(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
4 July 2013 (04.07.20 13)

W I P O | P C T

(10) International Publication Number
WO 2013/101180 Al

(51) International Patent Classification:
G06F 1/32 (2006.01)

(21) International Application Number:
PCT/US201 1/068096

(22) International Filing Date:
30 December 201 1 (30. 12.201 1)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant *(for all designated States except US):* **INTEL CORPORATION** [US/US]; 2200 Mission College Boulevard, Santa Clara, CA 95052 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only):* **KUMAR, Anil** [US/US]; 63 12 W Linda Lane, Chandler, AZ 85226 (US). **CRAWFORD, John** [US/US]; 20128 Chateau Drive, Saratoga, CA 95070 (US). **DIEFENBAUGH, Paul** [US/US]; 9016 NW Bartholomew Dr., Portland, OR 97229 (US).

(74) Agent: **BUCKLEY, Patrick;** Buckley, Maschoff & Talwalkar LLC, 50 Locust Avenue, New Canaan, CT 06840 (US).

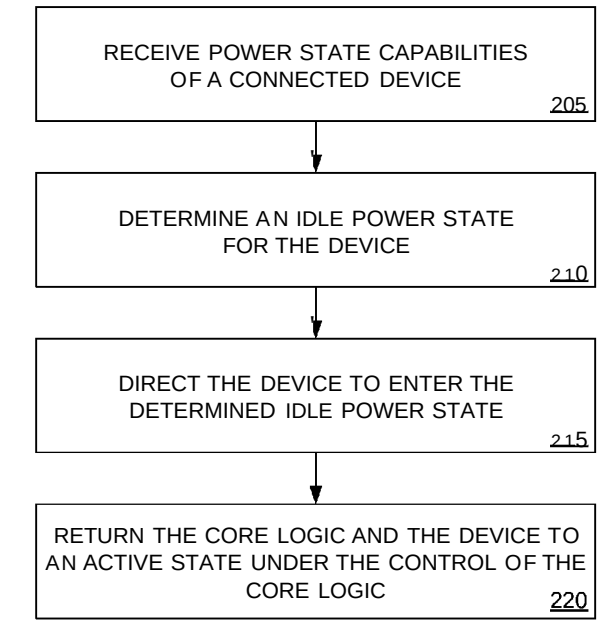
(81) Designated States *(unless otherwise indicated, for every kind of national protection available):* AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available):* ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

(54) **Title:** PCIE DEVICE POWER STATE CONTROL



(57) **Abstract:** An apparatus, system, and method, the method including receiving an indication of a idle state capability of a platform connected device; determining, by a chipset, an idle power state compatible with the device; and directing the device to enter the determined idle power state based on a power state of the chipset.

WO 2013/101180 A1

PCIE DEVICE POWER STATE CONTROL

BACKGROUND OF THE INVENTION

Power management for computing platforms is a subject of great concern. It is a major focus in the design and operation of computing platforms, systems and sub-systems. Ideally, devices and components should not operate at a higher level of activity than absolutely required. However, system and devices may not, for example, be fully aware of the present and/or future operational states of the components to which they are connected. In some aspects, software embodied techniques have been proposed to manage the power states of some computing platforms. A disadvantage in some such systems is that the software must be executing in order to manage the power. Thus, the running of the software may itself prevent the system from entering into a lower idle power state.

BRIEF DESCRIPTION OF THE DRAWINGS

Aspects of the present disclosure herein are illustrated by way of example and not by way of limitation in the accompanying figures. For purposes related to simplicity and clarity of illustration rather than limitation, aspects illustrated in the figures are not necessarily drawn to scale. Further, where considered appropriate, reference labels have been repeated among the figures to indicate corresponding or analogous elements.

FIG. 1 is a block diagram of a system compatible with some embodiments herein.

FIG. **2** is a flow diagram of a process, in accordance with an embodiment.

FIG. **3** is a flow diagram of a sequence of events, in accordance with one embodiment.

FIG. 4 is an illustrative depiction of a power state table, in accordance with one embodiment.

DETAILED DESCRIPTION

The disclosure herein provides numerous specific details such regarding a system for implementing various processes and operations. However, it will be appreciated by one skilled in the art(s) related hereto that embodiments of the present disclosure may be practiced without such specific details. Thus, in some instances aspects such as control mechanisms and full software instruction sequences have not been shown in detail in order not to obscure other aspects of the present disclosure. Those of ordinary skill in the art will be able to implement appropriate functionality without undue experimentation given the included descriptions herein.

References in the specification to "one embodiment", "some embodiments", "an embodiment", "an example embodiment", "an instance", "some instances" indicate that the embodiment described may include a particular feature, structure, or characteristic, but that every embodiment may not necessarily include the particular feature, structure, or characteristic. Moreover, such phrases are not necessarily referring to the same embodiment. Further, when a particular feature, structure, or characteristic is described in connection with an embodiment, it is submitted that it is within the knowledge of one skilled in the art to affect such feature, structure, or characteristic in connection with other embodiments whether or not explicitly described.

Some embodiments herein may be implemented in hardware, firmware, software, or any combinations thereof. Embodiments may also be implemented as executable instructions stored on a machine-readable medium that may be read and executed by one or more processors. A machine-readable storage medium may include any tangible non-transitory mechanism for storing information in a form readable by a machine (e.g., a computing device). In some aspects, a machine-readable storage medium may include read only memory (ROM); random access memory (RAM); magnetic disk storage media; optical storage media; flash memory devices; and electrical and optical forms of signals. While firmware, software, routines, and instructions may be described herein as performing certain actions, it should be appreciated that such descriptions are merely for convenience and that such actions are in fact result from computing devices, processors, controllers, and other devices executing the firmware, software, routines, and instructions.

The present disclosure relates, in some aspects, to a control scheme or algorithm for dynamically controlling a power state of a device connected to a platform based on the power state of the platform. In some embodiments, the device may be a PCIE (Peripheral Component Interconnect Express) device connected to an end point of a PCIE bus, under the control of a core logic chipset (also referred to herein as "core logic" and "chipset"). Some embodiments herein provide a mechanism to manage the power of the PCIE device based on the power state of the platform, under the control of the chipset and within the operational constraints of the PCIE device, the chipset, and other devices, (e.g., internal and external to the chipset). FIG. 1 is a simplified block diagram of a chipset 105 interconnected with a PCIE device 110. In some embodiments, device 110 may be any endpoint connected device compatible with the bus 115 protocol. Device 110 may be one or more devices. In some embodiments, concepts and aspects of the present disclosure may be applicable to platform other than those particularly referenced, shown, or discussed herein.

Process 200 is a flow diagram of a process 200 related to one embodiment of a platform power management method, algorithm, system, and apparatus. At operation 205, a chipset (e.g., chipset 105) may receive information including the power state capabilities of a platform connected device (e.g., PCIE device 110). The power state information may be received as the result of the device advertising its power state capabilities, whether in response to a solicitation or request for such information from the chipset, the chipset requesting such information, as part of a "boot" process, etc. The power state information may be accumulated, stored, or persisted for future reference by the chipset.

At operation 210, the chipset determines an idle power state for the device. The determined idle power state may consider and be based, at least in part, on the power state information gathered by the chipset at operation 205. In some embodiments, the idle power state determined for the device at operation 210 may consider additional factors other than the device's power capabilities. In some instances, the chipset may consider, for example, operational parameters of the chipset itself as well desired performance goals.

At operation 215 the chipset may direct the connected device to enter the determined power state. In some embodiments, the device may receive an indication of the determined idle power state and enter such an idle power state at a next available opportunity. In some embodiments, the information provided to the device regarding the determined idle power state may include an indication of the power state of the chipset. Based on the power state of the chipset as reported from the chipset, the connected device may make its own state transitions with confidence knowing the state of the chipset (i.e., platform). In some embodiments, the device may enter the determined idle power state at a next idle period or at all future idle periods unless directed otherwise.

At operation 220, the core logic and the device may be returned to an active state. Operation 220 may also be managed under the control of the chipset.

In some aspects, it is noted that the control scheme and methods disclosed herein may be strictly controlled by hardware.

FIG. 3 is a flow diagram 300 of a sequence of operations occurring between a chipset at 302 and a PCIE device at 304. According to FIG. 3, the device communicates its idle state capabilities to the core logic at 305. In this instance, as provide by the DMA Capability structure of the PCIE standard. In addition to collecting or gathering the power state capabilities of the device, the chipset may receive other platform related information. The chipset may use the received information, all or parts thereof, to determine or predict how long it can expect to be

idle, including expectations on exit latency. The core logic may predict how deep of an idle state it can expect to be in a next idle transition.

FIG. 4 is an illustrative example of power state information for the device connected at 304. As shown, the power state information may be captured in a table. It is noted however that any data structure may be used to contain or at least reference the device's power state parameters and values. The example of FIG. 4 details the idle states the device is capable of entering, as well as latency and power parameter values corresponding to the idle states. In this example, D O refers to an active state and states DOil and D0i2 refer to a light and a deeper idle state, respectively.

Based on the gathered or accumulated device power state and platform information, the core logic can inform the connected device to enter a lighter or deeper idle state (e.g., DOil and D0i2) that is supported by the device. That is, the core logic may provide guidance to the device regarding the idle power state(s) it should enter based, at least in part, the state of the chipset.

At operation 310, the device is armed to enter one of the idle states that it supports, DOil and D0i2, upon the next idle period by a communication from the core logic. In some aspects, this guiding of the device of which idle state to enter may operate to inform the device to prepare for an idle state transition. Such preparing may involve the device "cleaning up" and concluding other operations.

At operation 315, the platform has decided that it is in or entering an idle state, defined by all cores and activity generators coalescing. A signal, for example, OBFF Active to Idle, is sent to the external device (and internal devices in some embodiments) to indicate the platform desires to enter an idle state.

At this juncture of the process, the device may know that it should coalesce the activities it is generating, and that it is safe for it (i.e., the device) to enter into the one of the idle states determined and selected by the core logic.

The platform core logic expects to be idle for a certain period of time, as illustrated at 325. In some aspects, the core logic also knows the device recovery latency from the information received from the device and stored for use/reference by the chipset.

In some aspects, the core logic may cause the system to "warm up" and activate in advance of the device exiting the idle state and (potentially) needing platform resources. In some aspects, this "warming up" operation may work to minimize software latency by not adding to it.

At operation 320, the platform informs the device that is intends to exit the idle state. The device may then enter an active state as shown at 330. Thereafter, the device may resume DMA (and other) transactions.

It is seen that the present disclosure provides a mechanism to have devices enter a low (e.g., idle) power state in an orderly and sequenced manner. As illustrated by the foregoing discussion, a process in accordance herewith may include the selection of idle states for the device based, at least in part, on the device recovery latency.

In some embodiments, the operations and processes herein are implemented in hardware. This is in contrast to OS and other software directed means. The present disclosure may also be contrasted with link state power management protocols where a physical connection or link between end points is used, as opposed to the present disclosure that directly manages and controls the operation of the core logic and the device(s). In some aspects, the power management scheme herein is independent of a communication link state status.

All systems and processes discussed herein may be embodied in program code stored on one or more computer-readable media. Such media may include, for example, a floppy disk, a CD-ROM, a DVD-ROM, one or more types of "discs", magnetic tape, a memory card, a flash drive, a solid state drive, and solid state Random Access Memory (RAM) or Read Only Memory (ROM) storage units. Embodiments are therefore not limited to any specific combination of hardware and software.

Embodiments have been described herein solely for the purpose of illustration. Persons skilled in the art will recognize from this description that embodiments are not limited to those described, but may be practiced with modifications and alterations limited only by the spirit and scope of the appended claims.

CLAIMS

What is claimed is:

1. A method, the method comprising:

receiving an indication of a idle state capability of a platform connected device;

determining, by a chipset, an idle power state compatible with the device; and

directing the device to enter the determined idle power state based on a power state of the chipset.

2. The method of claim 1, wherein the indication of an idle state capability of the platform connected device includes information regarding at least one idle state and a corresponding latency and a power consumption associated therewith.

3. The method of claim 1, wherein the received information is stored for a future reference by the chipset.

4. The method of claim 1, wherein the determining of the idle power state compatible with the device is based, at least in part, on a latency of the device where information indicative of the latency is received from the device.

5. The method of claim 1, wherein the determining of the idle power state compatible with the device is based, at least in part, on a state of the chipset.

6. The method of claim 1, further comprising receiving an indication of at least one platform parameter value and further determining the idle power state compatible with the device based, at least in part, on the at least one received platform parameter value.

7. An apparatus, the apparatus comprising:

a platform connected device; and

a chipset comprising the platform, the chipset being operative to:

receive an indication of a idle state capability of the platform connected device;

determine an idle power state compatible with the device; and

direct the device to enter the determined idle power state based on a power state of the chipset.

8. The apparatus of claim 7, wherein the indication of an idle state capability of the platform connected device includes information regarding at least one idle state and a corresponding latency and a power consumption associated therewith.

9. The apparatus of claim 7, wherein the received information is stored for a future reference by the chipset.

10. The apparatus of claim 7, wherein the determining of the idle power state compatible with the device is based, at least in part, on a latency of the device where information indicative of the latency is received from the device.

11. The apparatus of claim 7, wherein the determining of the idle power state compatible with the device is based, at least in part, on a state of the chipset.

12. The apparatus of claim 7, the chipset being further operative to receive an indication of at least on platform parameter value and to determine the idle power state compatible with the device based, at least in part, on the at least one received platform parameter value.

13. A system, the system comprising:

a memory;

a platform connected device; and

a chipset comprising the platform, the chipset being operative to:

receive an indication of a idle state capability of the platform connected device;

determine an idle power state compatible with the device; and

direct the device to enter the determined idle power state based on a power state of the chipset.

14. The system of claim 13, wherein the indication of an idle state capability of the platform connected device includes information regarding at least one idle state and a corresponding latency and a power consumption associated therewith.

15. The system of claim 13, wherein the received information is stored for a future reference by the chipset.

16. The system of claim 13, wherein the determining of the idle power state compatible with the device is based, at least in part, on a latency of the device where information indicative of the latency is received from the device.

17. The system of claim 13, wherein the determining of the idle power state compatible with the device is based, at least in part, on a state of the chipset.

18. The system of claim 13, the chipset being further operative to receive an indication of at least on platform parameter value and to determine the idle power state compatible with the device based, at least in part, on the at least one received platform parameter value.

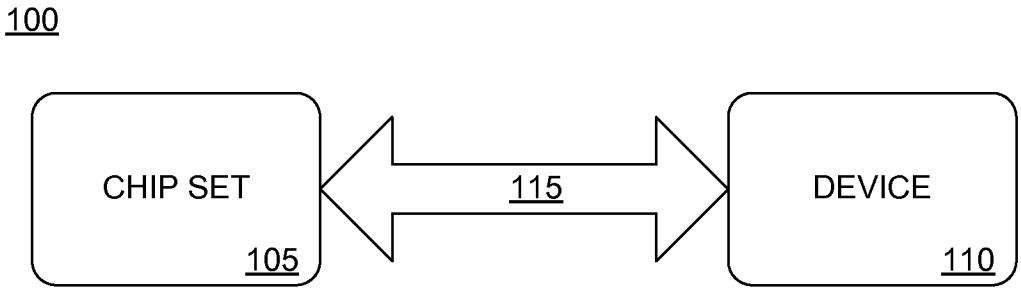

100
CHIP SET
105
115
DEVICE
110

FIG. 1

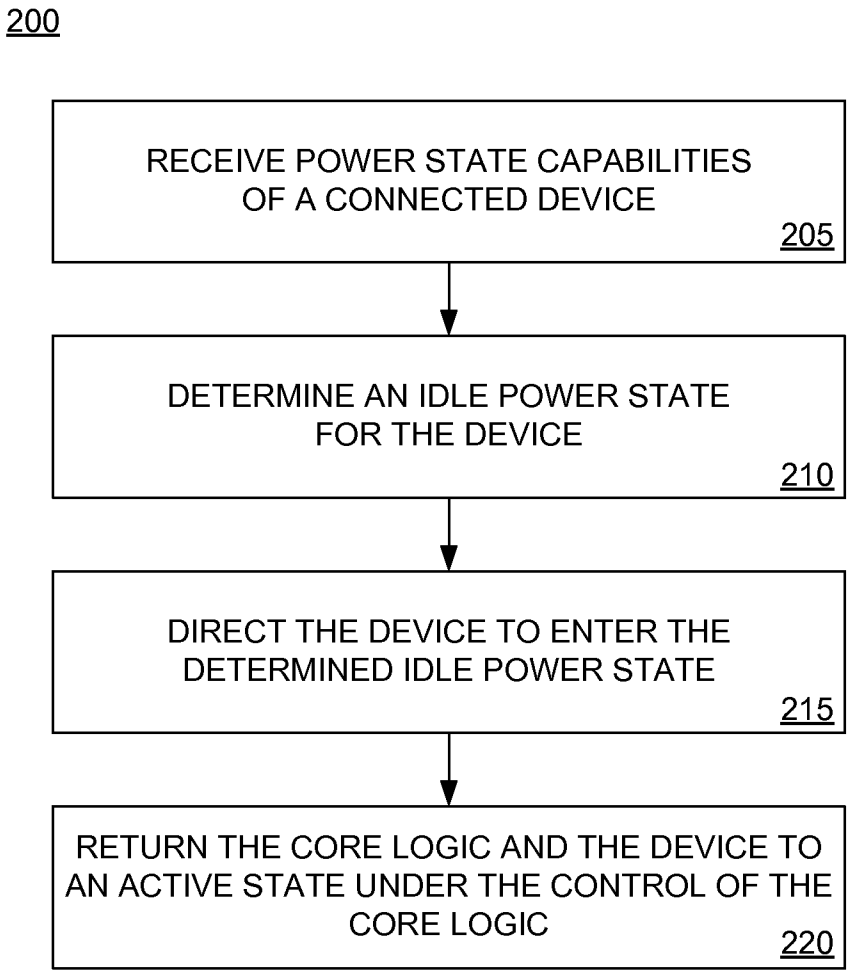

200
RECEIVE POWER STATE CAPABILITIES OF A CONNECTED DEVICE
205
DETERMINE AN IDLE POWER STATE FOR THE DEVICE
210
DIRECT THE DEVICE TO ENTER THE DETERMINED IDLE POWER STATE
215
RETURN THE CORE LOGIC AND THE DEVICE TO AN ACTIVE STATE UNDER THE CONTROL OF THE CORE LOGIC
220

FIG. 2

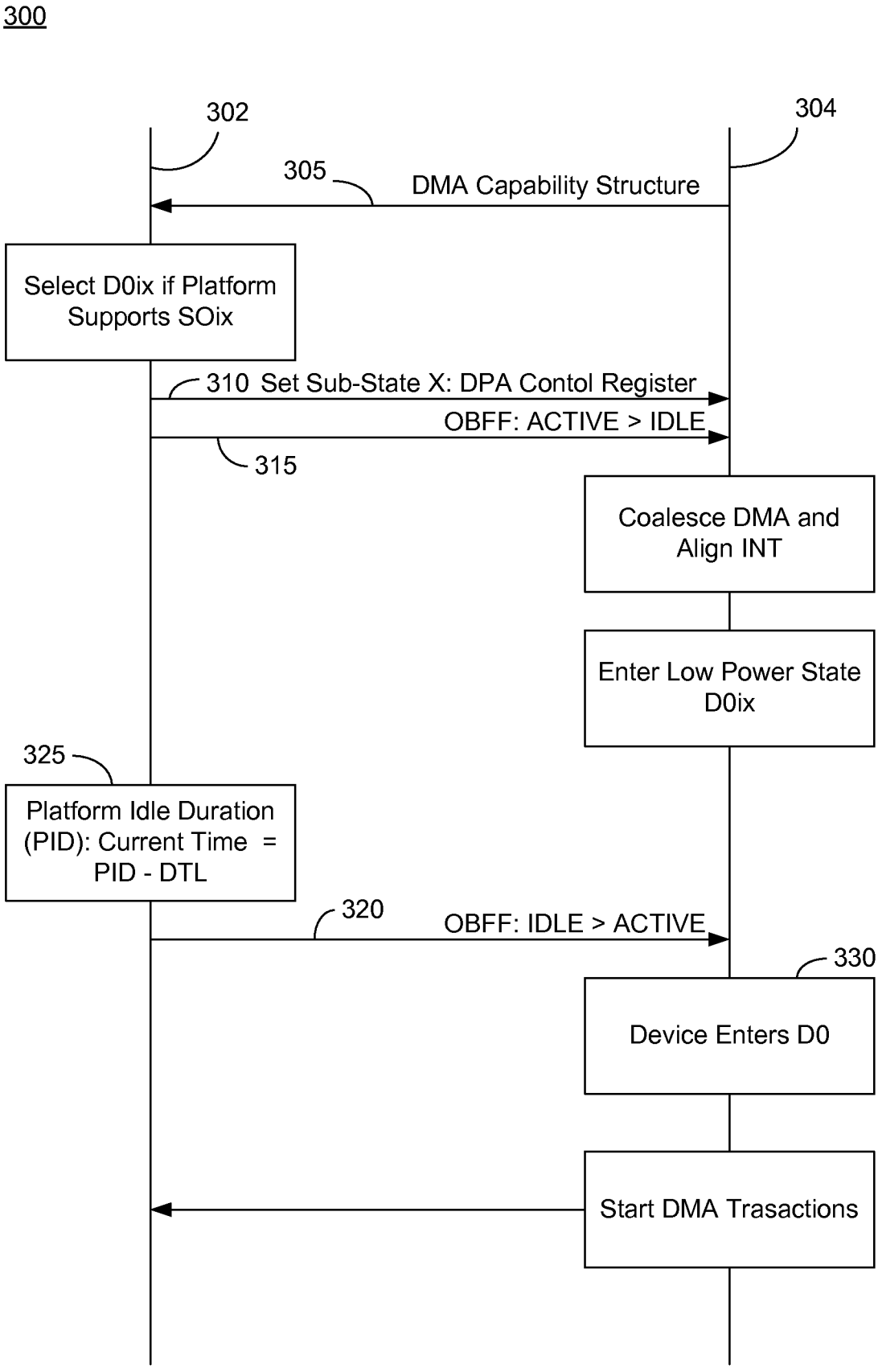

300
302
304
305
DMA Capability Structure
Select D0ix if Platform Supports SOix
310 Set Sub-State X: DPA Contol Register
OBFF: ACTIVE > IDLE
315
Coalesce DMA and Align INT
Enter Low Power State D0ix
325
Platform Idle Duration (PID): Current Time = PID - DTL
320
OBFF: IDLE > ACTIVE
330
Device Enters D0
Start DMA Trasactions

FIG. 3

Sub-State	DTL	Power (W)
D0	0	10
D0i1	100 µs	2
D0i2	100 µs	0.2

FIG. 4

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2011/068096

A. CLASSIFICATION OF SUBJECT MATTER

G06F l/32(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F 1/32; G06F 11/34; G06F 12/08

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Korean utility models and applications for utility models

Japanese utility models and applications for utility models

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

eKOMPASS(KIPO internal) & Keywords: plarform, idle state, power, management, capability, chipset

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2011-0154080 Al (WANG REN et al.) 23 June 2011 See figs. 1-4, paras .[0012]-[0014] , [0029]-0033] , [0044] and claims 1-4,12-15,18.	1-18
A	US 2010-0169683 Al (WANG REN et al.) 01 July 2010 See figs. 1-3, paras .[0010]-[0016] and claims 1-14.	1-18
A	US 2011-0264938 Al (HENROID ANDREW D. et al.) 27 October 2011 See figs. 1,2, 4 and claims 1-4.	1-18
A	US 05754869A A (HOLZHAMMER; GERALD S. et al.) 19 May 1998 See figs.l-5b and cliams 1-7.	1-18

☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
24 SEPTEMBER 2012 (24.09.2012)	**24 SEPTEMBER 2012 (24.09.2012)**

Name and mailing address of the ISA/KR	Authorized officer
Korean Intellectual Property Office 189 Cheongsa-ro, Seo-gu, Daejeon Metropolitan City, 302-701, Republic of Korea	Park Ji Eun
Facsimile No. 82-42-472-7140	Telephone No. 82-42-481-5696

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/US2011/068096

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2011-0154080 A1	23.06.2011	None	
US 2010-0169683 A1	01.07.2010	CN 101916137 A	15.12.2010
		JP 2010-191945 A	02.09.2010
		TW 201037506 A	16.10.2010
US 2011-0264938 A1	27.10.2011	None	
US 05754869A A	19.05.1998	US 05754869A A	19.05.1998