METHOD AND APPARATUS FOR REDUCING IDLE LINK POWER CONSUMPTION IN A PLATFORM

VERFAHREN UND VORRICHTUNG ZUR REDUZIERUNG DES STROMVERBRAUCHS EINER PLATTFORMVERBINDUNG IM RUHEMEODUS

PROCÉDÉ ET APPAREIL DESTINÉS À DIMINUER LA CONSOMMATION ÉLECTRIQUE D’UNE LIAISON DE COMMUNICATION INACTIVE

Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR

Priority: 01.07.2011 US 201113175574

Date of publication of application: 07.05.2014 Bulletin 2014/19

Divisional application: 18152405.9 / 3 330 838

Proprietor: Intel Corporation
Santa Clara, CA 95054 (US)

Inventors:
• DIEFENBAUGH, Paul S.
  Portland, Oregon 97229 (US)
• GOUGH, Robert E.
  Sherwood, Oregon 97140 (US)
• BACHRACH, Yuval
  34990 Haifa (IL)
• HUNSAKER, Mikal C.
  El Dorado Hills, California 95762 (US)
• BEN-TAL, Rafi
  20692 Yokneam (IL)
• PARDO, Ilan
  47246 Ramat-Hasharon (IL)

References cited:
US-B1- 7 188 263


RAVIBUDRUK ET AL: "PASSAGE - PCI EXPRESS SYSTEM ARCHITECTURE", PCI EXPRESS SYSTEM ARCHITECTURE, ADDISON-WESLEY, US, 4 September 2003 (2003-09-04), page 48pages, XP007901254,
This invention relates to a platform, and more specifically but not exclusively, to a method and apparatus to reduce the idle link power in a platform.

BACKGROUND DESCRIPTION

[0002] High speed serial communication links are often used in a platform to provide fast data access. However, these high speed serial communication links often require high idle power. In power sensitive platforms such as handheld and tablet computers, high speed serial communication links may not be suitable due to the high idle power requirements.


BRIEF DESCRIPTION OF THE DRAWINGS

[0004] The features and advantages of embodiments of the invention will become apparent from the following detailed description of the subject matter in which:

- Figure 1 illustrates a block diagram of a platform in accordance with one embodiment of the invention;
- Figure 2 illustrates a block diagram of a host device and an endpoint device in accordance with one embodiment of the invention;
- Figure 3 illustrates a block diagram of the link states in accordance with one embodiment of the invention;
- Figure 4 illustrates a timing diagram of an upstream port initiated link state exit in accordance with one embodiment of the invention;
- Figure 5 illustrates a timing diagram of a downstream port initiated link state exit in accordance with one embodiment of the invention;
- Figure 6 illustrates a timing diagram of an upstream port initiated link state exit in accordance with one embodiment of the invention;
- Figure 7 illustrates a timing diagram of a downstream port initiated link state exit in accordance with one embodiment of the invention; and
- Figure 8 illustrates a system to implement the methods disclosed herein in accordance with one embodiment of the invention.

DETAILED DESCRIPTION

[0005] Embodiments of the invention described herein are illustrated by way of example and not by way of limitation in the accompanying figures. For simplicity and clarity of illustration, elements illustrated in the figures are not necessarily drawn to scale. For example, the dimensions of some elements may be exaggerated relative to other elements for clarity. Further, where considered appropriate, reference numerals have been repeated among the figures to indicate corresponding or analogous elements.

[0006] Reference in the specification to "one embodiment" or "an embodiment" of the invention means that a particular feature, structure, or characteristic described in connection with the embodiment is included in at least one embodiment of the invention. Thus, the appearances of the phrase "in one embodiment" in various places throughout the specification are not necessarily all referring to the same embodiment.

[0007] The terms "upstream" and "downstream" are used to illustrate the direction of the traffic or data flow in the platform in one embodiment of the invention and are not meant to be limiting. The terms "upstream" and "downstream" may be interchanged in another embodiment of the invention. Other terminology to describe the direction of the traffic or data flow in the platform can be used without affecting the workings of the invention.

[0008] Embodiments of the invention provide a method and apparatus to reduce the idle power in a platform according to appended claims 1 and 3. In one embodiment of the invention, the host and its coupled endpoint(s) in the platform each has a low power idle link state(s) that allows disabling of the high speed link circuitry in both the host and its coupled endpoint(s). This allows the platform to reduce its idle power as both the host and its coupled endpoint(s) are able to turn off their high speed link circuitry in one embodiment of the invention. The disabling of the high speed link circuitry includes, but is not limited to, power gating of the voltage input(s), shutting off the input clock(s) and any other technique to reduce the power consumption of the high speed link circuitry.

[0009] In one embodiment of the invention, the low power idle link state is a sub-state of an existing low power link state. In another embodiment of the invention, the low power idle link is added as a new link state of the platform. The platform includes, but is not limited to, a desktop computer, a laptop computer, a netbook computer, a handheld computer, a tablet computer, a notebook computer, a cellular telephone, a smartphone, a mobile computing device, an Internet appliance or any other type of computing device.

[0010] Figure 1 illustrates a block diagram 100 of a platform 105 in accordance with one embodiment of the invention. In one embodiment of the invention, the platform 105 is compliant at least in part with the Peripheral
Component Interconnect (PCI) Express (PCIe) standard or specification maintained by the PCI Special Interest Group (PCI-SIG). In one embodiment of the invention, the platform is compliant at least in part with the power states of the advanced configuration and power interface specification (ACPI standard, "Advanced Configuration and Power Interface Specification", Revision 4.0a, published April 5 2010). In another embodiment of the invention, the platform is compliant at least in part with earlier and/or future revisions of the ACPI standard.

In one embodiment of the invention, the platform has a processing core(s) coupled to a Root Complex. The Root Complex is coupled with a bus and a memory and supports three Root Ports. The Root Ports are coupled with the PCIe endpoint devices respectively via the PCIe communication links. In one embodiment of the invention, each of the Root Ports supports more than three Root Ports. In yet another embodiment of the invention, the Root Complex supports less than three Root Ports.

In one embodiment of the invention, the Root Complex generates transaction requests on behalf of the processor core(s). In one embodiment of the invention, the Root Complex is integrated with the processing core. In another embodiment of the invention, the Root Complex supports more than three Root Ports. In yet another embodiment of the invention, the Root Complex supports less than three Root Ports.

In one embodiment of the invention, the Root Ports and their respective PCIe endpoint devices support a low power idle link state that allows disabling of their high speed circuitry. Each of the Root Ports and their respective PCIe endpoint devices can initiate an entry or transition to the low power idle link state by sending an indication in one embodiment of the invention.

In one embodiment of the invention, a side band signal between the Root Port and its coupled PCIe endpoint device is used for the indication whether the platform is to transition to the low power idle link state. The out of band signal includes, but is not limited to, the system management bus (SMBus), Inter-IC (I2C) Bus or any other communication protocol.

Figure 1 is not meant to be limiting and serves only as one embodiment of the invention that is compliant at least in part with the PCIe standard. One of ordinary skill in the relevant art will readily appreciate how to adapt the workings of the invention to other high speed serial communication protocols and shall not be described herein.

Figure 2 illustrates a block diagram of a host device and an endpoint device in accordance with one embodiment of the invention. For clarity of illustration, the host device and the endpoint device are illustrated as being compliant at least in part with the PCIe standard. Figure 2 illustrates how the clock request (CLKREQ#) signal (not shown in figure 2) which is an unidirectional open drain signal from the host device to the endpoint device is repurposed as a bidirectional open drain signal clock request (CLKREQ#) signal in one embodiment of the invention. In one embodiment of the invention, both host device and the endpoint device use the bidirectional CLKREQ# signal to indicate that it wants to enter the low power idle state.

The host device has an upstream port that is communicatively coupled with the downstream port in the endpoint device in one embodiment of the invention. The upstream port has a link Power Management (PM) control logic, an Electrical Idle (EI) exit detect circuitry, a Reference Clock (CLK) source, a Transmitter (TX) circuit, and a Receiver (RX) circuit in one embodiment of the invention. The downstream port has a link Power Management (PM) control logic, an Electrical Idle (EI) exit detect circuitry, a Reference Clock (CLK) sink, a Transmitter circuit, and a Receiver (RX) circuit in one embodiment of the invention.

The Reference CLK source of the upstream port is coupled with the Reference CLK sink via the bidirectional CLKREQ# signal and the reference clock signal. The Alternating Current (AC) coupled data signaling illustrates the AC common mode differential signaling between the TX circuit and the RX circuit of the downstream port, and the AC common mode differential signaling between the TX circuit of the downstream port and the RX circuit of the upstream port. The high speed link circuitry of each of the upstream port and the downstream port includes, but is not limited to, the TX circuit, the RX circuit, the phase lock loop(s) (PLLS), and any other circuitry or logic required for the high speed communication between the upstream port and the downstream port.
port 220 and the downstream port 240 enter the low power idle link state. For example, in one embodiment of the invention, the PLL(s) in the reference CLK source 224 in the upstream port 220 is disabled in the low power idle link state. In one embodiment of the invention, the TX circuit 227 and RX circuit 228 in the upstream port 220 and the TX circuit 248 and RX circuit 247 in the downstream port 240 are disabled in the low power idle link state.

[0022] In another embodiment of the invention, the AC common mode differential signaling between the TX circuit 227 of the upstream port 220 and the RX circuit 247 of the downstream port 240 and the AC common mode differential signaling between the TX circuit 248 of the downstream port 240 and the RX circuit 228 of the upstream port 220 are disabled in the low power idle link state.

[0023] Different combinations of disabling the various components in the host device 210 and the endpoint device 230 in the low power idle link state can be used in one embodiment of the invention. In one embodiment of the invention, the low power idle link state is defined as a sub state of the Link Training Status State Machine (LTSSM). In one embodiment of the invention, two sub states, link 1 off (L1.OFF) link state, and link 1 snooze (L1.SNOOZ) link state are defined.

[0024] In the L1.OFF link state, the TX circuit, RX circuit, PLL(s) and the AC common mode keepers of the high speed link circuitry are fully power gated or disabled in one embodiment of the invention. In the L1.SNOOZ link state, the TX circuit, RX circuit, PLL(s) of the high speed link circuitry are fully power gated or disabled but the AC common mode keepers of the high speed link circuitry are enabled to maintain the AC common mode voltages of the differential signaling in one embodiment of the invention.

[0025] In figure 2, for clarity of illustration, it is assumed that it is possible for the RX and TX circuits and the PLL(s) associated with the port to be power gated while all the port context is maintained in one embodiment of the invention. If it is not possible for the RX and TX circuits and the PLL(s) associated with the port to maintain all the port context when they are power gated or disabled, one of ordinary skill in the relevant art will readily appreciate how to add additional logic to maintain all the port context when the RX and TX circuits and the PLL(s) associated with the port are disabled and it shall not be described herein.

[0026] Figure 3 illustrates a block diagram 300 of the link states in accordance with one embodiment of the invention. For clarity of illustration, figure 3 is discussed with reference to figure 2. The link state L1 330 illustrates one embodiment of the invention where it has two low power idle sub states: L1.OFF link state 380, and L1.SNOOZ link state 360 in one embodiment of the invention.

[0027] The L0 link state 310 is the normal operational state where the data and control packets can be transmitted and received. All power management states are entered from the L0 link state 310 in one embodiment of the invention. When a power management L1 link state request is received, the platform transitions to the L1 link state 330. The existing L1 link state 350 is the L1 state as defined in the LTSSM. In one embodiment of the invention, the platform transitions to the existing L1 link state 350 to maintain all the port context.

[0028] In one embodiment of the invention, the platform checks the link state policy in step 340 to determine whether it should remain in the existing L1 link state 350 or transition to L1.SNOOZ link state 360 or the L1.OFF link state 380. The link state policy is based on, but not limited to, the service latency requirement of the up/downstream port, and the idle duration information about the up/downstream port. In one embodiment of the invention, the link state policy of the host or upstream port dynamically determines which of the low power idle states to enter. The link state policy of the endpoint device or downstream port determines the guidelines and constraints to send to the host or upstream port in one embodiment of the invention.

[0029] For example, in one embodiment of the invention, both the Root Port and the endpoint device support Latency Tolerance Reporting (LTR). In one embodiment of the invention, the Root Port in the upstream port receives a LTR message from the endpoint device. In another embodiment of the invention, the endpoint device receives a LTR message from the Root Port in the upstream port.

[0030] The LTR message includes, but is not limited to, the service latency requirement of the up/downstream port. In one embodiment of the invention, the link state policy of the Root Port uses the last advertised service latency requirement of the downstream port. In another embodiment of the invention, the link state policy of the downstream port uses the last advertised service latency requirement of the Root Port.

[0031] In another embodiment of the invention, the estimated or projected idle duration is used by the link state policy to determine which low power idle link state to enter. For example, in one embodiment of the invention, when the link state policy determines that the idle duration is short, it sets or keeps the platform at the existing L1 link state 350. When the link state policy determines that the idle duration is long, it transitions the platform to the L1.OFF link state 380 or L1.SNOOZ link state 360 based on the service latency requirement in one embodiment of the invention.

[0032] In one embodiment of the invention, the platform has the ability to detect and enable the low power idle link states in the upstream port and the endpoint device. For example, in one embodiment of the invention, both the upstream port and the endpoint device have configuration registers to indicate whether it supports the low power idle link states. In another embodiment of the invention, the upstream port and the endpoint device use boot strapping to indicate whether the low power idle link
In one embodiment of the invention, in step 340, the platform checks if the service latency requirement of the downstream port received from the LTR is greater than a threshold LTR_L1OFF_THRESHOLD and if the L1.OFF link state 380 is enabled. In another embodiment of the invention, in step 340, the platform checks if the service latency requirement of the upstream port received from the LTR is greater than a threshold LTR_L1OFF_THRESHOLD and if the L1.OFF link state 380 is enabled. In one embodiment of the invention, the LTR_L1OFF_THRESHOLD is stored in a register in the root port and the endpoint device.

If both the service latency requirement of the up/downstream port received from the LTR is greater than the threshold LTR_L1OFF_THRESHOLD and the L1.OFF link state 380 is enabled, the platform transitions to the L1.OFF link state 380. The platform transitions to the L1.OFF link state 380 to achieve the greatest power savings by disabling all the high speed link circuitry in one embodiment of the invention.

If either the service latency requirement of the up/downstream port received from the LTR is less than the threshold LTR_L1OFF_THRESHOLD or the L1.OFF link state 380 is not enabled, the platform transitions to the L1.SNOOZ link state 360.

In one embodiment of the invention, the L1.SNOOZ link state 360 offers an intermediate power benefit as the TX circuit, RX circuit, PLL(s) of the high speed link circuitry are fully powered or disabled but the AC common mode keepers of the high speed link circuitry are enabled to maintain the AC common mode voltages of the differential signaling. By maintaining the AC common mode voltages of the differential signaling, the L1.SNOOZ link state 360 needs shorter exit latency compared to the L1.OFF link state 380. The L1.OFF link state 360 needs longer exit latency as it is required to reestablish the AC common voltages of the differential signaling in one embodiment of the invention.

The L1.SNOOZ link state 360 transitions to the recovery link state 320 via the existing L1 link state 350 in one embodiment of the invention. The L1.OFF link state 380 transitions to the recovery link state 320 via the existing L1 link state 350 in another embodiment of the invention. The Recovery link state 320 allows both the Transmitter and Receiver to send and receive data in one embodiment of the invention. The Recovery link state 320 also allows a configured link to change the speed data rate of operation if desired, re-establish bit lock, symbol lock or block alignment, and Lane-to-Lane de-skew.

Figure 4 illustrates a timing diagram or flow 400 of an upstream port initiated link state exit in accordance with one embodiment of the invention. For clarity of illustration, figure 4 is discussed with reference to figure 3. Figure 4 illustrates the transition from the L0 link state 310, the existing L1 link state 350, the L1.OFF link state 380, the existing L1 link state 350, the recovery link state 320 and back to the L0 link state 310. In figure 4, for clarity of illustration, it is assumed that both the service latency requirement of the downstream port received from the LTR is greater than the threshold LTR_L1OFF_THRESHOLD and the L1.OFF link state is enabled.

Figure 4 illustrates the upstream port state 410, the downstream port state 420, the upstream port CLKREQ# signal 430, the downstream port CLKREQ# signal 440, the bidirectional CLKREQ# signal 450, and the reference clock (REFCLK) 460. The CLKREQ# signal 450 is a bidirectional open drain signal that is coupled between the upstream port and the downstream port in one embodiment of the invention.

The upstream port CLKREQ# signal 430 shows the signal driven the upstream port on the bidirectional CLKREQ# signal 450. The downstream port CLKREQ# signal 440 shows the signal driven the downstream port on the CLKREQ# signal 450. In one embodiment of the invention, the bidirectional CLKREQ# signal 450 is coupled with a pull-up resistor that pulls the bidirectional CLKREQ# signal 450 to a logic one voltage when the upstream port and the downstream port are not driving the bidirectional CLKREQ# signal 450 low.

At the start of the flow 400, the upstream port state 410 and the downstream port state 420 are in the L0 link state 411 and 421 respectively. The upstream port sends a Power Management L1 state transition request (PM_L1_REQ) 412 to the downstream port and transitions to the existing L1 link state 413 in preparation for entering the L1.OFF link state 414. In one embodiment of the invention, the existing L1 link state 413 maintains all port context. The downstream port receives the PM_L1_REQ 412 and sends a PM acknowledgement (PM_ACK) 422 to the upstream port. The downstream port transitions to the L1 link state 423 in one embodiment of the invention.

The downstream port does not drive the downstream port CLKREQ# signal 440, i.e., the signal is left floating or at a tristate level. The upstream port drives or keeps the upstream port CLKREQ# signal 430 low as illustrated by the signal 431. When the upstream port is in the existing L1 link state 413, the upstream port initiates the transition to the L1.OFF link state 414 by de-asserting the upstream port CLKREQ# signal 430 as illustrated by the tristate signal 432.

The upstream and downstream ports sample the CLKREQ# signal 450 and when it determines that the CLKREQ# signal 450 is de-asserted or inactive, the upstream and downstream ports enter the L1.OFF link state 414 and 424 respectively. After waiting for a minimum Tpoweroff 470 time, the upstream and downstream ports may power gate their Physical layer (PHY) circuitry
in one embodiment of the invention. When the downstream port samples that the bidirectional CLKREQ# signal 450 is inactive, it turns off the link REFCLK 460 within the time Trefclk_off 480. The REFCLK 460 is valid 461 until it is switched off in one embodiment of the invention. If either the upstream or downstream port needs to exit the L1.OFF state, it asserts its respective CLKREQ# signal 430 and 440. In one embodiment of the invention, if the downstream port is initiating an exit from the L1.OFF link state 424, it drives the CLKREQ# signal 440 low until the link enters the L0 link state 427. When the upstream port is initiating an exit from the L1.OFF link state 414, it drives the CLKREQ# signal 430 until the link enters the recovery link state 416.

[0045] The upstream port initiates the exit from the L1.OFF link state 414 when it drives the upstream port CLKREQ# signal 430 low as illustrated by the signal 433. Both the upstream and downstream ports monitor CLKREQ# signal 450 when they are in the L1.OFF link state 414 and 424. When the upstream and downstream ports determine that the bidirectional CLKREQ# signal 450 has been asserted, they perform an exit from their L1.OFF link state and enter the L1 link state 415 and 425 respectively.

[0046] In one embodiment of the invention, the upstream and downstream ports wait for a minimum time of Tpowerup 472 before powering up their PHYs and actively driving their link interface. When the downstream port samples that the bidirectional CLKREQ# signal 450 is de-asserted, the REFCLK 460 remains gated or disabled for the minimum time of Tpowerup 472. The Trefclk_on 482 illustrates the amount of time before the REFCLK 460 becomes valid 462.

[0047] In one embodiment of the invention, all the link and PHY states are maintained during the L1.OFF link state 414 and 424 and are restored upon exit from the L1.OFF link state 414 and 424. When the REFCLK 460 becomes valid 462 and their PLL(s) are locked, the upstream and downstream ports enter the recovery link state 416 and 426 respectively.

[0048] In one embodiment of the invention, the electrical state of the link during an exit from the existing L1 link state 415 and 425 may not match the PCIe standard L1 exit specification because the TX common mode was not maintained. The TX common mode is re-established by the time of exiting the recovery training sequence. In order to ensure that the TX common mode has been established prior to entering the the L0 link state 427, the downstream port LTSSM maintains a timer and does not send any TS2 Training Sequences until a minimum time of Tcommonmode 474 has elapsed since the downstream port started both transmitting and receiving TS1 Training Sequences.

[0049] Figure 5 illustrates a timing diagram or flow 500 of a downstream port initiated link state exit in accordance with one embodiment of the invention. Figure 5 illustrates a scenario where the upstream port initiates the entry to the L1.OFF link state 514 and the downstream port initiates the exit from the L1.OFF link state 524 when it drives or asserts the downstream port CLKREQ# signal 540 low as illustrated by the signal 533. The workings of the link state transitions in figure 5 are similar to the workings of the link state transitions in figure 4 and shall not be described herein.

[0050] Figure 6 illustrates a timing diagram or flow 600 of an upstream port initiated link state exit in accordance with one embodiment of the invention. For clarity of illustration, figure 6 is discussed with reference to figure 3. Figure 6 illustrates the transition from the L0 link state 310, the existing L1 link state 350, the L1.SNOOZ link state 360, the existing L1 link state 350, the recovery link state 320 and back to the L0 link state 310.

[0051] The flow 600 is applicable when the service latency requirement of the upstream port is less than the LTR_L1OFF_THRESHOLD in one embodiment of the invention. For example, in one embodiment of the invention, the flow 600 is applicable when Max Snooped or Non-Snooped LTR value advertised by the upstream port is less than the LTR_L1OFF_THRESHOLD.

[0052] In one embodiment of the invention, the flow 600 is applicable when the L1.OFF is not enabled. In another embodiment of the invention, the platform has a L1.SNOOZ enable function and the flow 600 is applicable when the L1.SNOOZ is not enabled. The enabling feature of L1.SNOOZ link state is implemented using, but not limited to, register bits, non-volatile storage medium and the like.

[0053] Figure 6 illustrates the upstream port state 610, the downstream port state 620, the upstream port CLKREQ# signal 630, the downstream port CLKREQ# signal 640, the Up/Downstream port EI Exit Detect signal 650, the bidirectional CLKREQ# signal 660, and the REFCLK 670.

[0054] At the start of the flow 600, the upstream port state 610 and the downstream port state 620 are in the L0 link state 611 and 621 respectively. The upstream port sends a PM_L1_REQ 612 to the downstream port and transitions to the existing L1 link state 613. The downstream port receives the PM_L1_REQ 612 and sends a PM_ACK 622 to the upstream port. The downstream port transitions to the existing L1 link state 623.

[0055] The downstream port does not drive the downstream port CLKREQ# signal 640, i.e., the signal is left floating or at a tristate level. The upstream port drives or keeps the upstream port CLKREQ# signal 630 low as illustrated by the signal 631. In the existing L1 link state 613 and 623, the link electrical state is maintained based on the PCIe specification. When the upstream port is in the existing L1 link state 613, the upstream port initiates the transition to the L1.SNOOZ link state 614 by de-asserting the upstream port CLKREQ# signal 630 as illustrated by the tristate signal 632.

[0056] When both the output driver of the upstream port CLKREQ# signal 630 and the downstream port CLKREQ# signal 640 are switched off, the pull up resistor on the bidirectional CLKREQ# signal 660 pulls the volt-
age on the bidirectional CLKREQ# signal 640 to logic one voltage.

[0057] The upstream and downstream ports sample the CLKREQ# signal 660 and when it determines that the CLKREQ# signal 660 is de-asserted or inactive, the upstream and downstream ports enter the L1.SNOOZ link state 614 and 624 respectively. Both upstream and downstream ports turn off their EI exit detect circuitry when they enter the L1.SNOOZ link state 614 and 624 respectively as illustrated as the signal 652 in one embodiment of the invention. In one embodiment of the invention, both upstream and downstream ports power down any active logic that is not required to maintain TX link state transitions in figure 7 are similar to the workings of the L1.SNOOZ link state 614 and 624 when they are in the L1.SNOOZ link state 614 and 624 respectively. Both upstream and downstream ports determine that the bidirectional CLKREQ# signal 660 has been asserted, they perform an exit from the L1.SNOOZ link state 624, it drives the downstream port CLKREQ# signal 640 low until the link enters the L0 link state 627. If the upstream port is initiating an exit from the L1.SNOOZ link state 614, it drives the upstream port CLKREQ# signal 630 low until the link has passed through the L0 link state 617 and re-entered the next existing L1 link state.

[0058] If either the upstream or downstream port needs to exit the L1.SNOOZ state, it asserts their respective CLKREQ# signal 630 and 640. In one embodiment of the invention, if the downstream port is initiating an exit from the L1.SNOOZ link state 624, it drives the downstream port CLKREQ# signal 640 low until the link enters the L0 link state 627. If the upstream port is initiating an exit from the L1.SNOOZ link state 614, it drives the upstream port CLKREQ# signal 630 low until the link has passed through the L0 link state 617 and re-entered the next existing L1 link state.

[0059] The upstream port initiates the exit from the L1.SNOOZ link state 614 when it drives the upstream port CLKREQ# signal 630 low as illustrated by the signal 633. Both the upstream and downstream ports monitor CLKREQ# signal 660 when they are in the L1.SNOOZ link state 614 and 624. When the upstream and downstream ports determine that the bidirectional CLKREQ# signal 660 has been asserted, they perform an exit from the L1.SNOOZ link state and enter the existing L1 link state 615 and 625 respectively. The upstream and downstream ports enable the EI exit detect circuitry as illustrated as the signal 653 in one embodiment of the invention. The upstream port turns on its REFCLK activity detection circuit as illustrated by the valid signal 672 in one embodiment of the invention.

[0060] Figure 7 illustrates a timing diagram or flow 700 of a downstream port initiated link state exit in accordance with one embodiment of the invention. Figure 7 illustrates a scenario where the upstream port initiates the entry to the L1.SNOOZ link state and the downstream port initiates the exit from the L1.SNOOZ link state when it drives or asserts the downstream port CLKREQ# signal 740 low as illustrated by the signal 733. The workings of the link state transitions in figure 7 are similar to the workings of the link state transitions in figure 6 and shall not be described herein.

[0061] Figure 8 illustrates a system or platform 800 to implement the methods disclosed herein in accordance with one embodiment of the invention. The system 800 includes, but is not limited to, a desktop computer, a laptop computer, a netbook computer, a handheld computer, a tablet computer, a notebook computer, a personal digital assistant (PDA), a server, a workstation, a cellular telephone, a smartphone, a mobile computing device, an Internet appliance or any other type of computing device. In another embodiment, the system 800 used to implement the methods disclosed herein may be a system on a chip (SOC) system.

[0062] The processor 810 has a processing core 812 to execute instructions of the system 800. The processing core 812 includes, but is not limited to, pre-fetch logic to fetch instructions, decode logic to decode the instructions, execution logic to execute instructions and the like. The processor 810 has a cache memory 816 to cache instructions and/or data of the system 800. In another embodiment of the invention, the cache memory 816 includes, but is not limited to, level one, level two and level three, cache memory or any other configuration of the cache memory within the processor 810.

[0063] The memory control hub (MCH) 814 performs functions that enable the processor 810 to access and communicate with a memory 830 that includes a volatile memory 832 and/or a non-volatile memory 834. The volatile memory 832 includes, but is not limited to, Synchronous Dynamic Random Access Memory (SDRAM), Dynamic Random Access Memory (DRAM), Rambus Dynamic Random Access Memory (RDRAM), and/or any other type of random access memory device. The non-volatile memory 834 includes, but is not limited to, NAND flash memory, phase change memory (PCM), read only memory (ROM), electrically erasable programmable read only memory (EEPROM), or any other type of non-volatile memory device.

[0064] The memory 830 stores information and instructions to be executed by the processor 810. The memory 830 may also stores temporary variables or other intermediate information while the processor 810 is executing instructions. The chipset 820 connects with the processor 810 via Point-to-Point (PtP) interfaces 817 and 822. In another embodiment of the invention, the chipset 820 is a platform control hub. The I/O subsystem optionally includes the platform control hub in one embodiment of the invention.

[0065] The chipset 820 enables the processor 810 to connect to other modules in the system 800. In one embodiment of the invention, the interfaces 817 and 822 operate in accordance with a PtP communication protocol such as the Intel® QuickPath Interconnect (QPI) or the like. The chipset 820 connects to a display device 840 that includes, but is not limited to, liquid crystal display (LCD), cathode ray tube (CRT) display, or any other form of visual display device.

[0066] In addition, the chipset 820 connects to one or more buses 850 and 860 that interconnect the various modules 874, 880, 882, 884, and 886. Buses 850 and 860 may be interconnected together via a bus bridge 872 if there is a mismatch in bus speed or communication protocol. The chipset 820 couples with, but is not limited
to, a non-volatile memory 880, a mass storage device(s) 882, a keyboard/mouse 884 and a network interface 886. The mass storage device 882 includes, but is not limited to, a solid state drive, a hard disk drive, an universal serial bus flash memory drive, or any other form of computer data storage medium. The network interface 886 is implemented using any type of well-known network interface standard including, but not limited to, an Ethernet interface, a universal serial bus (USB) interface, a Peripheral Component Interconnect (PCI) Express interface, a wireless interface and/or any other suitable type of interface. The wireless interface operates in accordance with, but is not limited to, the IEEE 802.11 standard and its related family, Home Plug AV (HPAV), Ultra Wide Band (UWB), Bluetooth, WiMax, or any form of wireless communication protocol.

[0067] While the modules shown in Figure 8 are depicted as separate blocks within the system 800, the functions performed by some of these blocks may be integrated within a single semiconductor circuit or may be implemented using two or more separate integrated circuits. For example, although the cache memory 816 is depicted as a separate block within the processor 810, the cache memory 816 can be incorporated into the processor core 812 respectively. The system 800 may include more than one processor / processing core in another embodiment of the invention.

[0068] The methods disclosed herein can be implemented in hardware, software, firmware, or any other combination thereof. Although examples of the embodiments of the disclosed subject matter are described, one of ordinary skill in the relevant art will readily appreciate that many other methods of implementing the disclosed subject matter may alternatively be used. In the preceding description, various aspects of the disclosed subject matter have been described. For purposes of explanation, specific numbers, systems and configurations were set forth in order to provide a thorough understanding of the subject matter. However, it is apparent to one skilled in the relevant art having the benefit of this disclosure that the subject matter may be practiced without the specific details. In other instances, well-known features, components, or modules were omitted, simplified, combined, or split in order not to obscure the disclosed subject matter.

[0069] The term "is operable" used herein means that the device, system, protocol etc, is able to operate or is adapted to operate for its desired functionality when the device or system is in off-powered state. Various embodiments of the disclosed subject matter may be implemented in hardware, firmware, software, or combination thereof, and may be described by reference to or in conjunction with program code, such as instructions, functions, procedures, data structures, logic, application programs, design representations or formats for simulation, emulation, and fabrication of a design, which when accessed by a machine results in the machine performing tasks, defining abstract data types or low-level hardware contexts, or producing a result.

[0070] The techniques shown in the figures can be implemented using code and data stored and executed on one or more computing devices such as general purpose computers or computing devices. Such computing devices store and communicate (internally and with other computing devices over a network) code and data using machine-readable media, such as machine readable storage media (e.g., magnetic disks; optical disks; random access memory; read only memory; flash memory devices; phase-change memory) and machine readable communication media (e.g., electrical, optical, acoustical or other form of propagated signals - such as carrier waves, infrared signals, digital signals, etc.).

[0071] While the disclosed subject matter has been described with reference to illustrative embodiments, this description is not intended to be construed in a limiting sense. Various modifications of the illustrative embodiments, as well as other embodiments of the subject matter, which are apparent to persons skilled in the art to which the disclosed subject matter pertains are deemed to lie within the scope of the disclosed subject matter.

Claims

1. An apparatus (105) compliant at least in part with a Peripheral Component Interface Express 'PCIe' standard comprising logic that is part of an upstream port (220), the logic configured to:

   determine whether the apparatus is in a first link state (310);
   determine whether an indication that the apparatus is to transition to a second link state (360), the second link state to consume less power than the first link state, has been received in response to the determination that the apparatus is in the first link state; and
   transition from the first link state to the second link state in response to the determination that the indication has been received, wherein:
   the logic is further configured to determine whether a service latency requirement of a downstream port (240) coupled with the upstream port is less than a threshold; and
   transition from the first link state to the second link state in response to the determination that the indication has been received and the determination that the service latency requirement of the downstream port coupled with the upstream port is less than the threshold; characterized in that the logic is further configured to:
   disable a receiver circuitry, a transmitter circuitry, and a phase locked loop 'PLL', and to enable an Alternating Current 'AC' common mode circuitry of the upstream port.

5
10
15
20
25
30
35
40
45
50
55

8
2. The apparatus of claim 1, wherein the indication comprises a signal coupled with the upstream port and a downstream port, and wherein the logic to determine whether the indication that the apparatus is to transition to the second link state has been received is to determine whether a setting of the signal has been asserted.

3. A method (340) for reducing a link power consumption comprising:

- determining whether an apparatus (105) compliant at least in part with a Peripheral Component Interface Express (PCIe) standard is in a first link state (310);
- determining whether an indication that the apparatus is to transition to a second link state (360), the second link state to consume less power than the first link state, has been received in response to the determination that the apparatus is in the first link state; and
- transitioning from the first link state to the second link state in response to the determination that the indication has been received, wherein:
  - the method is to be performed by an upstream port (220);
  - the method further comprises determining by the upstream port, whether a service latency requirement of a downstream port (240) coupled with the upstream port is less than a threshold; and
  - transitioning from the first link state to the second link state in response to the determination that the indication has been received and the determination that the service latency requirement of the downstream port coupled with the upstream port is less than the threshold;

characterized in that the method further comprises:

- disabling a receiver circuitry, a transmitter circuitry, and a phase locked loop (PLL) and
- enabling an Alternating Current (AC) common mode circuitry of the upstream port.

4. The method of claim 3, wherein determining whether the indication that the apparatus is to transition to the second link state has been received in response to the determination that the apparatus is in the first link state comprises determine whether a setting of a signal has been asserted.

5. The method of claim 3, further comprising: receiving a Latency Tolerance Reporting (LTR) message from the downstream port, wherein the service latency requirement of the downstream port is part of the LTR message.

Patentansprüche

1. Vorrichtung (105), die zumindest teilweise einem Peripheral Component Interface Express (PCIe)-Standard entspricht, die Logik umfasst, die Teil eines Upstream-Ports (220) ist, wobei die Logik konfiguriert ist, um:

- zu bestimmen, ob sich die Vorrichtung in einem ersten Verbindungszustand (310) befindet; zu bestimmen, ob eine Angabe, dass die Vorrichtung in einen zweiten Verbindungszustand (360) übergehen soll, wobei der zweite Verbindungszustand weniger Leistung verbraucht als der erste Verbindungszustand, in Reaktion auf das Bestimmen empfangen wurde, dass sich die Vorrichtung in dem ersten Verbindungszustand befindet; und
- von dem ersten Verbindungszustand in den zweiten Verbindungszustand in Reaktion auf das Bestimmen überzugehen, dass die Angabe empfangen wurde, wobei:
  - die Logik ferner konfiguriert ist, um zu bestimmen, ob eine Dienstlizenzanforderung eines Downstream-Ports (240) deaktivieren, kleiner als ein Schwellenwert ist; und
dadurch gekennzeichnet, dass die Logik ferner konfiguriert ist, um:
  - eine Empfängerschaltung, eine Senderschaltung und einen Phasenregelkreis (Phase Locked Loop, PLL) zu deaktivieren und eine Wechselstrom (Alternating Current, AC)-Gleichtaktschaltung des Upstream-Ports zu aktivieren.

2. Vorrichtung nach Anspruch 1, wobei die Angabe ein Signal umfasst, das mit dem Upstream-Port gekoppelt ist, wobei die Logik zum Bestimmen, ob die Angabe, dass die Vorrichtung in den zweiten Verbindungszustand übergehen soll, empfangen wurde, bestimmen soll, ob
eine Einstellung des Signals vorgenommen wurde.

3. Verfahren (340) zum Reduzieren eines Verbindungsleistungsverbrauchs, umfassend:

Bestimmen, ob eine Vorrichtung (105), die zumindest teilweise einem Peripheral Component Interface Express (PCIe)-Standard entspricht, sich in einem ersten Verbindungszustand (310) befindet;
Bestimmen, ob eine Angabe, dass die Vorrichtung in einen zweiten Verbindungszustand (360) übergehen soll, wobei der zweite Verbindungszustand weniger Leistung verbraucht als der erste Verbindungszustand, in Reaktion auf das Bestimmen empfangen wurde, dass sich die Vorrichtung in dem ersten Verbindungszustand befindet; und
Übergehen von dem ersten Verbindungszustand in den zweiten Verbindungszustand in Reaktion auf das Bestimmen, dass die Angabe empfangen wurde, wobei:

- das Verfahren von einem Upstream-Port (220) ausgeführt werden soll; das Verfahren ferner das Bestimmen durch den Upstream-Port umfasst, ob eine Dienstlatenzanforderung eines Downstream-Ports (240), der mit dem Upstream-Port gekoppelt ist, kleiner als ein Schwellenwert ist; und
- Übergehen von dem ersten Verbindungszustand in den zweiten Verbindungszustand in Reaktions auf das Bestimmen, dass die Angabe empfangen wurde, das Verfahren von einem Upstream-Port (220) ausgeführt werden soll; das Verfahren ferner das Bestimmen durch den Upstream-Port umfasst, ob eine Dienstlatenzanforderung eines Downstream-Ports (240), der mit dem Upstream-Port gekoppelt ist, kleiner als ein Schwellenwert ist; und

dadurch gekennzeichnet, dass das Verfahren von einem Upstream-Port (220) ausgeführt werden soll; das Verfahren ferner das Bestimmen durch den Upstream-Port umfasst, ob eine Dienstlatenzanforderung eines Downstream-Ports (240), der mit dem Upstream-Port gekoppelt ist, kleiner als ein Schwellenwert ist; und

4. Verfahren nach Anspruch 3, wobei das Bestimmen, ob die Angabe, dass die Vorrichtung in den zweiten Verbindungszustand übergehen soll, in Reaktion auf das Bestimmen empfangen wurde, dass sich die Vorrichtung in dem ersten Verbindungszustand befindet, das Bestimmen umfasst, ob eine Einstellung eines Signals vorgenommen wurde.

5. Verfahren nach Anspruch 3, ferner umfassend:
Empfangen einer Latenz-Toleranzberichts (Latency Tolerance Reporting, LTR)-Nachricht von dem Downstream-Port, wobei die Dienstlatenzanforderung des Downstream-Ports Teil der LTR-Nachricht ist.

Revendications

1. Appareil (105) conforme, au moins en partie, à une norme d’Interface de Composants Périphériques Express ( «PCIe» ) comprenant une logique qui fait partie d’un pot amont (220), la logique étant configurée :

   - pour déterminer si l’appareil se trouve dans un premier état de liaison (310) ;
   - pour déterminer si une indication que l’appareil doit passer à un second état de liaison (360), le second état de liaison étant destiné à consommer moins d’énergie que le premier état de liaison, a été reçue à la suite de la détermination que l’appareil se trouve dans le premier état de liaison ; et
   - pour passer du premier état de liaison au second état de liaison à la suite de la détermination que l’indications a été reçue, dans lequel :
     - la logique est en outre configurée pour déterminer si une exigence de latence de service d’un port aval (240) couplé au port amont est inférieure à un seuil ; et
     - pour passer du premier état de liaison au second état de liaison à la suite de la détermination que l’indications a été reçue, et de la détermination que l’exigence de latence de service du port aval couplé au port amont est inférieure au seuil ; caractérisé en ce que la logique est en outre configurée :
     - pour désactiver un ensemble de circuits de récepteur, un ensemble de circuits d’émetteur et une boucle à verrouillage de phase ( «PLL ») et pour activer un ensemble de circuits de mode commun en courant alternatif ( «AC») du port amont.

2. Appareil selon la revendication 1, dans lequel l’indications comprend un signal couplé au port amont et à un port aval, et dans lequel la logique pour déterminer si l’indications que l’appareil doit passer au second état de liaison, a été reçue, doit déterminer si un régulation du signal a été activé.

3. Procédé (340) pour réduire une consommation d’énergie de liaison consistant :

   - à déterminer si un appareil (105) conforme, au moins en partie, à une norme d’Interface de
Composants Périphériques Express (« PCIe ») se trouve dans un premier état de liaison (310) ; à déterminer si une indication que l’appareil doit passer à un second état de liaison (360), le second état de liaison étant destiné à consommer moins d’énergie que le premier état de liaison, a été reçue à la suite de la détermination que l’appareil se trouve dans le premier état de liaison ; et à passer du premier état de liaison au second état de liaison à la suite de la détermination que l’indication a été reçue, dans lequel :

le procédé doit être réalisé par un port amont (220) ;

le procédé consiste en outre à déterminer, au moyen du port amont, si une exigence de latence de service d’un port aval (240) couplé au port amont est inférieure à un seuil ; et à passer du premier état de liaison au second état de liaison à la suite de la détermination que l’indication a été reçue, et de la détermination que l’exigence de latence de service du port aval couplé au port amont est inférieure au seuil ; caractérisé en ce que le procédé consiste en outre :

à désactiver un ensemble de circuits de récepteur, un ensemble de circuits d’émetteur et une boucle à verrouillage de phase (« PLL ») et à activer un ensemble de circuits de mode commun en courant alternatif (« AC ») du port amont.

4. Procédé selon la revendication 3, dans lequel la détermination que l’indication que l’appareil doit passer au second état de liaison, a été reçue à la suite de la détermination que l’appareil se trouve dans le premier état de liaison, consiste à déterminer si un réglage d’un signal a été activé.

5. Procédé selon la revendication 3, consistant en outre :

à recevoir un message de rapport de tolérance de latence (« LTR ») en provenance du port aval, dans lequel l’exigence de latence de service du port aval fait partie du message de rapport LTR.
1701
Upstream port state

710 PM_L1_REQ 714 L1.SNOOZ 715 L1 716 Recovery 717 L0

720 Downstream port state

721 L0 722 PM_ACK 723 L1 724 L1.SNOOZ 725 L1 726 Recovery 727 L0

730 Upstream port CLKREQ# 731 Drive Low 732 TriState 733 Drive Low

740 Downstream port CLKREQ# 741 TriState 733 Drive Low 734 TriState

750 Up/Downstream port El Exit Detect 751 Enabled 752 Disabled 753 Enabled

760 Bidirectional CLKREQ# 770 REFCLK 771 VALID 772 VALID

FIG. 7
REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader’s convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Non-patent literature cited in the description

- Advanced Configuration and Power Interface Specification. ACPI standard, 05 April 2010 [0010]