

TRANSPARENT PCI TO PCI BRIDGE WITH DYNAMIC MEMORY AND I/O MAP PROGRAMMING

5

The invention relates to an apparatus and method of bridging two PCI buses together, and more particularly to a software transparent PCI-to-PCI repeater.

10

The performance of a personal computer (PC) is dependant upon many factors, such as the speed of the processor, memory and input/output (I/O) subsystem. With the introduction of the peripheral component interconnect (PCI) bus in 1992, the I/O subsystem was given a high performance bus from which to operate.

15

Originally, the PCI bus was not intended to supplant the existing expansion buses, such as the industry standard architecture (ISA) bus or extended industry standard architecture (EISA). However, pressure from the computer industry and competing buses caused the PCI bus to be available for expansion bus purposes. Thus, computer systems may incorporate PCI devices onto the motherboard or provide support for add-in boards.

20

The PCI bus is referred to as a mezzanine bus, or a local bus, because it usually resides between the very high performance processor bus and the lower performance ISA or EISA bus. The logic that connects one computer bus to another, allowing an agent on one bus to access an agent on the other, is known as a bridge. In PCI vernacular, an agent is any entity or device that operates on a computer bus. An agent can be either a bus master or bus slave. Bus masters, or initiators, initiate bus transactions and bus

25

slaves, or targets, respond to a bus transaction initiated by a bus master. Oftentimes, an initiator is on one bus and the target is on another.

The bridge provides a low latency path through which the processor may directly access PCI devices mapped anywhere in the memory or I/O address space of the computer system. The primary function of a bridge is to map the address space of one bus into the address space of another bus. The PCI bus defines three physical address spaces: memory, I/O space, and configuration space. Address decoding on the PCI bus is distributed; i.e. each device coupled to the PCI bus performs address decoding. The PCI specification defines two styles of address decoding: positive and subtractive. Positive decoding is faster since each PCI device is looking for accesses in the address range(s) the device has been assigned. Subtractive decoding can be implemented by only one device on the PCI bus since the subtractive decoding device accepts all accesses not positively decoded by some other agent.

Every PCI transfer begins with an address phase, during which an address/data bus (AD[31:0]) transfers the address and a command/byte enable (C/BE_[3:0]) bus transfers a command code. One or more data phases follow this phase, during which the same address/data bus transfers data and the command/byte enable bus transfers byte-enable signals. In a burst cycle, multiple data phases can follow a single address phase. In PCI terminology, the requesting PCI device is known as the initiator, and the addressed PCI device as the target. Every transfer starts with the activation of the frame (FRAME_) signal.

A device selection (DEVSEL#) signal is driven by the target to indicate that it is responding to the transaction. A device decodes the address/data lines and asserts a device select (DEVSEL#) signal if it owns the starting address. DEVSEL# may be driven with slow, medium or fast timing. If no agent asserts DEVSEL# within the slow timing parameters, the agent performing subtractive decoding may claim and assert DEVSEL#. The “#” or “_” refers to active low signals. More details on the PCI bus and particularly PCI addressing are found in the PCI Local Bus Specification, Production Version, Revision 2.1, dated June 1, 1995, which is published by the PCI Special Interest Group of Hillsboro, Oregon and hereby incorporated by reference.

The target indicates its readiness with an active target ready (TRDY_) signal. An active TRDY_ during a write access indicates that the target can take the data from the address/data bus. An active TRDY_ during a read access indicates that the requested data is available on the address/data bus.

5 In addition, the initiator must also indicate its readiness to the PCI bridge, through an active initiator ready (IRDY_) signal. An active IRDY_ during a write access indicates that the initiator has sent the write data on the address/data bus. In a read access, an active IRDY_ indicates that it accepts the data from the address/data bus. The initiator ends or interrupts the transfer by deactivating the FRAME_ signal. The target
10 can also stop the transfer by activation of a stop (STOP_) signal.

As defined, the PCI bus is limited to ten loads. A PCI device incorporated onto the motherboard is essentially one load and a PCI slot is considered two loads. Hence, a computer system with a processor/PCI bridge, three PCI slots and a PCI/ISA bridge is limited to two motherboard devices. Since oftentimes two motherboard devices is too
15 limiting, it is desirable to exceed the ten load limit.

One method of extending the number of loads described in the above-referenced PCI Specification is to use multiple PCI buses. Multiple PCI buses provide support for more devices than can be directly connected to one PCI bus. There are two ways to organize multiple PCI buses: as peers of each other or hierarchically. Peer buses require
20 multiple bridges on the processor bus that may affect loading. A hierarchical configuration has advantages if I/O patterns tend to be from one PCI device to another. If most I/O traffic goes into and out of memory, then peer buses make more sense. However, either bus configuration requires that the bridges be configured at startup to respond to an access on their primary bus only if the address falls into a specified range.
25 Furthermore, bridges separate the bus into two logical buses, thereby further complicating the configuration.

Each bridge includes address registers that are programmable through configuration space, so that the bridge responds to an access on its primary bus only if the address falls into the range specified by those registers; otherwise the access is
30 claimed by a subtractive decode agent. Only one set of address range response registers

is required by the PCI Specification, however if multiple buses are provided the complexity of the address may increase and multiple sets of registers are needed. Generally, the address range response registers are programmed to correspond to the addresses unused by the primary bus, instead of the memory space required by the secondary bus. Thus, the secondary side of the bridge responds to all memory accesses except those that fall into the ranges specified by the address response range registers. All transactions initiated on the secondary bus, outside of the programmed range, are responded to on the primary bus. It is the responsibility of system software to maintain the address response range registers of the bridges so that address decoding is properly performed.

The configuration of the address range register grows even more complicated if devices are allowed to be hot-pluggable. Hot pluggable devices, such as PC Cardbus cards, cause problems since the address range changes with the insertion or removal of the hot pluggable device. Therefore, it is desirable to remove this level of complexity and simultaneously provide a high number of PCI loads for sufficient functionality and expandability.

Transparent bridges have attempted to solve the configuration problem by making the bridge appear invisible to software. However, the performance of such bridges are often less than desirable. Cycles that were performed on the primary bus at top PCI bus speed take three times as long to complete on the secondary bus, if proper PCI timing conventions are followed. Therefore, a higher performing transparent bridge is desirable.

A computer system of the present invention includes a PCI bridge or repeater to increase the number of capacitive loads on a PCI bus without requiring significant changes to software. The PCI repeater connects a primary portion of the PCI bus to a secondary portion of the PCI bus. The portions act as one logical bus but are electrically separate for loading purposes. An arbiter controls access to the buses.

Transactions initiated on the primary bus and intended for a target on the secondary bus are downstream transactions. Transactions initiated on the secondary bus

and intended for a target on the primary bus are upstream transactions. Transactions initiated on the primary bus are echoed, passed or reflected to the secondary bus, and vice versa.

Signals are clocked through the PCI repeater, hence, a one clock delay is built in.
5 Because of the inherent delay, one of the buses will complete the transaction before the other. To prevent the earlier finishing bus from starting another transaction while the later finishing bus is completing the transaction, the arbiter removes any pending grants and will not grant the bus to any device on either bus until the later finishing bus has completed the transaction. This technique works especially well for bursted transactions
10 where the target is unable to move data at the same speed as an initiator.

Upstream transactions are handled like downstream transactions, except if there is a subtractive decode agent on the secondary such as an ISA bus bridge. Since only one subtractive decode agent can reside on a bus, the transaction is not subtractively decodable both upstream to the primary bus and downstream to the ISA bus. In a first
15 alternative, subtractive decoding to the ISA bus by a PCI to ISA bridge is disabled if the transaction is initiated on the secondary bus. This provides peer to peer transactions only between devices on the primary and secondary buses. In a second method, ISA subtractive decoding is enabled. After the transaction starts on the secondary bus, the secondary bus clock is halted to allow a target on the primary bus to claim the
20 transaction. If the transaction is not claimed the transaction finishes running on the secondary bus for either a secondary bus target to positively claim or the PCI to ISA bus bridge to subtractively claim. Thus, upstream address decoding is not needed on the PCI repeater.

In a third but preferred alternative, the repeater snoops configuration cycles on
25 the primary PCI bus to build an address map of memory and I/O address ranges corresponding to devices attached to the primary PCI bus. Configuration cycles of devices attached to the secondary PCI bus or ISA bus are not preferably snooped. The address map is built at initialization time or anytime a device is added or removed according to standard PCI and cardbus conventions. The address map provides a lookup
30 table for transactions initiated on the secondary PCI bus. The lookup table provides a

positive indication of whether the targeted device is on the primary PCI bus. If the target is on the primary PCI bus, the repeater will forward the transaction upstream. Otherwise, the transaction is assumed to be for a target on the secondary PCI bus or any devices attached to a subtractive decode agent, such as a PCI-to-ISA bridge.

5

A better understanding of the present invention can be obtained when the following detailed description of the preferred embodiment is considered in conjunction with the following drawings, in which:

10 Figure 1 is a block diagram illustrating a computer system according to the preferred embodiment;

 Figure 2 is a block diagram illustrating a PCI repeater according to the preferred embodiment;

 Figure 3 is a timing diagram illustrating a downstream single data phased write transaction followed by a downstream single data phase read transaction;

15 Figure 4 is a timing diagram illustrating a target disconnect write transaction followed by a target retry read transaction;

 Figure 5 is a timing diagram illustrating a downstream burst write sequence;

 Figure 6 is a timing diagram illustrating a downstream burst read sequence;

20 Figure 7 is a timing diagram illustrating a downstream burst read sequence without arbiter intervention;

 Figure 8 is a timing diagram illustrating a downstream burst read sequence with arbiter intervention;

 Figure 9 is a timing diagram illustrating an upstream single data phase write transaction followed by an upstream single data phase read transaction;

25 Figure 10 is an upstream target disconnect write transaction followed by an upstream read retry transaction;

 Figure 11 is a timing diagram illustrating an upstream burst write sequence;

 Figure 12 is a timing diagram illustrating an upstream burst read sequence;

Figure 13 is a timing diagram illustrating an upstream single phase write transaction positively claimed on the primary bus;

Figure 14 is a timing diagram illustrating an upstream single data phase write transaction echoed on the primary bus and subtractively claimed by the secondary bus;

5 Figure 15 is a timing diagram illustrating an upstream burst read sequence claimed on the primary bus;

Figure 16 is a block diagram illustrating logic of the PCI repeater for snooping configuration cycles; and

10 Figure 17 is a block diagram illustrating further logic for snooping configuration cycles.

Turning now to the drawings, a computer system C according to the preferred embodiment of the present invention is shown. To provide sufficient process and capability, the computer C deploys one or more processors 100 such as the Pentium Pro™ Processor available from Intel Corporation located in Santa Clara, California. The Pentium Pro™ Processor 100 includes a primary and secondary cache. Of course, other types of processors can be used with minimal changes. The Processor 100 is connected to a Pentium Pro™ host bus called the Processor Bus 102 which is a high performance bus utilizing gunning transceiver logic (GTL) as generally defined by the Pentium Pro™ specification.

In addition to the Processor 100, the processor bus 102 is connected to a data path unit (DP) 104, such as an Intel 82452GX, and a memory controller (MC) 106, such as an Intel 82453GX, which collectively form a memory control subsystem for a memory unit 108, which further connect to several memory interface components, such as an Intel 82451KX (not shown). The data path unit 104, the memory controller 106 and the memory interface components collectively form a memory control subsystem for the memory unit 108. The memory unit 108 includes several slots for receiving memory modules, such as a 72 pin extended data output (EDO) dynamic random access memory (DRAM) module. The memory controller 106 provides address, control and timing to

the memory unit 108, while the data path unit 104 interfaces a 72 bit data portion of the processor bus 102 to the memory unit 108. The memory controller 106 and the data path unit 104 are capable of receiving a memory request from the processor 100, queuing it and responding after the requested operation has completed. Additionally, the memory controller 106 provides memory error correction, including the capability of single bit and multi bit error detection on the fly. The memory controller 106 can handle up to four gigabytes of DRAM. Memory arrangements having non-interleaved, x2 and x4 interleaving configurations are supported by the memory control subsystem.

In addition to the memory subsystem, the processor bus 102 connects to one or more peripheral component interconnect (PCI) bridges 110, such as an Intel 82454GX. The PCI bridge 110 provides the necessary logic and control for passing bus cycles between the processor bus 102 and a primary PCI bus 112. It is noted here that although a hierarchical configuration is shown, the present invention works equally as well in peer-to-peer configurations.

Attached to the primary PCI bus 112 are one or more PCI devices 114, such as a small computer system interface (SCSI) controller 114a and a video system 114b. The SCSI controller 114a is connected to a hard disk drive 124 and the video system 114b, including video memory, is connected to a monitor 126. The PCI devices 114 also include Cardbus devices. In addition to the PCI devices 114, a PCI repeater or bridge 116 according to the present invention is connected between the primary PCI bus 112 and a secondary PCI bus 118. The PCI repeater 116 electrically isolates the secondary PCI bus 118 from the primary PCI bus 112 but causes both buses to appear as one logical PCI bus. The PCI repeater 116 performs this transparently to system software thereby increasing the number of PCI loads without adding significant complexity to the system software.

Attached to the secondary PCI bus 118 are one or more PCI devices or slots 120 and a PCI to ISA bridge 122. The PCI to ISA bridge 122 passes cycles between the secondary PCI bus 118 and an industry standard architecture (ISA) bus 128. Attached to the ISA bus 128 are a keyboard controller 130, a ROM 134 and a multi I/O unit 136 supporting two serial ports, a parallel port, a floppy disk controller connected to a floppy

drive and an integrated drive electronics (IDE) interface for connecting to an optional IDE hard disk drive or CD ROM drive.

Now turning to Figure 2, there is illustrated a block diagram of the PCI repeater 116. The repeater 116 is comprised of a primary bus unit 200 connected to the primary bus 112 and a secondary bus unit 202 connected to the secondary bus 118. Each bus unit has selectable voltage translators 204 for converting signal voltage levels to a common voltage, such as 3.3V. For example, if the primary PCI bus 112 and the PCI repeater 116 operated at 3.3V and the secondary PCI bus operated at 5V, the voltage translators 204 of the secondary side would convert output signals to 5V and input signals from 5V to 3.3V. To facilitate the voltage translation, the primary side is connected to a +V_p voltage and the secondary side is connected to a +V_s voltage to select the voltage. This is useful for applications such as a portable computer and docking station where the portable computer may operate at a 3.3V level to conserve power, but the docking station operates at the conventional 5V level. The repeater 116 can be placed in either the portable computer or the docking station to provide a connectable PCI bus to the docking station.

Also included in the repeater 116 are read prefetch buffers (RPB) 208 and write posting buffers (WPB) 210 for each side, and an arbiter 206 for arbitrating access between the primary PCI bus 112 and secondary PCI bus 118. The read prefetch buffers 208 are preferably two DWORDs deep, but could be any depth. An alternative embodiment can have an external arbiter. A single clock input is used to provide timing for both the primary and secondary bus units 200 and 202. Therefore, the primary and secondary PCI buses are required to operate at the same frequency. An alternative embodiment includes a second clock input and synchronization logic for operating the buses at different frequencies, although the PCI repeater 116 will only operate as fast as the slowest bus. The PCI repeater 116 also supports a zero power consumption state.

The primary bus unit 200 contains conventional configurable address decode logic for positively and subtractively claiming transactions from the primary PCI bus 112. The secondary bus unit 202 lacks any address decoding for reasons disclosed below.

The PCI repeater 116 is different than a conventional bridge because it prefers not to store and forward data. Generally, incoming signals are clocked on a rising edge of the PCI bus clock and passed through to the other bus. However, it is a function of the repeater to control the echoing or repeating of signals from one bus to another without
 5 violating any of the PCI operating rules. The functionality of the PCI repeater will be discussed with respect to downstream cycles and upstream cycles. Single and burst data cycles are subsets of each.

Downstream Transactions

10 Downstream transactions or cycles refer to transactions initiating on the primary PCI bus 112 and targeting a device on the secondary PCI bus 118, or buses below. The PCI repeater 116 is the only subtractive decode device on the primary PCI bus 112, in compliance with PCI conventions. Subtractively decoding the downstream transactions, rather than positively decoding, further eliminates BIOS (basic input output system) and
 15 operating system overhead. The PCI repeater 116 generally echoes on the secondary PCI bus 118 the entire transaction as it happens on the primary PCI bus 112 regardless of whether the transaction is intended for the secondary bus 118. The PCI repeater 116 reflects the primary PCI bus 112 delayed by one clock, therefore, the PCI repeater 116 adds two clocks to all downstream transactions - one for cycle start and one for data
 20 return. Hence, a transaction positively claimed with medium decode timing on the secondary bus, will appear to devices on the primary bus as a subtractively decoded transaction.

In a downstream transaction, signals echoed from the initiator on the primary PCI bus 112 to the target on the secondary PCI bus 118 are FRAME_p, AD[31:0]p,
 25 C/BE[3:0]p and IRDY_p. Signals echoed from the secondary PCI bus 118 to the initiator on the primary PCI bus 112 are TRDY_s, STOP_s and DEVSEL_s. In an upstream transaction, signals echoed from the initiator on the secondary PCI bus 118 to the target on the primary PCI bus 112 are FRAME_s, AD[31:0]s, C/BE[3:0]s and IRDY_s. Signals echoed from the primary PCI bus 112 to the initiator on the secondary PCI bus 118 are
 30 TRDY_p, STOP_p and DEVSEL_p.

Now referring to Figure. 3, there is illustrated a timing diagram of two single DWORD transactions initiated on the primary PCI bus 112 and targeted to the secondary PCI bus 118. The first transaction is a write transaction positively claimed by a device on the secondary bus 118 using medium timing and the second transaction is a read transaction subtractively decoded on the primary bus 112 and secondary bus 118. In the following Figures, signal names ending with a small "p" are primary PCI bus 112 signals and signal names ending with a small "s" are secondary PCI bus 118 signals, as shown in Figure 2. The underscore after the signal name indicates the signal is active low. The broken circles indicate a turn-around cycle as defined in the PCI specification.

10 In response to a transaction initiated on the primary PCI bus 112, the PCI repeater 116 asserts FRAME_s, AD[31:0]s, and C/BE_[3:0]s one clock delayed (clocks 2 and 8) on the secondary PCI bus 118. IRDY_s is also echoed onto the secondary PCI bus 118 on clocks 3 and 9. Although not shown, other master signals, such as a LOCK_ and IDSEL signals are also echoed as needed on the secondary PCI bus 118.

15 The PCI repeater 116 senses DEVSEL_p deasserted on clocks 5 and 11, and therefore, accepts the cycle using subtractive decode timing on behalf of the secondary PCI bus 118. For write transactions, once DEVSEL_s is sampled low (clock 5), the PCI repeater 116 copies the states of the slave signals (DEVSEL_s, TRDY_s and STOP_s) from the secondary PCI bus 118 to the primary PCI bus 112 (clock 6). After the PCI
20 repeater 116 senses DEVSEL_s and TRDY_s asserted low (clock 5), IRDY_s is deasserted to complete the transaction on the secondary PCI bus 118. The transaction completes on the primary PCI bus 112 one clock later (clock 6). Echoing DEVSEL_s and TRDY_s to the primary PCI bus 112 ensures that the issuing master has the final responsibility for executing the cycle.

25 For read cycles, once the PCI repeater 116 senses DEVSEL_s asserted (clock 13), the PCI repeater 116 copies the states of the slave signals (DEVSEL_s, TRDY_s, STOP_s and AD[31:0]s) from the secondary PCI bus 118 to the primary PCI bus 112. The PCI repeater 116 preferably does not post any data transfer that consists of a single data phase. This simplifies the design of the PCI repeater 116 and allows the repeater

116 to back out from any transfer which is retried by a target. The resulting latency across the PCI repeater 116 is 2 clocks in addition to the target latency.

Now referring to Figure 4 there is illustrated a target disconnect write transaction followed by a target retry read transaction. In response to a write transaction initiated on the primary bus 112, the PCI repeater 116 asserts `FRAME_s`, `AD[31:0]s`, and `C/BE_[3:0]s` one clock later (clocks 2 and 8) on the secondary PCI bus 118. The assertion `IRDY_p` is also echoed on the secondary PCI bus 118 on clocks 3 and 9.

The PCI repeater 116 senses that `DEVSEL_p` remains deasserted on clocks 5 and 11, and therefore, accepts the transaction using subtractive decode timing on behalf of the secondary PCI bus 118. For write cycles, once `DEVSEL_s` is sampled asserted (clock 5), the PCI repeater 116 copies the states of the slave signals (`DEVSEL_s`, `TRDY_s` and `STOP_s`) from the secondary PCI bus 118 to the primary PCI bus 112 (clock 6). A write transaction disconnected on the secondary bus 118 is also disconnected on the primary PCI bus 112 because the `STOP_p` signal is driven on the primary side concurrently with `TRDY_p`. Waiting on `DEVSEL_s` and `TRDY_s` before accepting the data on the primary PCI bus 112 ensures that the initiating master has the final responsibility for executing the transaction.

For read transactions, once the repeater 116 senses `DEVSEL_s` asserted (clock 13), the repeater 116 copies the states of the slave signals (`AD[31:0]s`, `DEVSEL_s`, `TRDY_s` and `STOP_s`) from the secondary PCI bus 118 to the primary PCI bus 112 (clock 14). A retried transaction on the secondary bus 118 is also retried on the primary bus 112 because the `STOP_p` signal is driven on the primary side concurrently with `TRDY_p` negated.

Now referring to Figure 5, there is illustrated a downstream burst write transaction. Write bursting across the PCI repeater 116 is achieved by accepting the burst data with zero wait states, once the transfer is accepted by the target. The PCI repeater 116 senses a burst sequence at clock 3 because both `IRDY_p` and `FRAME_p` are asserted. Since the transaction is a burst write transaction, the repeater does not assert `IRDY_s` at clock 3, as would happen for a single DWORD transaction. Instead, the PCI repeater 116 delays the assertion of `IRDY_s` on the secondary bus 118 until clock 6, the

clock following the repeater's 116 accepting the first data transaction using subtractive decode timing on the primary bus 112 at clock 5. The delay is needed to insure that the PCI repeater has the next transaction's byte enables (which are available at clock 8, the first clock following TRDY_p asserted) in time to meet the PCI Specification timing.

- 5 In response to the transaction initiated on the primary PCI bus 112, the PCI repeater 116 asserts FRAME_s, AD[31:0]s and C/BE[3:0]s one clock delayed (clock 2) on the secondary bus 118. A request/grant signal pair (REQ1_/GNT1_) corresponds to a master on the primary side requesting/gaining access to the bus 112. The PCI repeater 116 senses DEVSEL_p deasserted on clock 5, and so subtractively decodes the
- 10 transaction on behalf of the secondary bus 118 by asserting DEVSEL_p on the primary bus in clock 5. On the secondary bus 118, a target 120 decodes the address and command using medium decode timing and asserts DEVSEL_s and TRDY_s in clock 4. In turn, the PCI repeater 116 senses the assertion of DEVSEL_s and TRDY_s and echoes TRDY_p onto the primary bus one clock later (clock 5) to begin the subsequent
- 15 data transactions. Although medium decode timing is illustrated on the secondary bus, any timing is supported with the effect being a longer first data transaction. The PCI repeater 116 accepts the data with zero wait states. If the target 120 is unable to accept the data with zero wait states, the PCI repeater buffers the data in the write posting buffer 210. The secondary side delivers the data at a rate limited to the speed of the target.
- 20 Due to possible differences in speed between the initiator and the target, the PCI repeater could end up with unfinished cycles in its write posting buffer 210. If a bus master begins another transaction before the PCI repeater 116 empties its write posting buffer 210, a deadlock condition could develop because the PCI repeater 116 would not relinquish the secondary PCI bus 118 until it empties its write posting buffer 210.
- 25 Therefore, according to the preferred embodiment, to prevent any initiator from obtaining the primary bus 112 while the PCI repeater 116 is busy, the PCI repeater 116 notifies the PCI arbiter 111 of the need to stop granting the primary PCI bus 112 while it is finishing the burst sequence on the secondary PCI bus 118. At the same time no other initiator can access the secondary bus 118 because the repeater will not relinquish the secondary PCI
- 30 bus 118 until it empties its write posting buffer 210. A sideband signal no more grants

(NOMOGNTS_) forces the arbiter 111 to stop granting requests. While the NOMOGNTS_ signal is asserted, the arbiter 111 deasserts any pending grant signals, for example (GNT1). When the burst sequence is completed on the secondary PCI bus 118 (clock 10), the PCI repeater 116 deasserts the NOMOGNTS_ signal. After the
 5 NOMOGNTS_ signal is deasserted, the PCI arbiter 111 is free to issue grant signals again.

On clock 6, the first data is accepted by the repeater 116, on the primary side, which causes the data and byte enables for the next transaction to become available. The PCI repeater 116 then asserts IRDY_s to allow the write data transfer to occur. The write
 10 transfers then continue until the initiator is done, or disconnected by the PCI repeater 116 asserting STOP_p as shown in Figure 5, clock 9. The STOP_p signal is asserted if the write posting buffers are approaching their limit, as in this example. Although the PCI repeater 116 adds additional latency to the transaction, the effective burst rate approaches the limit of the PCI bus.

15 Now referring to Figure 6, there is illustrated an attempted burst read transaction of the prior art between a master on a primary PCI bus, such as PCI bus 112 and a target on a secondary PCI bus, such as secondary PCI bus 118 using the memory read command. The dotted lines illustrate an undesirable condition which the present invention solves.

20 If a master on the primary PCI bus 112 attempts to read more than one DWORD across the PCI repeater 116 (or prior art repeater), the PCI repeater 116 will terminate the transaction after a single data phase (clock 7) because the PCI repeater does not have the next set of byte enables from the requesting master. The master must subsequently perform another transaction to read the remaining data. Hence, the transaction is broken
 25 into multiple single data transfers.

Figure 6 also shows a secondary bus grant (SBGNT_) signal. This signal is provided by an arbiter, such as arbiter 111, to enable a secondary bus arbiter, such as secondary bus arbiter 206. Normally, the SBGNT_ signal is asserted on clock 2 to allow the secondary arbiter to issue grants, such as GNT2_. Thus, once the secondary PCI bus

118 has completed the transaction (clock 5), the secondary bus device asserting the request corresponding to GNT2_ will own the secondary PCI bus (clock 6).

Because of the inherent latencies of the PCI repeater 116 (and prior art repeater), a transaction on the secondary PCI bus 118 could begin (clock 6) before the primary PCI
5 bus 112 is ready to receive the echoed transaction, as shown by the transaction in dotted lines starting on the secondary PCI bus in clock 6.

To prevent this problem from occurring, the PCI repeater 116 of the present invention asserts the NOMOGNTS_ signal (clock 3) as soon as the PCI repeater 116 detects a memory read command and FRAME_p is held asserted indicating that the
10 master on the primary PCI bus 112 wants to burst the read transaction. The NOMOGNTS_ signal is held asserted until the PCI repeater 116 signals a disconnect to the master on the primary PCI bus 112 (TRDY_p and STOP_p asserted in clock 6). This prevents the arbiters 111 and 206 from asserting grants signals until the transaction is completed on both buses. If the repeater 116 detects that the master and the target are
15 on the same bus (i.e. primary PCI bus 112), it instead immediately deasserts the NOMOGNTS_ signal after DEVSEL_p is asserted (not shown) to allow the arbiter 111 to pipeline the grants as normal. Thus, the PCI repeater 116 handles read transactions to non-prefetchable regions where the read transaction is broken into multiple single data transfers.

Referring now to Figure 7, there is illustrated another problem of the prior art PCI
20 repeater with regard to memory read line and memory read multiple PCI commands. The memory read line and memory read multiple commands are used to access data in address ranges that are prefetchable. Figure 7 illustrates a case where a master on the primary PCI bus initiates a memory read line or a memory read multiple command. The
25 prior art PCI repeater would start the cycle on the secondary side and continue requesting new data until it samples FRAME_p deasserted. A problem develops if the target adds waits states to the transfer and therefore keeps the secondary PCI bus busy while the primary PCI bus is free to begin another cycle that targets a slave on the primary PCI bus (clock 10). In this case, the prior art PCI repeater would miss an entire transaction since
30 the transaction on the secondary bus does not end until clock 12.

Now referring to Figure 8, there is illustrated a downstream burst read sequence with arbiter intervention according to the preferred embodiment, solving the problem of Figures 6-7. A master initiates a memory read line or memory read multiple command on the primary PCI bus 112 to a target on the secondary PCI bus 118. The PCI repeater 116 initiates the command on the secondary PCI bus 118 on clock 2. The PCI repeater 116 sets the byte enables to all zeros regardless of the requesting master byte enables to read all the bytes and to be able to read ahead (or prefetch) of the master on the primary bus 112 (since byte enables are not pipelined). On clock 5, because the transaction is not positively claimed on the primary PCI bus 112, the PCI repeater 116 determines that the transaction is heading downstream and asserts the `NOMOGNTS_` signal to notify the PCI arbiter 111 to remove the current grant (`GNT1_`) on the primary PCI bus 112 and any subsequent grants until the secondary PCI bus 118 finishes its current transaction. The `NOMOGNTS_` signal is deasserted on clock 12 when the read completes on the secondary PCI bus 118.

The read prefetching is terminated on the secondary PCI bus 118 as soon as the PCI repeater 116 detects the last data phase on the primary PCI bus 112. The last data phase is signaled by the deassertion of `FRAME_p` on clock 8 and of `IRDY_p` on clock 9. On clock 9 the PCI repeater 116 signals its last data phase is completed on the secondary PCI bus 118 when the `FRAME_s` signal is deasserted on clock 9. Thus, the master on the primary PCI bus 112 finishes reading while the PCI repeater 116 is still reading the next DWORD from the target on the secondary PCI bus 118. So as not to violate PCI protocol, the PCI repeater holds the `C/BE[3:0]s` and `IRDY_s` signals on the secondary PCI bus 118 until the last data phase can be completed on the secondary PCI bus 118 at clock 12. After the PCI repeater 116 detects the last data phase, the repeater 116 drives the `AD[31:0]p`, and `C/BE[3:0]p` buses to a valid state (from clock 11 until clock 12), when the secondary PCI bus 118 read transaction is ended. Therefore, the PCI repeater 116 attempts to stay ahead of the requesting master by requesting more data. Unused data is discarded by the PCI repeater 116.

30 Upstream Transactions

Upstream transactions are transactions initiating on the secondary PCI bus 118 and targeting a device on the primary PCI bus 112. The PCI repeater handles upstream transactions the same way it handles downstream transactions, but with a few exceptions. The PCI repeater 116 will not respond to upstream configuration cycles.

5 One problem facing the PCI repeater 116 is how to determine which cycles are headed upstream and which are headed downstream. Two alternative solutions are possible. In a first alternative, the subtractive decode logic of the PCI to ISA bridge is enabled during downstream transactions only. The PCI repeater 116 broadcasts every transaction originated on the secondary PCI bus 118 to the primary PCI bus 112. If the
10 transaction is not positively claimed by a device on the secondary PCI bus 118, it is subtractively claimed by the repeater 116. Thus, the transaction is sent upstream, however, peer to peer transactions between devices on the secondary PCI bus 118 and devices on the ISA bus 128 are not available. In a second alternative, the PCI repeater 116 halts operation on the secondary PCI bus 118 and echoes the transaction to the
15 primary PCI bus 112. If the target is on the primary PCI bus 112, the target positively claims the transaction. If the transaction is not positively claimed by a primary bus PCI agent, the PCI repeater 116 subtractively claims the transaction and runs it on the secondary PCI bus 118. If the target is an ISA device, the PCI to ISA bridge 122 subtractively claims the transaction from the secondary PCI bus. This preferred
20 alternative has the advantage of handling a hierarchy of buses.

Figures 9-12 correspond to the first alternative. The principles illustrated in the Figures and described below apply equally to memory or I/O transactions. Referring now to Figure 9, there is illustrated two single DWORD transactions initiated on the secondary PCI bus 118 and terminated on the primary PCI bus 112. The first transaction
25 is a write transaction positively claimed by a device on the primary PCI bus 112 using medium decode timing and the second transaction is a read transaction claimed using medium decode timing. The first upstream transaction starts on clock 1 on the secondary PCI bus 118. The transaction is echoed upstream to the primary PCI bus 112 on clock 2. The PCI repeater 116 senses `FRAME_s` deasserted on clock 3, determines that this
30 is a single data phase transaction, and asserts `IRDY_p` to allow the write transaction to

complete. The transaction is positively decoded on the primary PCI bus 112 on clock 4 using medium decode timing. As soon as the repeater 116 samples DEVSEL_p asserted on clock 5, it copies the state of the DEVSEL_p on to DEVSEL_s along with the remaining slave signals (TRDY_p and STOP_p).

5 The upstream read transaction follows in a similar manner. The transaction starts on the secondary PCI bus 118 on clock 7 and is echoed to the primary PCI bus 112 on clock 8. The transaction is accepted on the primary PCI bus 112 on clock 10 and DEVSEL_s is echoed to the secondary PCI bus 118 on clock 11. On clock 12, the target places the requested data on the primary PCI bus 112 and asserts TRDY_p to terminate
10 the transaction. On clock 13, AD[31:0]p and TRDY_p are echoed to the secondary PCI bus 118.

Now referring to Figure 10, there is illustrated a target disconnect write transaction followed by a target retry read transaction. Because the PCI repeater 116 does not post single data phase transactions, the PCI repeater 116 handles upstream target
15 disconnect and retries the same way it handles the downstream transactions. On clock 4, a target on the primary PCI bus 112 signals a disconnect which in turn is echoed to the initiator on the secondary PCI bus 118 by the PCI repeater 116 on clock 5. In a similar fashion, a retry is signaled by a target on the primary PCI bus 112 on clock 12, which in turn is echoed to an initiator on the secondary PCI bus 118 by the PCI repeater 116 on
20 clock 13.

Now referring to Figures 11 and 12, there are illustrated an upstream burst write sequence and upstream read sequence respectively. Upstream burst transactions are similar to downstream burst transactions in that the arbiter 111 must not grant the PCI bus 112 and 118 to any agent until after the current transaction completes. This is
25 accomplished by asserting the NOMOGNTS_ signal as soon as the PCI repeater 116 determines that the transaction originated on the secondary PCI bus 118 is accepted by a target on the primary PCI bus 112. The target on the primary PCI bus 112 indicates its acceptance of the transaction by asserting DEVSEL_p, as shown at clock 3 with fast decode timing. The NOMOGNTS_ signal remains asserted until the transaction
30 completes on the primary PCI bus 112, as shown from clocks 4-11. In these examples,

it can be seen that the NOMOGNTS_ signal remains asserted even while the target on the primary PCI bus 112 inserts wait states (clocks 8-9 in Figure 11 and clock 9 in Figure 12) before accepting the final write cycle.

Figures 13-15 correspond to the second alternative. Figure 13 illustrates an upstream single data phase write transaction positively claimed on the primary PCI bus 112. Figure 14 illustrates a upstream single data phase write transaction not positively claimed on the primary PCI bus 112 or the secondary PCI bus 118, but subtractively claimed on the ISA bus 128. The principles illustrated in these Figures apply equally to memory and I/O transactions. In Figure 13, the transaction starts on the secondary PCI bus 118 on clock 1 and is echoed to the primary PCI bus 112 on clock 2.

In order to allow a prospective primary bus target sufficient time to respond, the secondary PCI bus 118 clock (CLKs) is halted for two PCI clock cycles by a clock disable (CLK_DIS) signal at the end of clock 2. The two clock delay allows the transaction to be claimed by either a primary PCI bus 112 target or secondary PCI bus 118 target before the ISA to PCI bridge 122 would subtractively claim the transaction. A primary bus agent could claim the transaction by asserting DEVSEL_p on clocks 3, 4 or 5. A secondary bus agent could claim the transaction by asserting DEVSEL_s on clocks 2, 5 or 6. At the end of clock 4, the CLKs is started again.

In Figure 13, a primary bus target claims the transaction with slow decode timing by asserting DEVSEL_p on clock 5. On clock 6, the PCI repeater senses the assertion of DEVSEL_p and echoes DEVSEL_p and TRDY_p onto the secondary PCI bus 118 and the transaction is terminated on the secondary bus 118 - that is, it will not be subtractively decoded by the ISA to PCI bridge 122.

Now turning to Figure 14, the transaction starts on the secondary PCI bus 118 on clock 1 and is echoed to the primary PCI bus 112 on clock 2 to allow the conventional decode logic of the primary bus unit 200 to determine if the address range corresponds to a target on the primary PCI bus 112. The secondary PCI bus clock (CLKs) is halted again for two clocks as in Figure 13. However, this time the transaction is not claimed on the primary PCI bus 112. On clock 6, the PCI repeater 116 senses the negated state of DEVSEL_p, determines that the target is not on the primary PCI bus 112 and asserts

DEVSEL_p to subtractively claim the transaction from the primary bus 112. On the secondary PCI bus 118, the transaction has also been unclaimed. On clock 7, the PCI repeater 116 senses that DEVSEL_s is still unasserted and the assertion of DEVSEL_p with subtractive decode timing and therefore determines that the target is on the ISA bus 128. On clock 7, the ISA to PCI bridge senses that DEVSEL_s is still unasserted and asserts DEVSEL_s to subtractively claim the transaction. The transaction completes as normal on clock 9. Thus, the target of a transaction is transparently determined without requiring any special upstream address decoding logic to be contained in the PCI repeater 116.

Now referring to Figure 15, there is illustrated an upstream prefetchable burst read sequence. In this second alternative, upstream burst transactions are similar to downstream burst transactions in that the arbiter 111 must not grant the PCI bus 112 and 118 to any agent until after the current transaction completes. The transaction starts on the secondary PCI bus on clock 1 and is echoed to the primary PCI bus 112 on clock 2 to determine if the target is on the primary bus. The byte enables C/BE[3:0]_p are forced to zero so that prefetching may occur. IRDY_s is asserted by the initiator on the secondary bus 118 on clock 2 and is echoed to the primary PCI bus 112 on clock 3.

While prospective targets on the primary PCI bus 112 are decoding the transaction, the PCI clock of the secondary PCI bus (CLKs) is halted by the PCI repeater 116 for two clocks. At clock 5, the secondary PCI bus clock is started again. A target on the primary PCI bus 112 accepts the transaction by asserting DEVSEL_p on clock 4 and the PCI repeater 116 echoes the signal (DEVSEL_s) onto the secondary bus on clock 5. Thus, the subtractive decode agent of the secondary PCI bus 118 assumes the transaction was positively claimed on the secondary bus 118.

On clock 5, the PCI repeater 116 senses FRAME_p and DEVSEL_p asserted and asserts NOMOGNTS_p in response to preclude the arbiter 111 from granting the PCI bus to a secondary bus initiator until the current sequence has completed on the primary PCI bus 112. On clock 9, the PCI repeater 116 senses the completion of the burst read sequence and in response deasserts the NOMOGNTS_p signal.

If the burst sequence had not been claimed by a target on the primary PCI bus 112, the sequence would have been claimed by a target on the secondary PCI bus 118. By first sending the transaction upstream for decoding on the primary bus before running the transaction on the secondary bus, the PCI repeater 116 can use the inherent address
5 decoding logic of the PCI devices. Thus, the PCI repeater 116 does not need any special downstream or upstream decode logic to handle these transactions.

Now referring to Fig. 16, there is illustrated a third but preferred alternative for the PCI repeater 116. In particular, this embodiment includes logic for deciding which transaction to forward from the secondary PCI bus 118 to the primary PCI bus 112. As
10 before, the PCI repeater 116 includes the primary and secondary bus units 200 and 202 to handle the reception and transmission of both master and slave cycles. Additionally, the PCI repeater 116 includes snooping logic 212 for snooping configuration cycles on the primary PCI bus 116.

Configuration cycles are bus transactions specific to configuration address space
15 for writing or reading configuration registers. Each PCI device 114 is required by the PCI Local Bus Specification to support a minimum set of configuration space registers. The configuration registers define a memory range and/or I/O space address range in which the PCI device 114 will respond. Thus, the PCI devices 114 are location programmable. Configuration cycles and configuration registers are more fully defined
20 in the PCI Local Bus Specification which is hereby incorporated by reference.

As configuration cycles are snooped by the snooping logic 212, the address ranges are stored in an address map 214 by the primary bus unit 200. The address map 214 includes ranges for both memory and I/O space. Since the repeater typically forwards transactions from the primary PCI bus to the secondary PCI bus, the snoop
25 cycles for the primary PCI bus are distinguished by configuration cycles terminated with the TRDY_P signal. Configuration cycles terminated on the secondary PCI bus 118 are snooped (since they originated on the primary PCI bus 112), but the data carried in the configuration cycle is not stored in the address map 214.

The configuration cycles are defined in the PCI Local Bus Specification to be
30 required at initialization time and anytime a PCI device 114 [or cardbus device] is added

or removed from the computer system C. Thus, the address map 214 is automatically built by the snooping logic 212 and the primary bus unit 200 without any special configuration software.

Once the address map 214 is built, transactions originating on the secondary PCI bus 118 are decodable in a positive fashion, as compared to the technique of Figs. 13-15. The secondary bus unit 202 responds to a secondary bus transaction by examining the address map to determine if the address belongs to a PCI device 114 attached to the primary PCI bus 112. If so, the transaction is positively claimed and forwarded to the primary PCI bus 112. If not so, the transaction remains on the secondary PCI bus to be subtractively claimed or positively claimed by a PCI device 114 attached to the secondary PCI bus 118. Thus, the PCI repeater 116 is able to route transactions according to the address map 214 without system intervention.

As an additional benefit to having the address map 214 in the PCI repeater 116, transactions originating on the primary PCI bus 112 no longer need to be subtractively claimed by the primary side of the PCI repeater 116. When a transaction is initiated on the primary PCI bus 112, the primary bus unit 200 examines the address map 214 to determine if the target device is attached to the primary PCI bus 112. If so, the transaction is not forwarded. However, if the address map indicates that the target device is not attached to the primary PCI bus 112, the transaction is positively claimed for forwarding to the secondary PCI bus 118.

Referring now to Fig. 17, there is illustrated a circuit diagram of the snoop detect logic 212. Included in the snoop detect logic 212 is an AND gate 220 receiving bits 1-3 of the command/byte enable (C/BE#) signals. Bit 2 is inverted so that when C/BE = 101 the output of the AND gate 220 is high or asserted. The output of the AND gate 220 is received into a non-inverted input of an AND gate 222. The FRAME# signal is received into a D input of a flip-flop 224 which is clocked with the PCI clock signal to provide a synchronous frame (SFRAME#) signal. The SFRAME# signal is a one-shot FRAME# signal provided when the FRAME# signal is first asserted. The FRAME# signal is received into a second inverted input of the AND gate 222 and the SFRAME# signal is

received into a third non-inverted input of the AND gate 222. The output of the AND gate 222 is a high or asserted when a configuration cycle is detected (SNOOP_CONFIG).

5 The SNOOP_CONFIG signal is qualified with the TRDY_P signal so that only configuration cycles of devices attached to the primary PCI bus 112 are stored in the address map. It is noted that both I/O space addresses and memory space addresses are stored in the address map.

Therefore, there is disclosed herein a novel apparatus and method for repeating PCI transactions across a bridge transparently of system software.

10 The foregoing disclosure and description of the invention are illustrative and explanatory thereof, and various changes in the size, shape, materials, components, circuit elements, wiring connections and contacts, as well as in the details of the illustrated circuitry and construction and method of operation may be made without departing from the spirit of the invention.

CLAIMS

What is claimed is:

- 1 1. A method of communicating a transaction between a device coupled to
2 a first bus and a device coupled to a second bus, the devices having configuration
3 registers for holding a programmable base address, the base address being written during
4 a configuration cycle, each device individually responding to a configuration cycle, the
5 method comprising the steps of:
6 (a) snooping configuration cycles communicated over the first bus, said
7 configuration cycles including a base address;
8 (b) storing the base address contained in a configuration cycle responded to
9 by a device attached to the first bus, the base address being stored in an address map;
10 (c) forwarding a transaction originating on the second bus to the first bus if
11 the transaction is addressed to a device contained in the address map.

- 1 2. The method of claim 1, wherein the configuration cycle responded to by
2 the device attached to the first bus is terminated with a first bus target ready signal.

- 1 3. The method of claim 1, wherein step (c) further comprises:
2 (d) positively claiming the bus transaction on the second bus.

- 1 4. The method of claim 1, wherein steps (a) and (b) are repeated for each
2 device configured.

- 1 5. The method of claim 1, wherein the base address is a memory address.



1 6. The method of claim 1, wherein the base address is an input/output space
2 address.

1 7. A bus repeater for coupling a first bus to a second bus, the first bus for
2 having a first device coupled thereto, the second bus for having a second device coupled
3 thereto, the repeater comprising:
4 a first bus unit operable to communicate bus transactions with the first
5 bus, the transactions including an address;
6 snooping logic operable to snoop configuration cycles on the first bus and
7 to provide a snoop indication when a device on the first bus acknowledges a
8 configuration cycle, the configuration cycles including data for indicating a device
9 address range;
10 an address map operable to hold a plurality of device address ranges, said
11 address map receiving the device address ranges from the first bus unit and storing the
12 device address ranges in response to the snoop indication; and
13 a second bus unit operable to communicate bus transactions with the
14 second bus and the first bus unit, said second bus unit operable to positively claim and
15 echo a bus transaction from the second bus to the first bus if the bus transaction address
16 is contained in the address map.

1 8. The repeater of claim 7, wherein the first bus unit is also operable to
2 positively claim bus transactions originating on the first bus for devices attached to the
3 second bus if the address corresponding to the bus transaction is not contained in the
4 address map.

1 9. The repeater of claim 7, wherein the first bus unit is operable to
2 subtractively claim bus transactions for the second bus.



1 10. The repeater of claim 7, further comprising:
2 said second bus unit operable to not claim a bus transaction from the
3 second bus to the first bus if the bus transaction is not contained in the address map.

1 11. A computer system, comprising:
2 a first bus;
3 a processor coupled to said first bus;
4 memory coupled to said first bus;
5 one or more devices coupled to said first bus, one of the devices being a
6 hard disk system;
7 a second bus;
8 one or more devices coupled to said second bus, one of the devices being
9 a being a repeater coupled between the first bus and the second bus, the repeater
10 comprising:
11 a first bus unit operable to communicate bus transactions with the
12 first bus, the transactions including an address;
13 snooping logic operable to snoop configuration cycles on the first
14 bus and to provide a snoop indication when a device on the first bus acknowledges a
15 configuration cycle, the configuration cycles including data for indicating a device
16 address range;
17 an address map operable to hold a plurality of device address
18 ranges, said address map receiving the device address ranges from the first bus unit and
19 storing the device address ranges in response to the snoop indication; and
20 a second bus unit operable to communicate bus transactions with
21 the second bus and the first bus unit, said second bus unit operable to positively claim
22 and echo a bus transaction from the second bus to the first bus if the bus transaction
23 address is contained in the address map.



1 12. The computer system of claim 11, wherein the first bus unit is also
2 operable to positively claim bus transactions originating on the first bus for devices
3 attached to the second bus if the address corresponding to the bus transaction is not
4 contained in the address map.

1 13. The computer system of claim 11, wherein the first bus unit is operable
2 to subtractively claim bus transactions for the second bus.

1 14. The computer system of claim 11, further comprising:
2 said second bus unit operable to not claim a bus transaction from the
3 second bus to the first bus if the bus transaction is not contained in the address map.

