ECC
VBE#[7:0]	8	Byte enables
VAID[3:0], VDID[3:0], VSID[3:0]	12	Bus IDs
VRDY#, VARTY#, VND#, VNA#, VERR[2:0]#, VHIT#, VHITM#, VSNW, VDTM#	11	Bus cycle responses
VADS#, VD/C#, VM/IO#, VW/R#, VCACHE#, VLOCK#, VAHOLD, VEADS#, VINV, VSNK	10	Bus cycle types
SAREQ, SAGNT, SDREQ, SDGNT	4	Arbitration signals*
VCLK, RESET	2	Bus clock, reset
Other	23	System dependent

Table 1. V-Bus has separate data and address signals to increase bus utilization by overlapping transactions. The data and address buses have their own ID and arbitration signals. *947/8 specific.

tion scheme as an implementation example. These chips use a relatively simple scheme. Each device sends a request signal, SAREQ, to the central arbiter and receives a signal, SAGNT, granting it the bus. When the token protocol is used, two additional signals, SDREQ and SDGNT, provide data bus arbitration. The arbiter must examine the requests from all devices and assert one grant signal to the arbitration winner.

Because a typical system will have multiple processors, the 947/8 chip set does not include an arbiter. The arbiter should be in centralized system logic or perhaps the memory controller. Vitesse has developed a VHDL model of an arbiter compatible with the 947/8 and plans to offer it free of charge to V-Bus system developers.
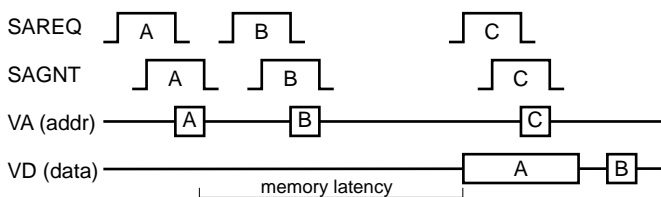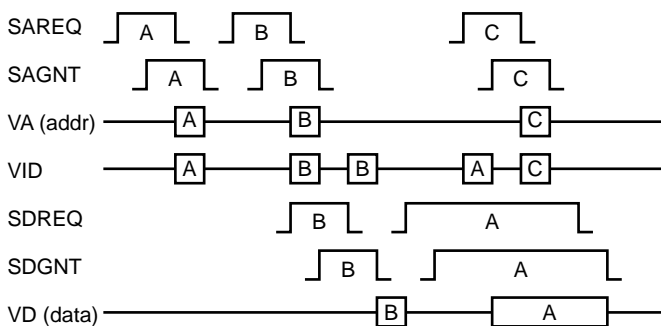

Figure 2. V-Bus supports both in-order and out-of-order transaction models. In the bottom example, the I/O write (B) completes during the latency of the memory read (A), but the read's data return is delayed by one cycle due to the need to arbitrate for the data bus.

System Logic Under Development

One thing that is missing from the V-Bus strategy is system logic. Vitesse is developing a V-Bus device, the 949, that interfaces with a third-party memory/PCI controller, and it is also working with another company on a high-speed SCI (scalable coherent interface) bridge. The memory/PCI chip is the most essential, as it provides a bridge to standard PC components needed to complete a low-cost MP system. This chip would also be the logical place for the V-Bus central arbiter. Until the 949 is ready, V-Bus system vendors must design their own ASICs for memory and I/O interfaces.

Vitesse also has no software support for its chip set. The chips are designed to be transparent to the operating system, leaving it to the system designer to handle interrupts and multiprocessor synchronization. The company expects that V-Bus will be used in systems that support the Intel MP Specification (MPS), which requires Intel's APIC for interrupt handling. Because the 947/8 works with the P54C chip, which already has a processor APIC, there is no need for an APIC in the Vitesse chip set. The central I/O APIC, however, must be implemented externally. There are no operating systems shipping today that are compliant with the recently announced MPS, but at least some MPS software should be available by the end of this year.

Vendors of proprietary "superservers" have the ability to design the additional hardware and software needed by the Vitesse chip set, and Tricord has already signed up for the 947/8 chip set. To attract volume PC server designs, however, all the necessary hardware and software must be openly available.

V-Bus is the first openly licensed bus for high-performance multiprocessor systems. It offers 25% more throughput than LSI Logic's MPI bus (used in its Hydra chip set) and 33% more than Corollary's C-Bus II. It doubles the performance of Sun's MBus and XDBus. HP's Runway bus *(see 080302.PDF)* offers a higher sustainable bandwidth than V-Bus, but HP does not plan to make its PA-7200 chip openly available.

Vitesse's uniprocessor chip set has seen slow acceptance in the desktop market, where few vendors are willing to add cost to their systems even if they gain performance. The server market is less cost-sensitive, however, and more willing to pay for performance.

Vitesse expects to offer a $250 volume price for the 947/8 set. LSI charges $190 in 1,000-unit volumes for its Hydra cache controller, which it expects to ship later this year. Corollary quotes $250 in 1,000-unit quantities for its C-Bus processor chip set, which it plans to ship in 3Q94. Both LSI and Corollary offer a compatible DRAM controller and a PCI bridge; these costs are not included in the above prices. Vitesse's fast cache will deliver better uniprocessor performance than these products, and the V-Bus bandwidth will support greater numbers of processors more efficiently.

Vitesse's openness won't offer an advantage over single-vendor buses until others deliver compatible chip sets; Vitesse hopes that such announcements will be made later this year. Also, while the higher bandwidth of V-Bus provides more headroom for systems with several processors, smaller configurations may see little performance advantage over a less expensive bus.

Vitesse's chip set is a good solution for server vendors wishing to support four or more P54C processors in a system, particularly as even faster Pentium chips debut. If Vitesse delivers the needed system logic, the 947/8 may be useful in PC servers with 2–4 processors, although the price tag is a bit high and the extra bandwidth is of less importance in these smaller systems. The chip set will first appear in custom superservers, but PC server vendors should keep an eye on V-Bus to see when the complete system solution becomes available. ♦