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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
5 October 2017 (05.10.2017)

WIPO | PCT

(10) International Publication Number
WO 2017/166272 A1

(51) International Patent Classification:
G06F 11/14 (2006.01)

(21) International Application Number:
PCT/CN2016/078282

(22) International Filing Date:
1 April 2016 (01.04.2016)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant: **INTEL CORPORATION** [US/US]; 2200 Mission College Boulevard, Santa Clara, California 95054 (US).

(72) Inventors; and
(71) Applicants *(for BZ only)*: **DONG, Yao Zu** [CN/CN]; Apt. 10I (1009), Building #5, Yanping Road 123, Shanghai 200242 (CN). **TIAN, Kun** [CN/CN]; No.880 Zi Xing Road, Shanghai Zizhu Science Park, Shanghai 200041 (CN).

(74) Agent: **CHINA PATENT AGENT (H.K.) LTD.**; 22/F., Great Eagle Center, 23 Harbour Road, Wanchai, Hong Kong (CN).

(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, BN, 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, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, 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, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, 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, KM, ML, MR, NE, SN, TD, TG).

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

(54) **Title**: METHOD AND APPARATUS OF PERIODIC SNAPSHOTTING IN GRAPHICS PROCESSING ENVIRONMENT

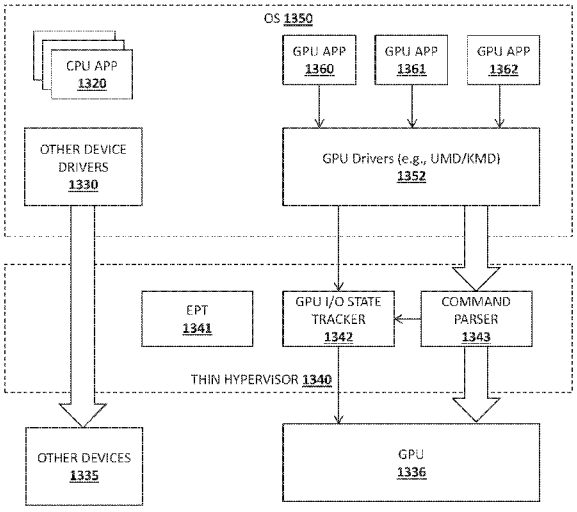


FIG. 13

(57) **Abstract**: An apparatus and method of performing debug and rollback operations using snapshots. The apparatus comprises: a graphics processing unit (GPU) (1336) to perform graphics processing operations by executing graphics commands; a command parser (1343) to parse graphics commands submitted to the GPU (1336) and generate a list of graphics memory pages which will be affected by the graphics commands; an I/O state tracker (1342) to track I/O accesses from a graphics driver (1352) to determine a list of registers affected by the I/O accesses; snapshot circuitry and/or logic to perform a memory snapshot and I/O snapshot based on the list of graphics memory pages and the list of registers, respectively; and rollback circuitry and/or logic to perform a rollback operation using the memory snapshot and I/O snapshot in response to detecting a GPU error condition.

WO 2017/166272 A1

METHOD AND APPARATUS PERIODIC SNAPSHOTTING IN GRAPHICS PROCESSING ENVIRONMENT

BACKGROUND

Field of the Invention

[0001] This invention relates generally to the field of computer processors. More particularly, the invention relates to an apparatus and method for periodic snapshotting in a graphics processing environment.

Description of the Related Art

[0002] Graphics Processing Unit (GPU) computing has been widely used for a variety of applications including 3D rendering, media transcoding, high performance computing, real time image processing, and packet filtering, to name a few. However, due to the complexity associated with GPU microarchitecture and programming GPU software may crash more frequently than software designed for general purpose CPUs. Existing approaches solve the problem by restarting the software and/or relying on Timeout Detection and Recovery (TDR) to reset the software. However these solutions are insufficient since they are incapable of determining the root cause of the problem. Fine-grain debuggers which may be used to perform GPU program debugging are also insufficient. For example, if a bug is encountered after long period of execution (e.g., a Mean Time Between Failures of 20 hours) it may be impossible to determine the cause. As such, fine-grain debuggers are impractical for comprehensive, hard-reproducible bugs.

BRIEF DESCRIPTION OF THE DRAWINGS

[0003] A better understanding of the present invention can be obtained from the following detailed description in conjunction with the following drawings, in which:

[0004] **FIG. 1** is a block diagram of an embodiment of a computer system with a processor having one or more processor cores and graphics processors;

[0005] **FIG. 2** is a block diagram of one embodiment of a processor having one or more processor cores, an integrated memory controller, and an integrated graphics processor;

[0006] **FIG. 3** is a block diagram of one embodiment of a graphics processor which may be a discreet graphics processing unit, or may be graphics processor integrated with a plurality of processing cores;

[0007] **FIG. 4** is a block diagram of an embodiment of a graphics-processing engine for a graphics processor;

[0008] **FIG. 5** is a block diagram of another embodiment of a graphics processor;

[0009] **FIG. 6** is a block diagram of thread execution logic including an array of processing elements;

[0010] **FIG. 7** illustrates a graphics processor execution unit instruction format according to an embodiment;

[0011] **FIG. 8** is a block diagram of another embodiment of a graphics processor which includes a graphics pipeline, a media pipeline, a display engine, thread execution logic, and a render output pipeline;

[0012] **FIG. 9A** is a block diagram illustrating a graphics processor command format according to an embodiment;

[0013] **FIG. 9B** is a block diagram illustrating a graphics processor command sequence according to an embodiment;

[0014] **FIG. 10** illustrates exemplary graphics software architecture for a data processing system according to an embodiment;

[0015] **FIG. 11** illustrates an exemplary IP core development system that may be used to manufacture an integrated circuit to perform operations according to an embodiment;

[0016] **FIG. 12** illustrates an exemplary system on a chip integrated circuit that may be fabricated using one or more IP cores, according to an embodiment;

[0017] **FIG. 13** illustrates an architecture in accordance with one embodiment of the invention; and

[0018] **FIG. 14** illustrates a method in accordance with one embodiment of the invention.

DETAILED DESCRIPTION

[0019] In the following description, for the purposes of explanation, numerous specific details are set forth in order to provide a thorough understanding of the embodiments of the invention described below. It will be apparent, however, to one skilled in the art that the embodiments of the invention may be practiced without some of these specific details. In other instances, well-known structures and devices are shown in block diagram form to avoid obscuring the underlying principles of the embodiments of the invention.

EXEMPLARY GRAPHICS PROCESSOR ARCHITECTURES AND DATA TYPES

[0020] **System Overview**

[0021] **Figure 1** is a block diagram of a processing system 100, according to an embodiment. In various embodiments the system 100 includes one or more processors 102 and one or more graphics processors 108, and may be a single processor desktop system, a multiprocessor workstation system, or a server system having a large number of processors 102 or processor cores 107. In on embodiment, the system 100 is a processing platform incorporated within a system-on-a-chip (SoC) integrated circuit for use in mobile, handheld, or embedded devices.

[0022] An embodiment of system 100 can include, or be incorporated within a server-based gaming platform, a game console, including a game and media console, a mobile gaming console, a handheld game console, or an online game console. In some embodiments system 100 is a mobile phone, smart phone, tablet computing device or mobile Internet device. Data processing system 100 can also include, couple with, or be integrated within a wearable device, such as a smart watch wearable device, smart eyewear device, augmented reality device, or virtual reality device. In some embodiments, data

processing system 100 is a television or set top box device having one or more processors 102 and a graphical interface generated by one or more graphics processors 108.

[0023] In some embodiments, the one or more processors 102 each include one or more processor cores 107 to process instructions which, when executed, perform operations for system and user software. In some embodiments, each of the one or more processor cores 107 is configured to process a specific instruction set 109. In some embodiments, instruction set 109 may facilitate Complex Instruction Set Computing (CISC), Reduced Instruction Set Computing (RISC), or computing via a Very Long Instruction Word (VLIW). Multiple processor cores 107 may each process a different instruction set 109, which may include instructions to facilitate the emulation of other instruction sets. Processor core 107 may also include other processing devices, such a Digital Signal Processor (DSP).

[0024] In some embodiments, the processor 102 includes cache memory 104. Depending on the architecture, the processor 102 can have a single internal cache or multiple levels of internal cache. In some embodiments, the cache memory is shared among various components of the processor 102. In some embodiments, the processor 102 also uses an external cache (e.g., a Level-3 (L3) cache or Last Level Cache (LLC)) (not shown), which may be shared among processor cores 107 using known cache coherency techniques. A register file 106 is additionally included in processor 102 which may include different types of registers for storing different types of data (e.g., integer registers, floating point registers, status registers, and an instruction pointer register). Some registers may be general-purpose registers, while other registers may be specific to the design of the processor 102.

[0025] In some embodiments, processor 102 is coupled to a processor bus 110 to transmit communication signals such as address, data, or control signals between processor 102 and other components in system 100. In one embodiment the system 100 uses an exemplary 'hub' system architecture, including a memory controller hub 116 and an Input Output (I/O) controller hub 130. A memory controller hub 116 facilitates communication between a memory device and other components of system 100, while an I/O Controller Hub (ICH) 130 provides connections to I/O devices via a local I/O bus. In one embodiment, the logic of the memory controller hub 116 is integrated within the processor.

[0026] Memory device 120 can be a dynamic random access memory (DRAM) device, a static random access memory (SRAM) device, flash memory device, phase-change memory device, or some other memory device having suitable performance to serve as process memory. In one embodiment the memory device 120 can operate as system memory for the system 100, to store data 122 and instructions 121 for use when the one or more processors 102 executes an application or process. Memory controller hub 116 also couples with an optional external graphics processor 112, which may communicate with the one or more graphics processors 108 in processors 102 to perform graphics and media operations.

[0027] In some embodiments, ICH 130 enables peripherals to connect to memory device 120 and processor 102 via a high-speed I/O bus. The I/O peripherals include, but are not limited to, an audio controller 146, a firmware interface 128, a wireless transceiver 126 (e.g., Wi-Fi, Bluetooth), a data storage device 124 (e.g., hard disk drive, flash memory, etc.), and a legacy I/O controller 140 for coupling legacy (e.g., Personal System 2 (PS/2)) devices to the system. One or more Universal Serial Bus (USB) controllers 142 connect input devices,

such as keyboard and mouse 144 combinations. A network controller 134 may also couple to ICH 130. In some embodiments, a high-performance network controller (not shown) couples to processor bus 110. It will be appreciated that the system 100 shown is exemplary and not limiting, as other types of data processing systems that are differently configured may also be used. For example, the I/O controller hub 130 may be integrated within the one or more processor 102, or the memory controller hub 116 and I/O controller hub 130 may be integrated into a discreet external graphics processor, such as the external graphics processor 112.

[0028] **Figure 2** is a block diagram of an embodiment of a processor 200 having one or more processor cores 202A-202N, an integrated memory controller 214, and an integrated graphics processor 208. Those elements of **Figure 2** having the same reference numbers (or names) as the elements of any other figure herein can operate or function in any manner similar to that described elsewhere herein, but are not limited to such. Processor 200 can include additional cores up to and including additional core 202N represented by the dashed lined boxes. Each of processor cores 202A-202N includes one or more internal cache units 204A-204N. In some embodiments each processor core also has access to one or more shared cached units 206.

[0029] The internal cache units 204A-204N and shared cache units 206 represent a cache memory hierarchy within the processor 200. The cache memory hierarchy may include at least one level of instruction and data cache within each processor core and one or more levels of shared mid-level cache, such as a Level 2 (L2), Level 3 (L3), Level 4 (L4), or other levels of cache, where the highest level of cache before external memory is classified as the LLC. In

some embodiments, cache coherency logic maintains coherency between the various cache units 206 and 204A-204N.

[0030] In some embodiments, processor 200 may also include a set of one or more bus controller units 216 and a system agent core 210. The one or more bus controller units 216 manage a set of peripheral buses, such as one or more Peripheral Component Interconnect buses (e.g., PCI, PCI Express). System agent core 210 provides management functionality for the various processor components. In some embodiments, system agent core 210 includes one or more integrated memory controllers 214 to manage access to various external memory devices (not shown).

[0031] In some embodiments, one or more of the processor cores 202A-202N include support for simultaneous multi-threading. In such embodiment, the system agent core 210 includes components for coordinating and operating cores 202A-202N during multi-threaded processing. System agent core 210 may additionally include a power control unit (PCU), which includes logic and components to regulate the power state of processor cores 202A-202N and graphics processor 208.

[0032] In some embodiments, processor 200 additionally includes graphics processor 208 to execute graphics processing operations. In some embodiments, the graphics processor 208 couples with the set of shared cache units 206, and the system agent core 210, including the one or more integrated memory controllers 214. In some embodiments, a display controller 211 is coupled with the graphics processor 208 to drive graphics processor output to one or more coupled displays. In some embodiments, display controller 211 may be a separate module coupled with the graphics processor via at least one

interconnect, or may be integrated within the graphics processor 208 or system agent core 210.

[0033] In some embodiments, a ring based interconnect unit 212 is used to couple the internal components of the processor 200. However, an alternative interconnect unit may be used, such as a point-to-point interconnect, a switched interconnect, or other techniques, including techniques well known in the art. In some embodiments, graphics processor 208 couples with the ring interconnect 212 via an I/O link 213.

[0034] The exemplary I/O link 213 represents at least one of multiple varieties of I/O interconnects, including an on package I/O interconnect which facilitates communication between various processor components and a high-performance embedded memory module 218, such as an eDRAM module. In some embodiments, each of the processor cores 202-202N and graphics processor 208 use embedded memory modules 218 as a shared Last Level Cache.

[0035] In some embodiments, processor cores 202A-202N are homogenous cores executing the same instruction set architecture. In another embodiment, processor cores 202A-202N are heterogeneous in terms of instruction set architecture (ISA), where one or more of processor cores 202A-N execute a first instruction set, while at least one of the other cores executes a subset of the first instruction set or a different instruction set. In one embodiment processor cores 202A-202N are heterogeneous in terms of microarchitecture, where one or more cores having a relatively higher power consumption couple with one or more power cores having a lower power consumption. Additionally, processor 200 can be implemented on one or more chips or as an SoC

integrated circuit having the illustrated components, in addition to other components.

[0036] **Figure 3** is a block diagram of a graphics processor 300, which may be a discrete graphics processing unit, or may be a graphics processor integrated with a plurality of processing cores. In some embodiments, the graphics processor communicates via a memory mapped I/O interface to registers on the graphics processor and with commands placed into the processor memory. In some embodiments, graphics processor 300 includes a memory interface 314 to access memory. Memory interface 314 can be an interface to local memory, one or more internal caches, one or more shared external caches, and/or to system memory.

[0037] In some embodiments, graphics processor 300 also includes a display controller 302 to drive display output data to a display device 320. Display controller 302 includes hardware for one or more overlay planes for the display and composition of multiple layers of video or user interface elements. In some embodiments, graphics processor 300 includes a video codec engine 306 to encode, decode, or transcode media to, from, or between one or more media encoding formats, including, but not limited to Moving Picture Experts Group (MPEG) formats such as MPEG-2, Advanced Video Coding (AVC) formats such as H.264/MPEG-4 AVC, as well as the Society of Motion Picture & Television Engineers (SMPTE) 421M/VC-1, and Joint Photographic Experts Group (JPEG) formats such as JPEG, and Motion JPEG (MJPEG) formats.

[0038] In some embodiments, graphics processor 300 includes a block image transfer (BLIT) engine 304 to perform two-dimensional (2D) rasterizer operations including, for example, bit-boundary block transfers. However, in one

embodiment, 2D graphics operations are performed using one or more components of graphics processing engine (GPE) 310. In some embodiments, graphics processing engine 310 is a compute engine for performing graphics operations, including three-dimensional (3D) graphics operations and media operations.

[0039] In some embodiments, GPE 310 includes a 3D pipeline 312 for performing 3D operations, such as rendering three-dimensional images and scenes using processing functions that act upon 3D primitive shapes (e.g., rectangle, triangle, etc.). The 3D pipeline 312 includes programmable and fixed function elements that perform various tasks within the element and/or spawn execution threads to a 3D/Media sub-system 315. While 3D pipeline 312 can be used to perform media operations, an embodiment of GPE 310 also includes a media pipeline 316 that is specifically used to perform media operations, such as video post-processing and image enhancement.

[0040] In some embodiments, media pipeline 316 includes fixed function or programmable logic units to perform one or more specialized media operations, such as video decode acceleration, video de-interlacing, and video encode acceleration in place of, or on behalf of video codec engine 306. In some embodiments, media pipeline 316 additionally includes a thread spawning unit to spawn threads for execution on 3D/Media sub-system 315. The spawned threads perform computations for the media operations on one or more graphics execution units included in 3D/Media sub-system 315.

[0041] In some embodiments, 3D/Media subsystem 315 includes logic for executing threads spawned by 3D pipeline 312 and media pipeline 316. In one embodiment, the pipelines send thread execution requests to 3D/Media

subsystem 315, which includes thread dispatch logic for arbitrating and dispatching the various requests to available thread execution resources. The execution resources include an array of graphics execution units to process the 3D and media threads. In some embodiments, 3D/Media subsystem 315 includes one or more internal caches for thread instructions and data. In some embodiments, the subsystem also includes shared memory, including registers and addressable memory, to share data between threads and to store output data.

[0042] **3D/Media Processing**

[0043] **Figure 4** is a block diagram of a graphics processing engine 410 of a graphics processor in accordance with some embodiments. In one embodiment, the GPE 410 is a version of the GPE 310 shown in **Figure 3**. Elements of **Figure 4** having the same reference numbers (or names) as the elements of any other figure herein can operate or function in any manner similar to that described elsewhere herein, but are not limited to such.

[0044] In some embodiments, GPE 410 couples with a command streamer 403, which provides a command stream to the GPE 3D and media pipelines 412, 416. In some embodiments, command streamer 403 is coupled to memory, which can be system memory, or one or more of internal cache memory and shared cache memory. In some embodiments, command streamer 403 receives commands from the memory and sends the commands to 3D pipeline 412 and/or media pipeline 416. The commands are directives fetched from a ring buffer, which stores commands for the 3D and media pipelines 412, 416. In one embodiment, the ring buffer can additionally include batch command buffers storing batches of multiple commands. The 3D and media pipelines 412, 416

process the commands by performing operations via logic within the respective pipelines or by dispatching one or more execution threads to an execution unit array 414. In some embodiments, execution unit array 414 is scalable, such that the array includes a variable number of execution units based on the target power and performance level of GPE 410.

[0045] In some embodiments, a sampling engine 430 couples with memory (e.g., cache memory or system memory) and execution unit array 414. In some embodiments, sampling engine 430 provides a memory access mechanism for execution unit array 414 that allows execution array 414 to read graphics and media data from memory. In some embodiments, sampling engine 430 includes logic to perform specialized image sampling operations for media.

[0046] In some embodiments, the specialized media sampling logic in sampling engine 430 includes a de-noise/de-interlace module 432, a motion estimation module 434, and an image scaling and filtering module 436. In some embodiments, de-noise/de-interlace module 432 includes logic to perform one or more of a de-noise or a de-interlace algorithm on decoded video data. The de-interlace logic combines alternating fields of interlaced video content into a single fame of video. The de-noise logic reduces or removes data noise from video and image data. In some embodiments, the de-noise logic and de-interlace logic are motion adaptive and use spatial or temporal filtering based on the amount of motion detected in the video data. In some embodiments, the de-noise/de-interlace module 432 includes dedicated motion detection logic (e.g., within the motion estimation engine 434).

[0047] In some embodiments, motion estimation engine 434 provides hardware acceleration for video operations by performing video acceleration

functions such as motion vector estimation and prediction on video data. The motion estimation engine determines motion vectors that describe the transformation of image data between successive video frames. In some embodiments, a graphics processor media codec uses video motion estimation engine 434 to perform operations on video at the macro-block level that may otherwise be too computationally intensive to perform with a general-purpose processor. In some embodiments, motion estimation engine 434 is generally available to graphics processor components to assist with video decode and processing functions that are sensitive or adaptive to the direction or magnitude of the motion within video data.

[0048] In some embodiments, image scaling and filtering module 436 performs image-processing operations to enhance the visual quality of generated images and video. In some embodiments, scaling and filtering module 436 processes image and video data during the sampling operation before providing the data to execution unit array 414.

[0049] In some embodiments, the GPE 410 includes a data port 444, which provides an additional mechanism for graphics subsystems to access memory. In some embodiments, data port 444 facilitates memory access for operations including render target writes, constant buffer reads, scratch memory space reads/writes, and media surface accesses. In some embodiments, data port 444 includes cache memory space to cache accesses to memory. The cache memory can be a single data cache or separated into multiple caches for the multiple subsystems that access memory via the data port (e.g., a render buffer cache, a constant buffer cache, etc.). In some embodiments, threads executing on an execution unit in execution unit array 414 communicate with the

data port by exchanging messages via a data distribution interconnect that couples each of the sub-systems of GPE 410.

[0050] **Execution Units**

[0051] **Figure 5** is a block diagram of another embodiment of a graphics processor 500. Elements of **Figure 5** having the same reference numbers (or names) as the elements of any other figure herein can operate or function in any manner similar to that described elsewhere herein, but are not limited to such.

[0052] In some embodiments, graphics processor 500 includes a ring interconnect 502, a pipeline front-end 504, a media engine 537, and graphics cores 580A-580N. In some embodiments, ring interconnect 502 couples the graphics processor to other processing units, including other graphics processors or one or more general-purpose processor cores. In some embodiments, the graphics processor is one of many processors integrated within a multi-core processing system.

[0053] In some embodiments, graphics processor 500 receives batches of commands via ring interconnect 502. The incoming commands are interpreted by a command streamer 503 in the pipeline front-end 504. In some embodiments, graphics processor 500 includes scalable execution logic to perform 3D geometry processing and media processing via the graphics core(s) 580A-580N. For 3D geometry processing commands, command streamer 503 supplies commands to geometry pipeline 536. For at least some media processing commands, command streamer 503 supplies the commands to a video front end 534, which couples with a media engine 537. In some embodiments, media engine 537 includes a Video Quality Engine (VQE) 530 for video and image post-processing and a multi-format encode/decode (MFX) 533

engine to provide hardware-accelerated media data encode and decode. In some embodiments, geometry pipeline 536 and media engine 537 each generate execution threads for the thread execution resources provided by at least one graphics core 580A.

[0054] In some embodiments, graphics processor 500 includes scalable thread execution resources featuring modular cores 580A-580N (sometimes referred to as core slices), each having multiple sub-cores 550A-550N, 560A-560N (sometimes referred to as core sub-slices). In some embodiments, graphics processor 500 can have any number of graphics cores 580A through 580N. In some embodiments, graphics processor 500 includes a graphics core 580A having at least a first sub-core 550A and a second core sub-core 560A. In other embodiments, the graphics processor is a low power processor with a single sub-core (e.g., 550A). In some embodiments, graphics processor 500 includes multiple graphics cores 580A-580N, each including a set of first sub-cores 550A-550N and a set of second sub-cores 560A-560N. Each sub-core in the set of first sub-cores 550A-550N includes at least a first set of execution units 552A-552N and media/texture samplers 554A-554N. Each sub-core in the set of second sub-cores 560A-560N includes at least a second set of execution units 562A-562N and samplers 564A-564N. In some embodiments, each sub-core 550A-550N, 560A-560N shares a set of shared resources 570A-570N. In some embodiments, the shared resources include shared cache memory and pixel operation logic. Other shared resources may also be included in the various embodiments of the graphics processor.

[0055] **Figure 6** illustrates thread execution logic 600 including an array of processing elements employed in some embodiments of a GPE. Elements of **Figure 6** having the same reference numbers (or names) as the elements of any

other figure herein can operate or function in any manner similar to that described elsewhere herein, but are not limited to such.

[0056] In some embodiments, thread execution logic 600 includes a pixel shader 602, a thread dispatcher 604, instruction cache 606, a scalable execution unit array including a plurality of execution units 608A-608N, a sampler 610, a data cache 612, and a data port 614. In one embodiment the included components are interconnected via an interconnect fabric that links to each of the components. In some embodiments, thread execution logic 600 includes one or more connections to memory, such as system memory or cache memory, through one or more of instruction cache 606, data port 614, sampler 610, and execution unit array 608A-608N. In some embodiments, each execution unit (e.g. 608A) is an individual vector processor capable of executing multiple simultaneous threads and processing multiple data elements in parallel for each thread. In some embodiments, execution unit array 608A-608N includes any number individual execution units.

[0057] In some embodiments, execution unit array 608A-608N is primarily used to execute “shader” programs. In some embodiments, the execution units in array 608A-608N execute an instruction set that includes native support for many standard 3D graphics shader instructions, such that shader programs from graphics libraries (e.g., Direct 3D and OpenGL) are executed with a minimal translation. The execution units support vertex and geometry processing (e.g., vertex programs, geometry programs, vertex shaders), pixel processing (e.g., pixel shaders, fragment shaders) and general-purpose processing (e.g., compute and media shaders).

[0058] Each execution unit in execution unit array 608A-608N operates on arrays of data elements. The number of data elements is the "execution size," or the number of channels for the instruction. An execution channel is a logical unit of execution for data element access, masking, and flow control within instructions. The number of channels may be independent of the number of physical Arithmetic Logic Units (ALUs) or Floating Point Units (FPUs) for a particular graphics processor. In some embodiments, execution units 608A-608N support integer and floating-point data types.

[0059] The execution unit instruction set includes single instruction multiple data (SIMD) instructions. The various data elements can be stored as a packed data type in a register and the execution unit will process the various elements based on the data size of the elements. For example, when operating on a 256-bit wide vector, the 256 bits of the vector are stored in a register and the execution unit operates on the vector as four separate 64-bit packed data elements (Quad-Word (QW) size data elements), eight separate 32-bit packed data elements (Double Word (DW) size data elements), sixteen separate 16-bit packed data elements (Word (W) size data elements), or thirty-two separate 8-bit data elements (byte (B) size data elements). However, different vector widths and register sizes are possible.

[0060] One or more internal instruction caches (e.g., 606) are included in the thread execution logic 600 to cache thread instructions for the execution units. In some embodiments, one or more data caches (e.g., 612) are included to cache thread data during thread execution. In some embodiments, sampler 610 is included to provide texture sampling for 3D operations and media sampling for media operations. In some embodiments, sampler 610 includes specialized texture or media sampling functionality to process texture or media

data during the sampling process before providing the sampled data to an execution unit.

[0061] During execution, the graphics and media pipelines send thread initiation requests to thread execution logic 600 via thread spawning and dispatch logic. In some embodiments, thread execution logic 600 includes a local thread dispatcher 604 that arbitrates thread initiation requests from the graphics and media pipelines and instantiates the requested threads on one or more execution units 608A-608N. For example, the geometry pipeline (e.g., 536 of **Figure 5**) dispatches vertex processing, tessellation, or geometry processing threads to thread execution logic 600 (**Figure 6**). In some embodiments, thread dispatcher 604 can also process runtime thread spawning requests from the executing shader programs.

[0062] Once a group of geometric objects has been processed and rasterized into pixel data, pixel shader 602 is invoked to further compute output information and cause results to be written to output surfaces (e.g., color buffers, depth buffers, stencil buffers, etc.). In some embodiments, pixel shader 602 calculates the values of the various vertex attributes that are to be interpolated across the rasterized object. In some embodiments, pixel shader 602 then executes an application programming interface (API)-supplied pixel shader program. To execute the pixel shader program, pixel shader 602 dispatches threads to an execution unit (e.g., 608A) via thread dispatcher 604. In some embodiments, pixel shader 602 uses texture sampling logic in sampler 610 to access texture data in texture maps stored in memory. Arithmetic operations on the texture data and the input geometry data compute pixel color data for each geometric fragment, or discards one or more pixels from further processing.

[0063] In some embodiments, the data port 614 provides a memory access mechanism for the thread execution logic 600 output processed data to memory for processing on a graphics processor output pipeline. In some embodiments, the data port 614 includes or couples to one or more cache memories (e.g., data cache 612) to cache data for memory access via the data port.

[0064] **Figure 7** is a block diagram illustrating a graphics processor instruction formats 700 according to some embodiments. In one or more embodiment, the graphics processor execution units support an instruction set having instructions in multiple formats. The solid lined boxes illustrate the components that are generally included in an execution unit instruction, while the dashed lines include components that are optional or that are only included in a sub-set of the instructions. In some embodiments, instruction format 700 described and illustrated are macro-instructions, in that they are instructions supplied to the execution unit, as opposed to micro-operations resulting from instruction decode once the instruction is processed.

[0065] In some embodiments, the graphics processor execution units natively support instructions in a 128-bit format 710. A 64-bit compacted instruction format 730 is available for some instructions based on the selected instruction, instruction options, and number of operands. The native 128-bit format 710 provides access to all instruction options, while some options and operations are restricted in the 64-bit format 730. The native instructions available in the 64-bit format 730 vary by embodiment. In some embodiments, the instruction is compacted in part using a set of index values in an index field 713. The execution unit hardware references a set of compaction tables based

on the index values and uses the compaction table outputs to reconstruct a native instruction in the 128-bit format 710.

[0066] For each format, instruction opcode 712 defines the operation that the execution unit is to perform. The execution units execute each instruction in parallel across the multiple data elements of each operand. For example, in response to an add instruction the execution unit performs a simultaneous add operation across each color channel representing a texture element or picture element. By default, the execution unit performs each instruction across all data channels of the operands. In some embodiments, instruction control field 714 enables control over certain execution options, such as channels selection (e.g., predication) and data channel order (e.g., swizzle). For 128-bit instructions 710 an exec-size field 716 limits the number of data channels that will be executed in parallel. In some embodiments, exec-size field 716 is not available for use in the 64-bit compact instruction format 730.

[0067] Some execution unit instructions have up to three operands including two source operands, src0 722, src1 722, and one destination 718. In some embodiments, the execution units support dual destination instructions, where one of the destinations is implied. Data manipulation instructions can have a third source operand (e.g., SRC2 724), where the instruction opcode 712 determines the number of source operands. An instruction's last source operand can be an immediate (e.g., hard-coded) value passed with the instruction.

[0068] In some embodiments, the 128-bit instruction format 710 includes an access/address mode information 726 specifying, for example, whether direct register addressing mode or indirect register addressing mode is used. When

direct register addressing mode is used, the register address of one or more operands is directly provided by bits in the instruction 710.

[0069] In some embodiments, the 128-bit instruction format 710 includes an access/address mode field 726, which specifies an address mode and/or an access mode for the instruction. In one embodiment the access mode to define a data access alignment for the instruction. Some embodiments support access modes including a 16-byte aligned access mode and a 1-byte aligned access mode, where the byte alignment of the access mode determines the access alignment of the instruction operands. For example, when in a first mode, the instruction 710 may use byte-aligned addressing for source and destination operands and when in a second mode, the instruction 710 may use 16-byte-aligned addressing for all source and destination operands.

[0070] In one embodiment, the address mode portion of the access/address mode field 726 determines whether the instruction is to use direct or indirect addressing. When direct register addressing mode is used bits in the instruction 710 directly provide the register address of one or more operands. When indirect register addressing mode is used, the register address of one or more operands may be computed based on an address register value and an address immediate field in the instruction.

[0071] In some embodiments instructions are grouped based on opcode 712 bit-fields to simplify Opcode decode 740. For an 8-bit opcode, bits 4, 5, and 6 allow the execution unit to determine the type of opcode. The precise opcode grouping shown is merely an example. In some embodiments, a move and logic opcode group 742 includes data movement and logic instructions (e.g., move (mov), compare (cmp)). In some embodiments, move and logic group 742

shares the five most significant bits (MSB), where move (mov) instructions are in the form of 0000xxxxb and logic instructions are in the form of 0001xxxxb. A flow control instruction group 744 (e.g., call, jump (jmp)) includes instructions in the form of 0010xxxxb (e.g., 0x20). A miscellaneous instruction group 746 includes a mix of instructions, including synchronization instructions (e.g., wait, send) in the form of 0011xxxxb (e.g., 0x30). A parallel math instruction group 748 includes component-wise arithmetic instructions (e.g., add, multiply (mul)) in the form of 0100xxxxb (e.g., 0x40). The parallel math group 748 performs the arithmetic operations in parallel across data channels. The vector math group 750 includes arithmetic instructions (e.g., dp4) in the form of 0101xxxxb (e.g., 0x50). The vector math group performs arithmetic such as dot product calculations on vector operands.

[0072] **Graphics Pipeline**

[0073] **Figure 8** is a block diagram of another embodiment of a graphics processor 800. Elements of **Figure 8** having the same reference numbers (or names) as the elements of any other figure herein can operate or function in any manner similar to that described elsewhere herein, but are not limited to such.

[0074] In some embodiments, graphics processor 800 includes a graphics pipeline 820, a media pipeline 830, a display engine 840, thread execution logic 850, and a render output pipeline 870. In some embodiments, graphics processor 800 is a graphics processor within a multi-core processing system that includes one or more general purpose processing cores. The graphics processor is controlled by register writes to one or more control registers (not shown) or via commands issued to graphics processor 800 via a ring interconnect 802. In some embodiments, ring interconnect 802 couples graphics processor 800 to

other processing components, such as other graphics processors or general-purpose processors. Commands from ring interconnect 802 are interpreted by a command streamer 803, which supplies instructions to individual components of graphics pipeline 820 or media pipeline 830.

[0075] In some embodiments, command streamer 803 directs the operation of a vertex fetcher 805 that reads vertex data from memory and executes vertex-processing commands provided by command streamer 803. In some embodiments, vertex fetcher 805 provides vertex data to a vertex shader 807, which performs coordinate space transformation and lighting operations to each vertex. In some embodiments, vertex fetcher 805 and vertex shader 807 execute vertex-processing instructions by dispatching execution threads to execution units 852A, 852B via a thread dispatcher 831.

[0076] In some embodiments, execution units 852A, 852B are an array of vector processors having an instruction set for performing graphics and media operations. In some embodiments, execution units 852A, 852B have an attached L1 cache 851 that is specific for each array or shared between the arrays. The cache can be configured as a data cache, an instruction cache, or a single cache that is partitioned to contain data and instructions in different partitions.

[0077] In some embodiments, graphics pipeline 820 includes tessellation components to perform hardware-accelerated tessellation of 3D objects. In some embodiments, a programmable hull shader 811 configures the tessellation operations. A programmable domain shader 817 provides back-end evaluation of tessellation output. A tessellator 813 operates at the direction of hull shader 811 and contains special purpose logic to generate a set of detailed geometric objects based on a coarse geometric model that is provided as input to graphics

pipeline 820. In some embodiments, if tessellation is not used, tessellation components 811, 813, 817 can be bypassed.

[0078] In some embodiments, complete geometric objects can be processed by a geometry shader 819 via one or more threads dispatched to execution units 852A, 852B, or can proceed directly to the clipper 829. In some embodiments, the geometry shader operates on entire geometric objects, rather than vertices or patches of vertices as in previous stages of the graphics pipeline. If the tessellation is disabled the geometry shader 819 receives input from the vertex shader 807. In some embodiments, geometry shader 819 is programmable by a geometry shader program to perform geometry tessellation if the tessellation units are disabled.

[0079] Before rasterization, a clipper 829 processes vertex data. The clipper 829 may be a fixed function clipper or a programmable clipper having clipping and geometry shader functions. In some embodiments, a rasterizer and depth test component 873 in the render output pipeline 870 dispatches pixel shaders to convert the geometric objects into their per pixel representations. In some embodiments, pixel shader logic is included in thread execution logic 850. In some embodiments, an application can bypass the rasterizer 873 and access un-rasterized vertex data via a stream out unit 823.

[0080] The graphics processor 800 has an interconnect bus, interconnect fabric, or some other interconnect mechanism that allows data and message passing amongst the major components of the processor. In some embodiments, execution units 852A, 852B and associated cache(s) 851, texture and media sampler 854, and texture/sampler cache 858 interconnect via a data port 856 to perform memory access and communicate with render output pipeline

components of the processor. In some embodiments, sampler 854, caches 851, 858 and execution units 852A, 852B each have separate memory access paths.

[0081] In some embodiments, render output pipeline 870 contains a rasterizer and depth test component 873 that converts vertex-based objects into an associated pixel-based representation. In some embodiments, the rasterizer logic includes a windower/masker unit to perform fixed function triangle and line rasterization. An associated render cache 878 and depth cache 879 are also available in some embodiments. A pixel operations component 877 performs pixel-based operations on the data, though in some instances, pixel operations associated with 2D operations (e.g. bit block image transfers with blending) are performed by the 2D engine 841, or substituted at display time by the display controller 843 using overlay display planes. In some embodiments, a shared L3 cache 875 is available to all graphics components, allowing the sharing of data without the use of main system memory.

[0082] In some embodiments, graphics processor media pipeline 830 includes a media engine 837 and a video front end 834. In some embodiments, video front end 834 receives pipeline commands from the command streamer 803. In some embodiments, media pipeline 830 includes a separate command streamer. In some embodiments, video front-end 834 processes media commands before sending the command to the media engine 837. In some embodiments, media engine 337 includes thread spawning functionality to spawn threads for dispatch to thread execution logic 850 via thread dispatcher 831.

[0083] In some embodiments, graphics processor 800 includes a display engine 840. In some embodiments, display engine 840 is external to processor 800 and couples with the graphics processor via the ring interconnect 802, or

some other interconnect bus or fabric. In some embodiments, display engine 840 includes a 2D engine 841 and a display controller 843. In some embodiments, display engine 840 contains special purpose logic capable of operating independently of the 3D pipeline. In some embodiments, display controller 843 couples with a display device (not shown), which may be a system integrated display device, as in a laptop computer, or an external display device attached via a display device connector.

[0084] In some embodiments, graphics pipeline 820 and media pipeline 830 are configurable to perform operations based on multiple graphics and media programming interfaces and are not specific to any one application programming interface (API). In some embodiments, driver software for the graphics processor translates API calls that are specific to a particular graphics or media library into commands that can be processed by the graphics processor. In some embodiments, support is provided for the Open Graphics Library (OpenGL) and Open Computing Language (OpenCL) from the Khronos Group, the Direct3D library from the Microsoft Corporation, or support may be provided to both OpenGL and D3D. Support may also be provided for the Open Source Computer Vision Library (OpenCV). A future API with a compatible 3D pipeline would also be supported if a mapping can be made from the pipeline of the future API to the pipeline of the graphics processor.

[0085] **Graphics Pipeline Programming**

[0086] **Figure 9A** is a block diagram illustrating a graphics processor command format 900 according to some embodiments. **Figure 9B** is a block diagram illustrating a graphics processor command sequence 910 according to an embodiment. The solid lined boxes in **Figure 9A** illustrate the components

that are generally included in a graphics command while the dashed lines include components that are optional or that are only included in a sub-set of the graphics commands. The exemplary graphics processor command format 900 of **Figure 9A** includes data fields to identify a target client 902 of the command, a command operation code (opcode) 904, and the relevant data 906 for the command. A sub-opcode 905 and a command size 908 are also included in some commands.

[0087] In some embodiments, client 902 specifies the client unit of the graphics device that processes the command data. In some embodiments, a graphics processor command parser examines the client field of each command to condition the further processing of the command and route the command data to the appropriate client unit. In some embodiments, the graphics processor client units include a memory interface unit, a render unit, a 2D unit, a 3D unit, and a media unit. Each client unit has a corresponding processing pipeline that processes the commands. Once the command is received by the client unit, the client unit reads the opcode 904 and, if present, sub-opcode 905 to determine the operation to perform. The client unit performs the command using information in data field 906. For some commands an explicit command size 908 is expected to specify the size of the command. In some embodiments, the command parser automatically determines the size of at least some of the commands based on the command opcode. In some embodiments commands are aligned via multiples of a double word.

[0088] The flow diagram in **Figure 9B** shows an exemplary graphics processor command sequence 910. In some embodiments, software or firmware of a data processing system that features an embodiment of a graphics processor uses a version of the command sequence shown to set up, execute,

and terminate a set of graphics operations. A sample command sequence is shown and described for purposes of example only as embodiments are not limited to these specific commands or to this command sequence. Moreover, the commands may be issued as batch of commands in a command sequence, such that the graphics processor will process the sequence of commands in at least partially concurrence.

[0089] In some embodiments, the graphics processor command sequence 910 may begin with a pipeline flush command 912 to cause any active graphics pipeline to complete the currently pending commands for the pipeline. In some embodiments, the 3D pipeline 922 and the media pipeline 924 do not operate concurrently. The pipeline flush is performed to cause the active graphics pipeline to complete any pending commands. In response to a pipeline flush, the command parser for the graphics processor will pause command processing until the active drawing engines complete pending operations and the relevant read caches are invalidated. Optionally, any data in the render cache that is marked 'dirty' can be flushed to memory. In some embodiments, pipeline flush command 912 can be used for pipeline synchronization or before placing the graphics processor into a low power state.

[0090] In some embodiments, a pipeline select command 913 is used when a command sequence requires the graphics processor to explicitly switch between pipelines. In some embodiments, a pipeline select command 913 is required only once within an execution context before issuing pipeline commands unless the context is to issue commands for both pipelines. In some embodiments, a pipeline flush command is 912 is required immediately before a pipeline switch via the pipeline select command 913.

[0091] In some embodiments, a pipeline control command 914 configures a graphics pipeline for operation and is used to program the 3D pipeline 922 and the media pipeline 924. In some embodiments, pipeline control command 914 configures the pipeline state for the active pipeline. In one embodiment, the pipeline control command 914 is used for pipeline synchronization and to clear data from one or more cache memories within the active pipeline before processing a batch of commands.

[0092] In some embodiments, return buffer state commands 916 are used to configure a set of return buffers for the respective pipelines to write data. Some pipeline operations require the allocation, selection, or configuration of one or more return buffers into which the operations write intermediate data during processing. In some embodiments, the graphics processor also uses one or more return buffers to store output data and to perform cross thread communication. In some embodiments, the return buffer state 916 includes selecting the size and number of return buffers to use for a set of pipeline operations.

[0093] The remaining commands in the command sequence differ based on the active pipeline for operations. Based on a pipeline determination 920, the command sequence is tailored to the 3D pipeline 922 beginning with the 3D pipeline state 930, or the media pipeline 924 beginning at the media pipeline state 940.

[0094] The commands for the 3D pipeline state 930 include 3D state setting commands for vertex buffer state, vertex element state, constant color state, depth buffer state, and other state variables that are to be configured before 3D primitive commands are processed. The values of these commands

are determined at least in part based the particular 3D API in use. In some embodiments, 3D pipeline state 930 commands are also able to selectively disable or bypass certain pipeline elements if those elements will not be used.

[0095] In some embodiments, 3D primitive 932 command is used to submit 3D primitives to be processed by the 3D pipeline. Commands and associated parameters that are passed to the graphics processor via the 3D primitive 932 command are forwarded to the vertex fetch function in the graphics pipeline. The vertex fetch function uses the 3D primitive 932 command data to generate vertex data structures. The vertex data structures are stored in one or more return buffers. In some embodiments, 3D primitive 932 command is used to perform vertex operations on 3D primitives via vertex shaders. To process vertex shaders, 3D pipeline 922 dispatches shader execution threads to graphics processor execution units.

[0096] In some embodiments, 3D pipeline 922 is triggered via an execute 934 command or event. In some embodiments, a register write triggers command execution. In some embodiments execution is triggered via a 'go' or 'kick' command in the command sequence. In one embodiment command execution is triggered using a pipeline synchronization command to flush the command sequence through the graphics pipeline. The 3D pipeline will perform geometry processing for the 3D primitives. Once operations are complete, the resulting geometric objects are rasterized and the pixel engine colors the resulting pixels. Additional commands to control pixel shading and pixel back end operations may also be included for those operations.

[0097] In some embodiments, the graphics processor command sequence 910 follows the media pipeline 924 path when performing media operations. In

general, the specific use and manner of programming for the media pipeline 924 depends on the media or compute operations to be performed. Specific media decode operations may be offloaded to the media pipeline during media decode. In some embodiments, the media pipeline can also be bypassed and media decode can be performed in whole or in part using resources provided by one or more general purpose processing cores. In one embodiment, the media pipeline also includes elements for general-purpose graphics processor unit (GPGPU) operations, where the graphics processor is used to perform SIMD vector operations using computational shader programs that are not explicitly related to the rendering of graphics primitives.

[0098] In some embodiments, media pipeline 924 is configured in a similar manner as the 3D pipeline 922. A set of media pipeline state commands 940 are dispatched or placed into in a command queue before the media object commands 942. In some embodiments, media pipeline state commands 940 include data to configure the media pipeline elements that will be used to process the media objects. This includes data to configure the video decode and video encode logic within the media pipeline, such as encode or decode format. In some embodiments, media pipeline state commands 940 also support the use one or more pointers to “indirect” state elements that contain a batch of state settings.

[0099] In some embodiments, media object commands 942 supply pointers to media objects for processing by the media pipeline. The media objects include memory buffers containing video data to be processed. In some embodiments, all media pipeline states must be valid before issuing a media object command 942. Once the pipeline state is configured and media object commands 942 are queued, the media pipeline 924 is triggered via an execute

command 944 or an equivalent execute event (e.g., register write). Output from media pipeline 924 may then be post processed by operations provided by the 3D pipeline 922 or the media pipeline 924. In some embodiments, GPGPU operations are configured and executed in a similar manner as media operations.

[00100] **Graphics Software Architecture**

[00101] **Figure 10** illustrates exemplary graphics software architecture for a data processing system 1000 according to some embodiments. In some embodiments, software architecture includes a 3D graphics application 1010, an operating system 1020, and at least one processor 1030. In some embodiments, processor 1030 includes a graphics processor 1032 and one or more general-purpose processor core(s) 1034. The graphics application 1010 and operating system 1020 each execute in the system memory 1050 of the data processing system.

[00102] In some embodiments, 3D graphics application 1010 contains one or more shader programs including shader instructions 1012. The shader language instructions may be in a high-level shader language, such as the High Level Shader Language (HLSL) or the OpenGL Shader Language (GLSL). The application also includes executable instructions 1014 in a machine language suitable for execution by the general-purpose processor core 1034. The application also includes graphics objects 1016 defined by vertex data.

[00103] In some embodiments, operating system 1020 is a Microsoft® Windows® operating system from the Microsoft Corporation, a proprietary UNIX-like operating system, or an open source UNIX-like operating system using a variant of the Linux kernel. When the Direct3D API is in use, the operating system 1020 uses a front-end shader compiler 1024 to compile any shader

instructions 1012 in HLSL into a lower-level shader language. The compilation may be a just-in-time (JIT) compilation or the application can perform shader pre-compilation. In some embodiments, high-level shaders are compiled into low-level shaders during the compilation of the 3D graphics application 1010.

[00104] In some embodiments, user mode graphics driver 1026 contains a back-end shader compiler 1027 to convert the shader instructions 1012 into a hardware specific representation. When the OpenGL API is in use, shader instructions 1012 in the GLSL high-level language are passed to a user mode graphics driver 1026 for compilation. In some embodiments, user mode graphics driver 1026 uses operating system kernel mode functions 1028 to communicate with a kernel mode graphics driver 1029. In some embodiments, kernel mode graphics driver 1029 communicates with graphics processor 1032 to dispatch commands and instructions.

[00105] **IP Core Implementations**

[00106] One or more aspects of at least one embodiment may be implemented by representative code stored on a machine-readable medium which represents and/or defines logic within an integrated circuit such as a processor. For example, the machine-readable medium may include instructions which represent various logic within the processor. When read by a machine, the instructions may cause the machine to fabricate the logic to perform the techniques described herein. Such representations, known as "IP cores," are reusable units of logic for an integrated circuit that may be stored on a tangible, machine-readable medium as a hardware model that describes the structure of the integrated circuit. The hardware model may be supplied to various customers or manufacturing facilities, which load the hardware model on fabrication machines that manufacture the integrated circuit. The integrated

circuit may be fabricated such that the circuit performs operations described in association with any of the embodiments described herein.

[00107] **Figure 11** is a block diagram illustrating an IP core development system 1100 that may be used to manufacture an integrated circuit to perform operations according to an embodiment. The IP core development system 1100 may be used to generate modular, re-usable designs that can be incorporated into a larger design or used to construct an entire integrated circuit (e.g., an SOC integrated circuit). A design facility 1130 can generate a software simulation 1110 of an IP core design in a high level programming language (e.g., C/C++). The software simulation 1110 can be used to design, test, and verify the behavior of the IP core. A register transfer level (RTL) design can then be created or synthesized from the simulation model 1100. The RTL design 1115 is an abstraction of the behavior of the integrated circuit that models the flow of digital signals between hardware registers, including the associated logic performed using the modeled digital signals. In addition to an RTL design 1115, lower-level designs at the logic level or transistor level may also be created, designed, or synthesized. Thus, the particular details of the initial design and simulation may vary.

[00108] The RTL design 1115 or equivalent may be further synthesized by the design facility into a hardware model 1120, which may be in a hardware description language (HDL), or some other representation of physical design data. The HDL may be further simulated or tested to verify the IP core design. The IP core design can be stored for delivery to a 3rd party fabrication facility 1165 using non-volatile memory 1140 (e.g., hard disk, flash memory, or any non-volatile storage medium). Alternatively, the IP core design may be transmitted (e.g., via the Internet) over a wired connection 1150 or wireless connection 1160. The fabrication facility 1165 may then fabricate an integrated circuit that is based

at least in part on the IP core design. The fabricated integrated circuit can be configured to perform operations in accordance with at least one embodiment described herein.

[00109] **Figure 12** is a block diagram illustrating an exemplary system on a chip integrated circuit 1200 that may be fabricated using one or more IP cores, according to an embodiment. The exemplary integrated circuit includes one or more application processors 1205 (e.g., CPUs), at least one graphics processor 1210, and may additionally include an image processor 1215 and/or a video processor 1220, any of which may be a modular IP core from the same or multiple different design facilities. The integrated circuit includes peripheral or bus logic including a USB controller 1225, UART controller 1230, an SPI/SDIO controller 1235, and an I^2S/I^2C controller 1240. Additionally, the integrated circuit can include a display device 1245 coupled to one or more of a high-definition multimedia interface (HDMI) controller 1250 and a mobile industry processor interface (MIPI) display interface 1255. Storage may be provided by a flash memory subsystem 1260 including flash memory and a flash memory controller. Memory interface may be provided via a memory controller 1265 for access to SDRAM or SRAM memory devices. Some integrated circuits additionally include an embedded security engine 1270.

[00110] Additionally, other logic and circuits may be included in the processor of integrated circuit 1200, including additional graphics processors/cores, peripheral interface controllers, or general purpose processor cores.

APPARATUS AND METHOD PERIODIC SNAPSHOTTING IN A GRAPHICS PROCESSING ENVIRONMENT

[00111] As mentioned, Graphics Processing Unit (GPU) computing has been widely used for a variety of applications including 3D rendering, media transcoding, high performance computing, real time image processing, and packet filtering, to name a few. However, due to the complexity associated with GPU microarchitecture and programming GPU software may crash more frequently than software designed for general purpose CPUs. Existing approaches solve the problem by restarting the software and/or relying on Timeout Detection and Recovery (TDR) to reset the software. However these solutions are insufficient since they are incapable of determining the root cause of the problem. Fine-grain debuggers which may be used to perform GPU program debugging are also insufficient. For example, if a bug is encountered after long period of execution (e.g., a Mean Time Between Failures of 20 hours) it may be impossible to determine the cause. As such, fine-grain debuggers are impractical for comprehensive, hard-reproducible bugs.

[00112] To address these limitations, one embodiment of the invention comprises a mediated command replay framework, which includes a thin hypervisor to snapshot the GPU context (including I/O and graphics memory state) at the boundary of command buffer submission. When a GPU hang condition occurs, the GPU context can be rolled back to the previous snapshot, referred to as a “replay.” This solution is transparent to the application and the operating system (OS), so it can be applied to massive GPU-based applications which are particularly important to GPU cloud.

[00113] A whole system snapshot would be useful but it would be challenging to manage all of the device states in the system, and could result in a massive amount of data. Moreover, PC devices may be unable to support the

rollback of their state in practice. Consequently, one embodiment of the invention implements niche targeted memory snapshots. This embodiment performs an on-demand snapshot of the GPU context only, and rolls back the GPU context to the selected previous snapshot, without the context of the CPU and other system components (e.g., I/O components).

[00114] Additionally, one embodiment implements a delta snapshot for both GPU memory and IO registers to minimize the snapshot size. The delta snapshot may take the difference between the context at an earlier time and the context at a later time rather than storing all of the actual context data at each time.

[00115] Moreover, one embodiment may decouple the execution engines such as the CPU and the rendering engine for speculative execution. Using these techniques, when the GPU is executing between two snapshots, the CPUs can execute in parallel as if they don't touch the state (which may potentially be clobbered by the GPU).

[00116] **Figure 13** illustrates an architecture in accordance with one embodiment of the invention which includes a thin hypervisor layer 1340 inserted between the original OS 1350 and the underlying platform which includes GPU hardware 1336 and one or more other devices 1335 (e.g., a CPU, I/O resources, memory, etc). A plurality of CPU applications 1320 are shown using device drivers 1330 and a plurality of GPU applications 1360-1362 are shown using GPU drivers 1352. In one embodiment, the thin hypervisor 1340 passes through all of the other devices 1335 to the OS 1350, other than privileged operations (e.g., I/O accesses and command submissions) directed to the GPU device

1336. Performance-critical GPU resources such as graphics memory accesses of the GPU may also be passed through

[00117] One embodiment of the thin hypervisor includes three components: an extended page table (EPT) 1341, a GPU I/O state tracker 1342, and a command parser 1343. In one embodiment, the EPT 1341 comprises a module used to add/remove mappings to the clobbered memory pages. As used herein, "clobbered" memory pages means pages which are dirtied or overwritten by the GPU. In one embodiment, it is utilized to manage parallel CPU/GPU accesses in these rare use cases.

[00118] In one embodiment, the I/O state tracker 1342 traps I/O accesses from the graphics drivers 1352 (i.e., user mode drivers (UMDs) and/or kernel mode driver (KMDs)) to keep track of the OS 1350 view of the I/O state of the GPU 1336.

[00119] One embodiment of the command parser 1343 parses the commands received from the GPU drivers 1352, parse the command, and generates a list of graphics memory pages which will be invalidated/clobbered by the commands per the result of command parser. This may be done, for example, at the time the commands are issued to the hardware GPU 1336. Note here that the GPU commands may be issued in batches. Thus, the context/state snapshot may be taken each time a batch of commands is submitted to the GPU 1336 and/or or after multiple batches are submitted.

[00120] As mentioned, in one embodiment, the GPU I/O snapshot can be tracked by the GPU I/O state tracker 1342 either fully or incrementally. It may intercept an I/O snapshot using the path of I/O register operation and tracks the

state of devices per semantics (e.g., using a device model similar to a virtual GPU (vGPU)). In one embodiment, the I/O snapshot is made by reading out all the vGPU registers. Alternatively, it may read out a set of registers the driver 1352 may touch, and protect the rest of the registers from GPU access by scanning the GPU commands received by the command parser 1343. In another embodiment, the latest state is maintained by capturing the incremental changes according to trapped I/O accesses and parsed commands (if the command execution may modify the I/O state).

[00121] The Graphics Memory snapshot can be made completely or incrementally as well. In one embodiment, the entire graphics memory is considered invalidated/clobbered for simplicity. In another embodiment, the snapshot is limited to memory pages which are actually invalidated/clobbered by the execution of the command buffers.

[00122] In one embodiment, the guest issued command may contain a flush command at the end, so the GPU context (cache) is flushed at the completion of the command execution in the last submission (i.e., at the TAIL register write), so the new memory snapshot is ensured to contain the latest data. However, not every guest command buffer includes a context flush operation. In one embodiment, the thin hypervisor 1340 may temporarily switch to its own command buffer to force the context flush. In another embodiment, the driver 1352 may be changed to force the flush. In yet another embodiment, the thin hypervisor 1340 may construct a shadow command buffer, into which it inserts the context flush commands.

[00123] Command buffers with mutual-dependency (e.g., through semaphore primitives) may be submitted and retired together. Making a snapshot between them may cause a restore operation to fail (e.g. by waiting for a semaphore released from an already-retired command buffer). In one embodiment, mutual-dependency can be detected by parsing the command buffers. In another embodiment, mutual-dependency hints may be provided by the driver.

[00124] In one embodiment, GPU state is snapshotted periodically (e.g., every 100ms), which means that the command buffers are buffered and submitted in a batch periodically. In one embodiment, the guest driver may be modified to submit the command buffers periodically. In another embodiment, even the guest driver is not submitting command buffer regularly, the command parser 1343 may be able to have a buffering component to temporary buffer the guest submitted command, and submit to the rest of 1343 periodically later on. All going-to-be-submitted-to-hardware command buffers are parsed to generate the invalidated register list and invalidated memory list. That is, the invalidated memory list determines which pages should be snapshotted, at the time when command buffers are submitted to real hardware. The invalidated register list determines which registers in the GPU I/O snapshot should be incrementally updated, at the time when command buffers are completed.

[00125] In one embodiment, the GPU memory snapshot and I/O snapshot is made just prior to submission of specified command buffers, based on the invalidated memory list generated by the command parser 1343. When all the command buffers are completed, the I/O snapshot is updated to catch the latest device state.

[00126] In another embodiment, a newly arriving command buffer may be submitted to GPU immediately, if to-be-dirtied/clobbered memory/register list of this new command buffer doesn't overlap with the dirty list of in-the-fly executing command buffers. It means the GPU state touched by new command buffer hasn't been snapshotted yet so it's safe to incrementally update the existing snapshot w/o waiting for completion of all in-the-fly command buffers (adding the newly-to-be-dirtied/clobbered memory pages to existing snapshot at the time new submission is made, and then reducing the memory snapshot and updating the I/O snapshot when this command buffer is completed).

[00127] After the snapshot (i.e. during the execution of the GPU), the CPU can keep running to maximize performance, until the CPU touches the shared memory resources, which may impact the GPU context. One embodiment modifies the hypervisor extended page table (EPT) 1341 to remove the mapping of those snapshotted pages, so that in the unlikely event that the CPU accesses this shared resource, the hypervisor 1340 can trap and temporarily pause the execution of the CPU for correctness, until the completion of GPU execution. Similarly, CPU access to GPU I/O registers may also cause the CPU to pause, at the time the command buffer is executing.

[00128] In one embodiment, once a new snapshot is taken, the waiting CPUs are released. Alternatively, if the GPU fails, the GPU context can be restored to the previous snapshot, and the GPU can re-execute with the CPUs waiting.

[00129] A method in accordance with one embodiment of the invention is illustrated in **Figure 14**. The method may be implemented within the context of

the system architectures described above, but is not limited to any particular system architecture.

[00130] At 1401, a plurality of graphics commands are received which are to be executed by a graphics processing unit (GPU). As mentioned, the graphics commands may be received in batches from a GPU driver. At 1402, the graphics commands are parsed to determine a list of graphics memory pages which will be affected by the graphics commands (e.g., which memory pages which will be updated in response to the execution of the graphics commands). At 1403, I/O accesses from the GPU driver are tracked to determine a list of registers affected by the I/O accesses. For example, I/O accesses from the graphics driver may be trapped by the hypervisor to keep track of the OS view of the GPU I/O state. At 1404, a memory snapshot and I/O snapshot are performed based on the list of graphics memory pages and the list of registers. As mentioned, the snapshots may be taken periodically (e.g., every 100ms) and/or in response to each new batch of GPU commands, and/or several batch of GPU commands base on the requirement of the frequency of snapshot, balancing the size of the snapshot and the impact of duration of replay. At 1405 a determination is made as to whether an error is detected with respect to the GPU such as a crash, hang or other type of fault condition. If so, then a rollback operation and command replay is performed at 1406 using the memory snapshot and the I/O snapshot. For example, in one embodiment, the GPU context is rolled back to the pervious memory snapshot and I/O snapshot

[00131] Because the embodiments of the invention described above are able to replay a partially-executed command buffer after a GPU hang in a

manner that is completely transparent to applications, these embodiments can meet the high availability requirements in specific GPU computing use cases.

[00132] The above examples may include specific combination of features. However, such the above examples are not limited in this regard and, in various implementations, the above examples may include the undertaking only a subset of such features, undertaking a different order of such features, undertaking a different combination of such features, and/or undertaking additional features than those features explicitly listed. For example, all features described with respect to the example methods may be implemented with respect to the example apparatus, the example systems, and/or the example articles, and vice versa.

[00133] Embodiments of the invention may include various steps, which have been described above. The steps may be embodied in machine-executable instructions which may be used to cause a general-purpose or special-purpose processor to perform the steps. Alternatively, these steps may be performed by specific hardware components that contain hardwired logic for performing the steps, or by any combination of programmed computer components and custom hardware components.

[00134] As described herein, instructions may refer to specific configurations of hardware such as application specific integrated circuits (ASICs) configured to perform certain operations or having a predetermined functionality or software instructions stored in memory embodied in a non-transitory computer readable medium. Thus, the techniques shown in the figures can be implemented using code and data stored and executed on one or more

electronic devices (e.g., an end station, a network element, etc.). Such electronic devices store and communicate (internally and/or with other electronic devices over a network) code and data using computer machine-readable media, such as non-transitory computer machine-readable storage media (e.g., magnetic disks; optical disks; random access memory; read only memory; flash memory devices; phase-change memory) and transitory computer machine-readable communication media (e.g., electrical, optical, acoustical or other form of propagated signals – such as carrier waves, infrared signals, digital signals, etc.). In addition, such electronic devices typically include a set of one or more processors coupled to one or more other components, such as one or more storage devices (non-transitory machine-readable storage media), user input/output devices (e.g., a keyboard, a touchscreen, and/or a display), and network connections. The coupling of the set of processors and other components is typically through one or more busses and bridges (also termed as bus controllers). The storage device and signals carrying the network traffic respectively represent one or more machine-readable storage media and machine-readable communication media. Thus, the storage device of a given electronic device typically stores code and/or data for execution on the set of one or more processors of that electronic device. Of course, one or more parts of an embodiment of the invention may be implemented using different combinations of software, firmware, and/or hardware. Throughout this detailed description, for the purposes of explanation, numerous specific details were set forth in order to provide a thorough understanding of the present invention. It will be apparent, however, to one skilled in the art that the invention may be practiced without some of these specific details. In certain instances, well known structures and functions were not described in elaborate detail in order to avoid obscuring the

subject matter of the present invention. Accordingly, the scope and spirit of the invention should be judged in terms of the claims which follow.

CLAIMS

What is claimed is:

1. An apparatus comprising:

a graphics processing unit (GPU) to perform graphics processing operations by executing graphics commands;

a command parser to parse graphics commands submitted to the GPU and generate a list of graphics memory pages which will be affected by the graphics commands;

an I/O state tracker to track I/O accesses from a graphics driver to determine a list of registers affected by the I/O accesses;

shapshot circuitry and/or logic to perform a memory snapshot and I/O snapshot based on the list of graphics memory pages and the list of registers, respectively; and

rollback circuitry and/or logic to perform a rollback operation using the memory snapshot and I/O snapshot in response to detecting a GPU error condition.

2. The apparatus as in claim 1 wherein performing the rollback operation resets the apparatus to a GPU context existing prior to the I/O accesses and graphics commands.

3. The apparatus as in claim 2 wherein following the rollback operation, one or more of the graphics commands are to be re-executed from the GPU context.

4. The apparatus as in claim 1 wherein the snapshot circuitry and/or logic is to perform the memory and I/O snapshots periodically.

5. The apparatus as in claim 1 wherein the snapshot circuitry and/or logic is to perform the memory and I/O snapshots in response to a detected submission of new graphics commands.

6. The apparatus as in claim 1 wherein the rollback circuitry and/or logic is to perform the rollback operation by flushing a current GPU context and resetting memory, cache, and registers in accordance with the memory snapshot and I/O snapshot.

7. The apparatus as in claim 1 further comprising:

a central processing unit (CPU) to execute program code and process data during and after the memory snapshot and I/O snapshot have been performed.

8. The apparatus as in claim 7 further comprising:

an extended page table (EPT) to be updated responsive to GPU operations following the memory snapshot and I/O snapshot, wherein the execution of the CPU is to be paused responsive to detecting a CPU access to an EPT entry which has been mofified by the GPU operations.

9. A method comprising:

executing graphics commands on a graphics processing unit (GPU) to perform graphics processing operations;

parsing the graphics commands submitted to the GPU and responsively generating a list of graphics memory pages which will be affected by the graphics commands;

tracking I/O accesses from a graphics driver to determine a list of registers affected by the I/O accesses;

performing a memory snapshot and an I/O snapshot based on the list of graphics memory pages and the list of registers, respectively; and

performing a rollback operation using the memory snapshot and I/O snapshot in response to detecting a GPU error condition.

10. The method as in claim 9 wherein performing the rollback operation resets the GPU to a GPU context existing prior to the I/O accesses and graphics commands.

11. The method as in claim 10 wherein following the rollback operation, one or more of the graphics commands are to be re-executed from the GPU context.

12. The method as in claim 9 wherein the snapshot circuitry and/or logic is to perform the memory and I/O snapshots periodically.

13. The method as in claim 9 wherein the snapshot circuitry and/or logic is to perform the memory and I/O snapshots in response to a detected submission of new graphics commands.

14. The method as in claim 9 wherein the rollback circuitry and/or logic is to perform the rollback operation by flushing a current GPU context and

resetting memory, cache, and registers in accordance with the memory snapshot and I/O snapshot.

15. The method as in claim 9 further comprising:

a central processing unit (CPU) to execute program code and process data during and after the memory snapshot and I/O snapshot have been performed.

16. The method as in claim 15 further comprising:

an extended page table (EPT) to be updated responsive to GPU operations following the memory snapshot and I/O snapshot, wherein the execution of the CPU is to be paused responsive to detecting a CPU access to an EPT entry which has been modified by the GPU operations.

17. A system comprising:

a memory to store program code and data;

a central processing unit (CPU) comprising an instruction cache for caching a portion of the program code and a data cache for caching a portion of the data, the CPU further comprising execution logic to execute at least some of the program code and responsively process at least some of the data, at least a portion of the program code comprising graphics commands;

a graphics processing subsystem to process the graphics commands and responsively render a plurality of images, the graphics processing subsystem comprising:

a graphics processing unit (GPU) to perform graphics processing operations by executing graphics commands;

a command parser to parse graphics commands submitted to the GPU and generate a list of graphics memory pages which will be affected by the graphics commands;

an I/O state tracker to track I/O accesses from a graphics driver to determine a list of registers affected by the I/O accesses;

shapshot circuitry and/or logic to perform a memory snapshot and I/O snapshot based on the list of graphics memory pages and the list of registers, respectively; and

rollback circuitry and/or logic to perform a rollback operation using the memory snapshot and I/O snapshot in response to detecting a GPU error condition.

18. The apparatus as in claim 17 wherein performing the rollback operation resets the apparatus to a GPU context existing prior to the I/O accesses and graphics commands.

19. The apparatus as in claim 18 wherein following the rollback operation, one or more of the graphics commands are to be re-executed from the GPU context.

20. The apparatus as in claim 17 wherein the snapshot circuitry and/or logic is to perform the memory and I/O snapshots periodically.

21. The apparatus as in claim 17 wherein the snapshot circuitry and/or logic is to perform the memory and I/O snapshots in response to a detected submission of new graphics commands.

22. The apparatus as in claim 17 wherein the rollback circuitry and/or logic is to perform the rollback operation by flushing a current GPU context and resetting memory, cache, and registers in accordance with the memory snapshot and I/O snapshot.

23. The apparatus as in claim 17 further comprising:

a central processing unit (CPU) to execute program code and process data during and after the memory snapshot and I/O snapshot have been performed.

24. The apparatus as in claim 23 further comprising:

an extended page table (EPT) to be updated responsive to GPU operations following the memory snapshot and I/O snapshot, wherein the execution of the CPU is to be paused responsive to detecting a CPU access to an EPT entry which has been modified by the GPU operations.

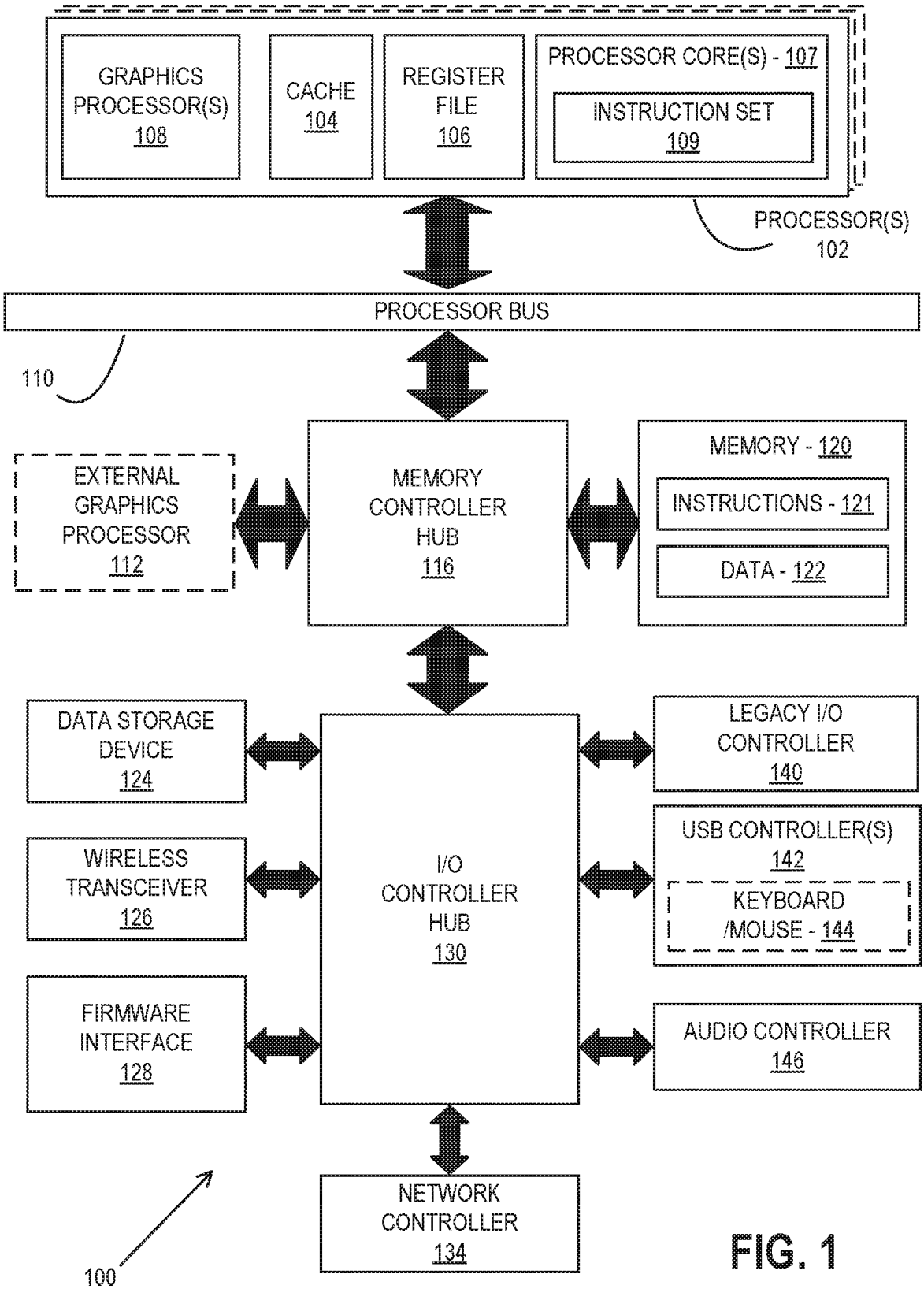

GRAPHICS PROCESSOR(S) 108
CACHE 104
REGISTER FILE 106
PROCESSOR CORE(S) - 107
INSTRUCTION SET 109
PROCESSOR(S) 102
PROCESSOR BUS
110
EXTERNAL GRAPHICS PROCESSOR 112
MEMORY CONTROLLER HUB 116
MEMORY - 120
INSTRUCTIONS - 121
DATA - 122
DATA STORAGE DEVICE 124
WIRELESS TRANSCEIVER 126
FIRMWARE INTERFACE 128
I/O CONTROLLER HUB 130
LEGACY I/O CONTROLLER 140
USB CONTROLLER(S) 142
KEYBOARD /MOUSE - 144
AUDIO CONTROLLER 146
NETWORK CONTROLLER 134
100

FIG. 1

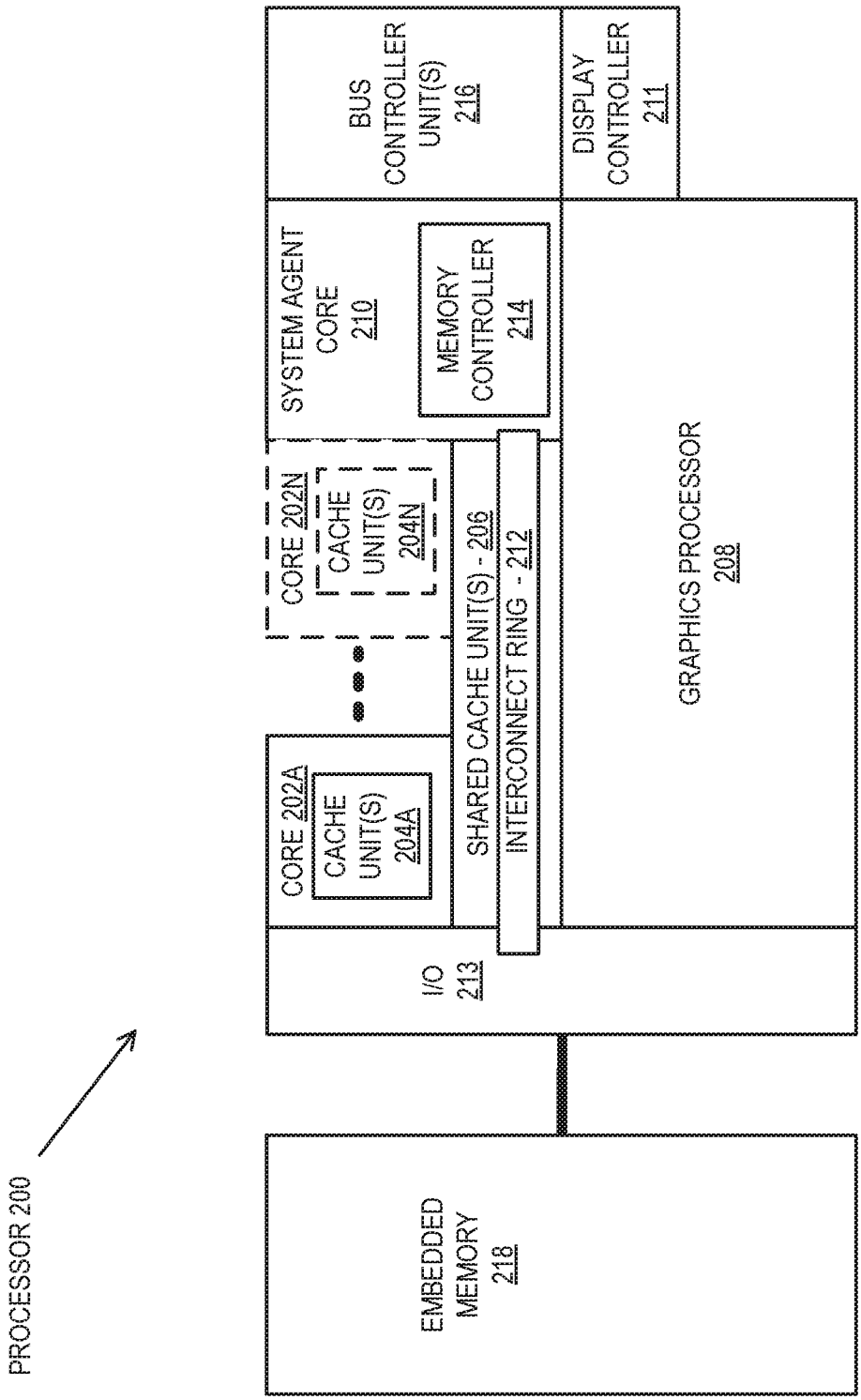

PROCESSOR 200
EMBEDDED MEMORY 218
I/O 213
CORE 202A
CACHE UNIT(S) 204A
CORE 202N
CACHE UNIT(S) 204N
SHARED CACHE UNIT(S) - 206
INTERCONNECT RING - 212
SYSTEM AGENT CORE 210
MEMORY CONTROLLER 214
BUS CONTROLLER UNIT(S) 216
DISPLAY CONTROLLER 211
GRAPHICS PROCESSOR 208

FIG. 2

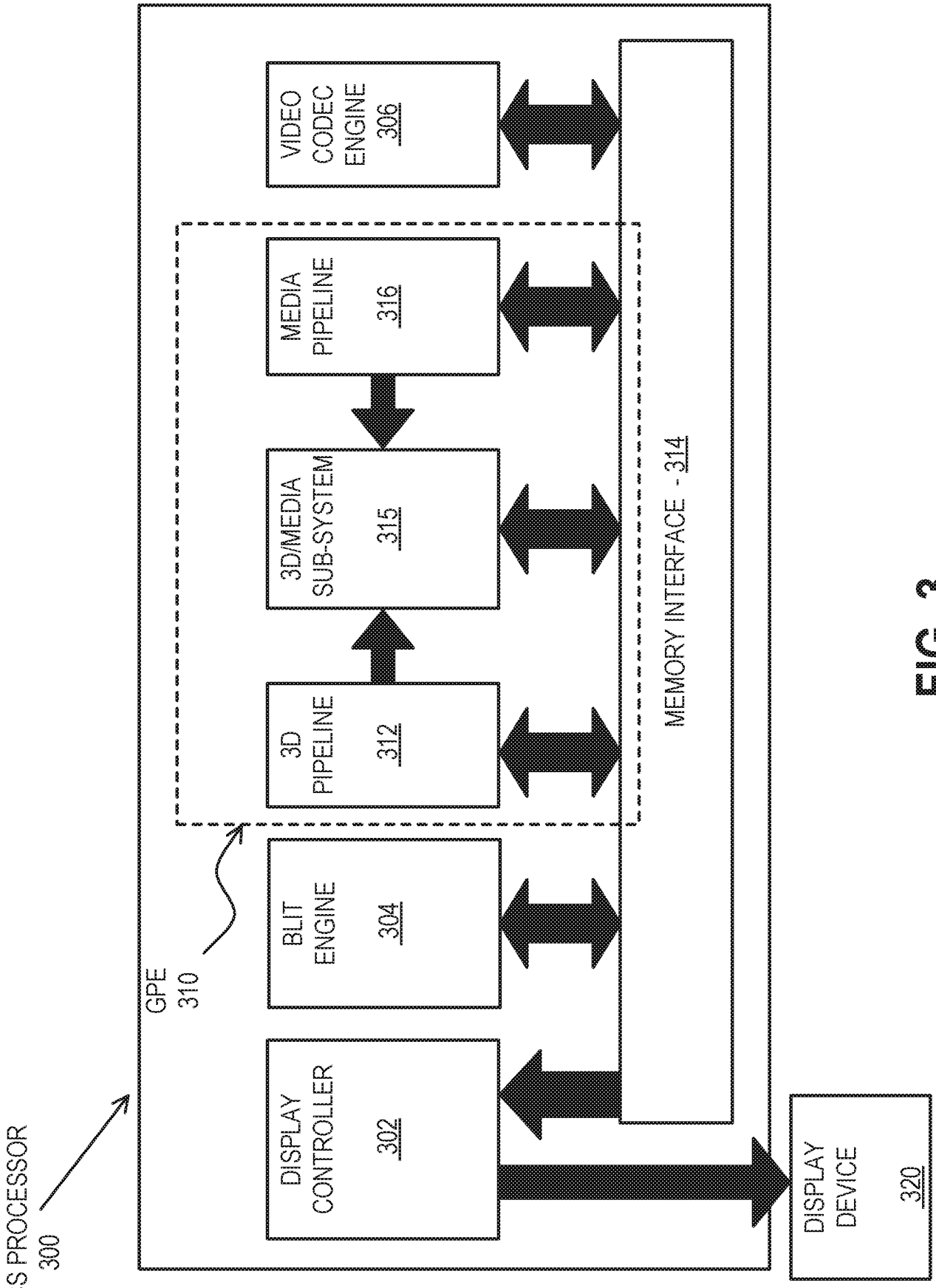

GRAPHICS PROCESSOR
300
GPE
310
DISPLAY CONTROLLER
302
BLIT ENGINE
304
3D PIPELINE
312
3D/MEDIA SUB-SYSTEM
315
MEDIA PIPELINE
316
VIDEO CODEC ENGINE
306
MEMORY INTERFACE - 314
DISPLAY DEVICE
320

FIG. 3

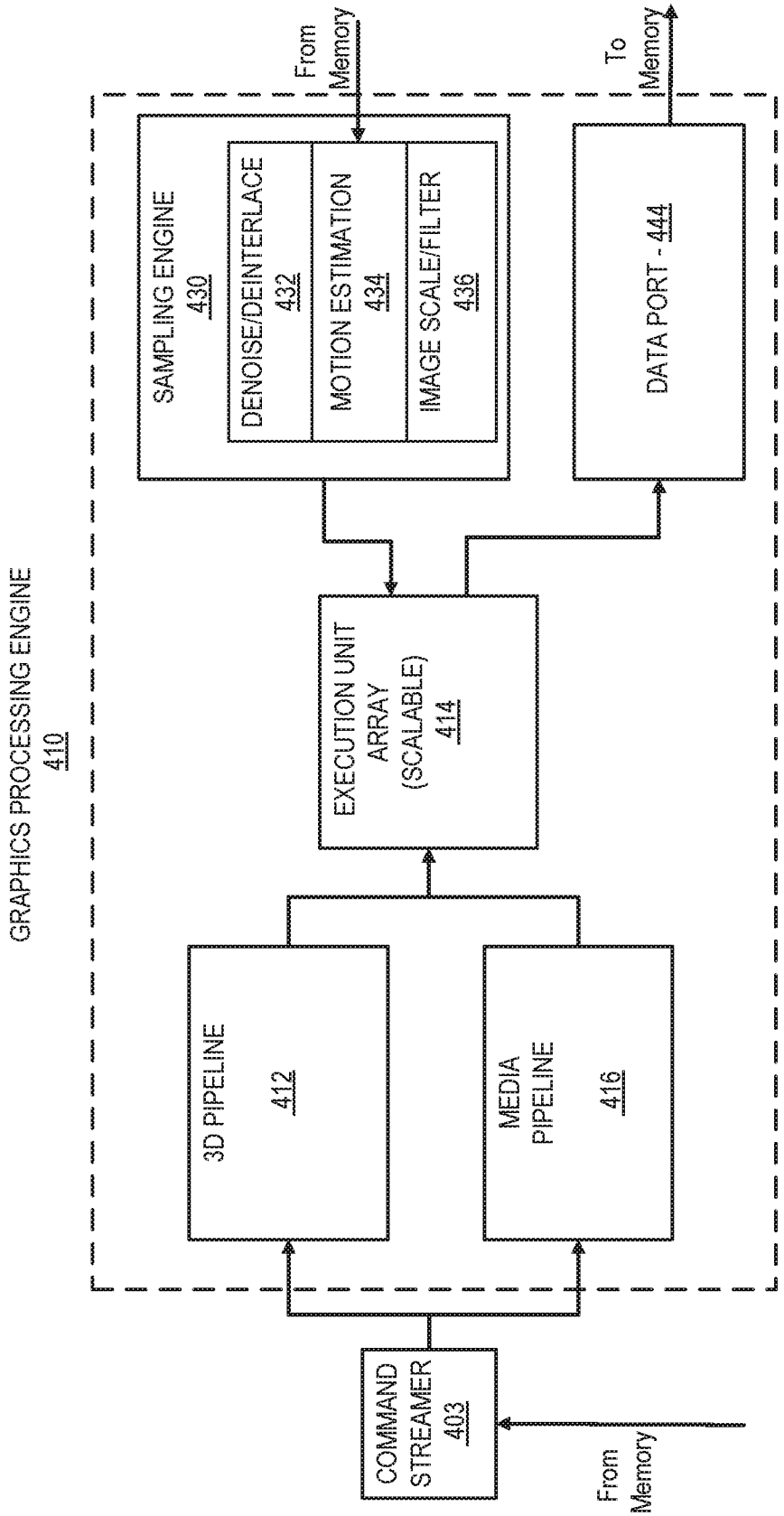

GRAPHICS PROCESSING ENGINE
410
From Memory
To Memory
SAMPLING ENGINE
430
DENOISE/DEINTERLACE
432
MOTION ESTIMATION
434
IMAGE SCALE/FILTER
436
DATA PORT - 444
EXECUTION UNIT ARRAY (SCALABLE)
414
3D PIPELINE
412
MEDIA PIPELINE
416
COMMAND STREAMER
403
From Memory

FIG. 4

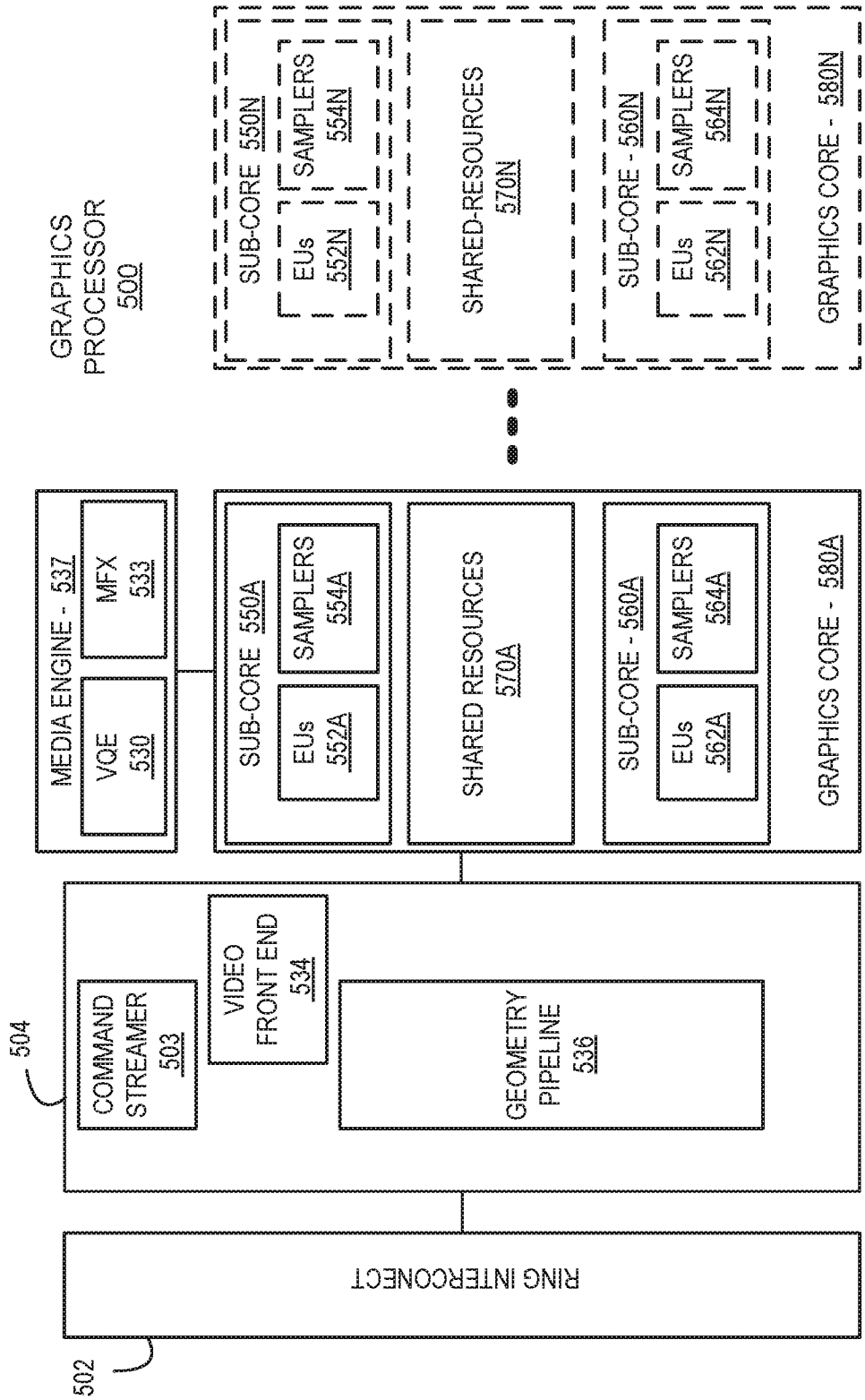

GRAPHICS PROCESSOR 500
MEDIA ENGINE - 537
VQE 530
MFX 533
SUB-CORE 550A
EUs 552A
SAMPLERS 554A
SHARED RESOURCES 570A
SUB-CORE - 560A
EUs 562A
SAMPLERS 564A
GRAPHICS CORE - 580A
SUB-CORE 550N
EUs 552N
SAMPLERS 554N
SHARED-RESOURCES 570N
SUB-CORE - 560N
EUs 562N
SAMPLERS 564N
GRAPHICS CORE - 580N
504
COMMAND STREAMER 503
VIDEO FRONT END 534
GEOMETRY PIPELINE 536
RING INTERCONNECT
502

FIG. 5

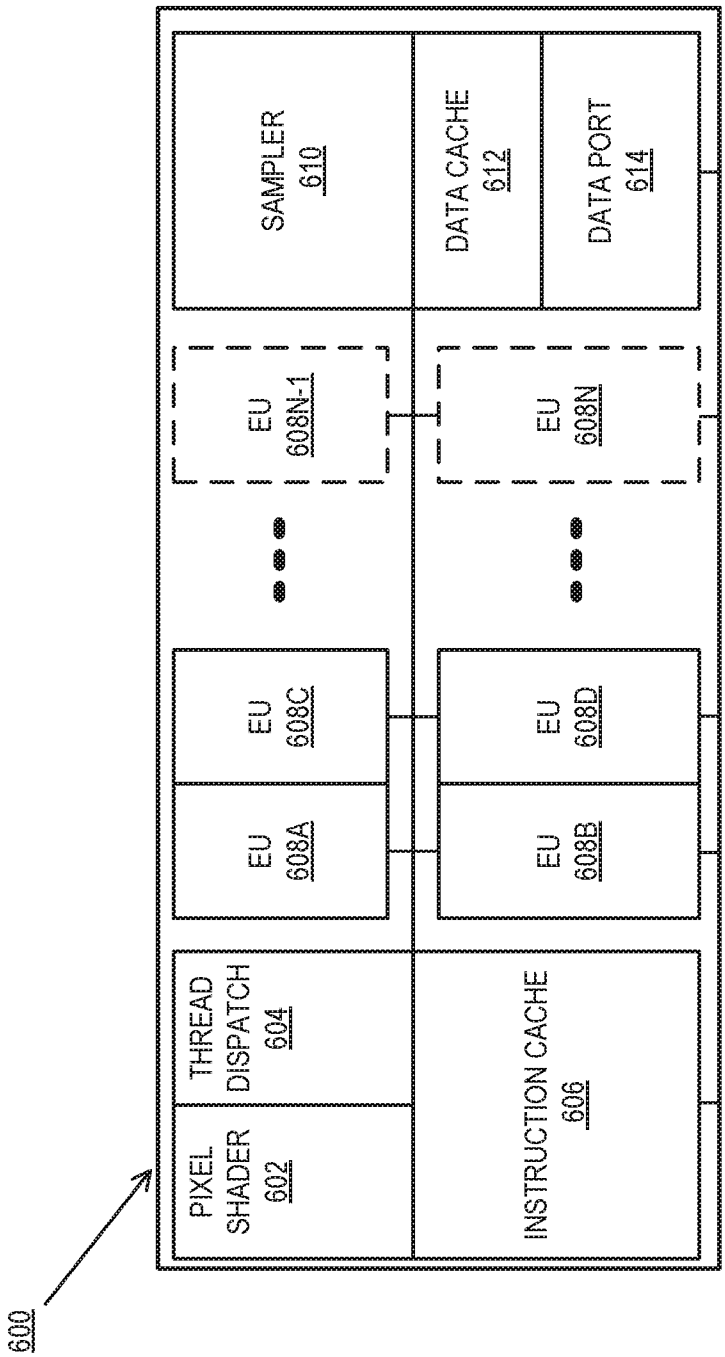

600
PIXEL SHADER 602
THREAD DISPATCH 604
INSTRUCTION CACHE 606
EU 608A
EU 608B
EU 608C
EU 608D
EU 608N-1
EU 608N
SAMPLER 610
DATA CACHE 612
DATA PORT 614

FIG. 6

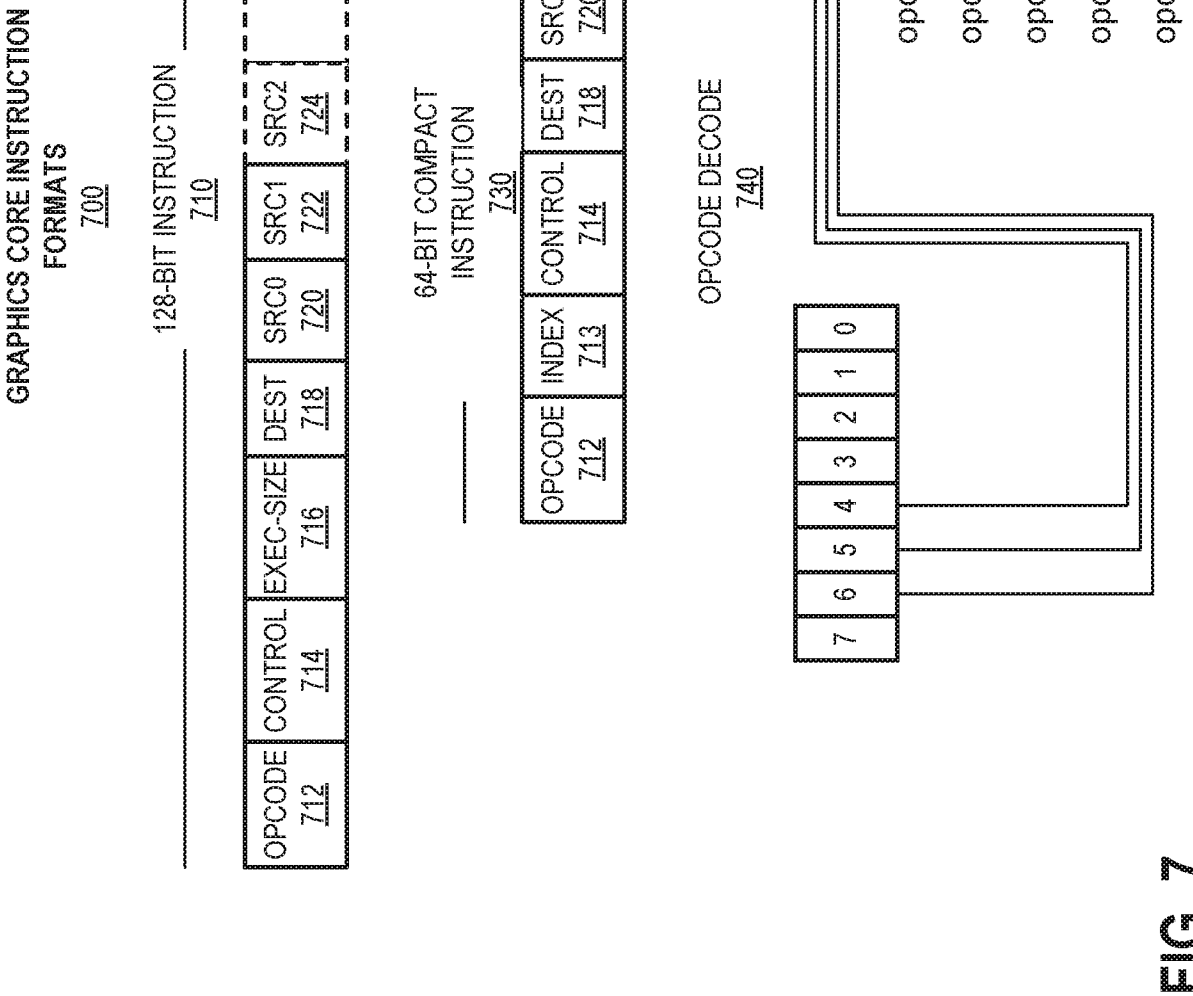

GRAPHICS CORE INSTRUCTION FORMATS
700
128-BIT INSTRUCTION
710
OPCODE 712
CONTROL 714
EXEC-SIZE 716
DEST 718
SRC0 720
SRC1 722
SRC2 724
ACCESS/ADDRESS MODE 726
64-BIT COMPACT INSTRUCTION
730
OPCODE 712
INDEX 713
CONTROL 714
DEST 718
SRC0 720
SRC1 722
OPCODE DECODE
740
7 6 5 4 3 2 1 0
opcode=0000xxxxb
opcode=0010xxxxb
opcode=0011xxxxb
opcode=0100xxxxb
opcode=0101xxxxb
Move/Logic - 742
Flow Control - 744
Miscellaneous - 746
Parallel Math - 748
Vector Math - 750

FIG. 7

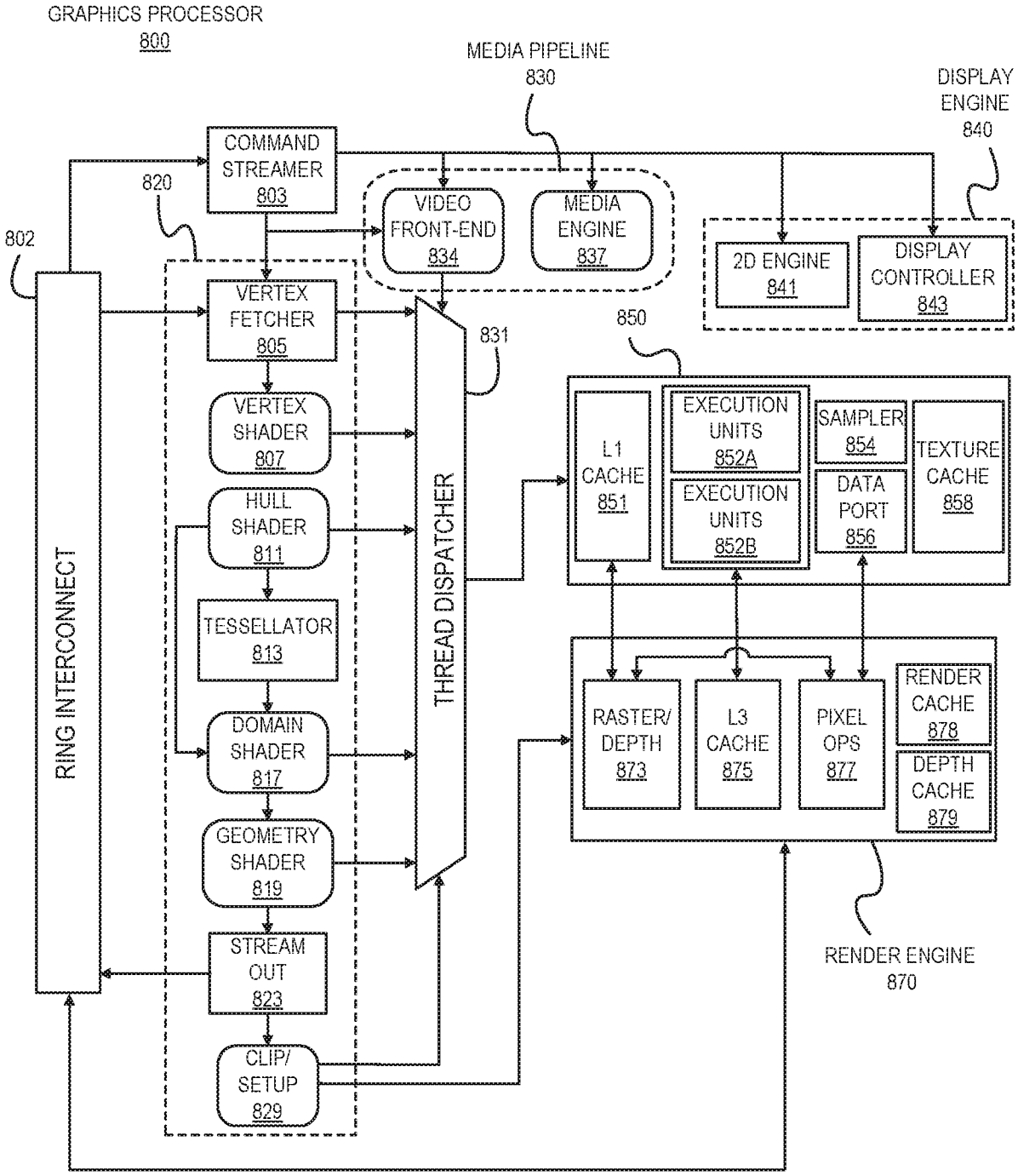
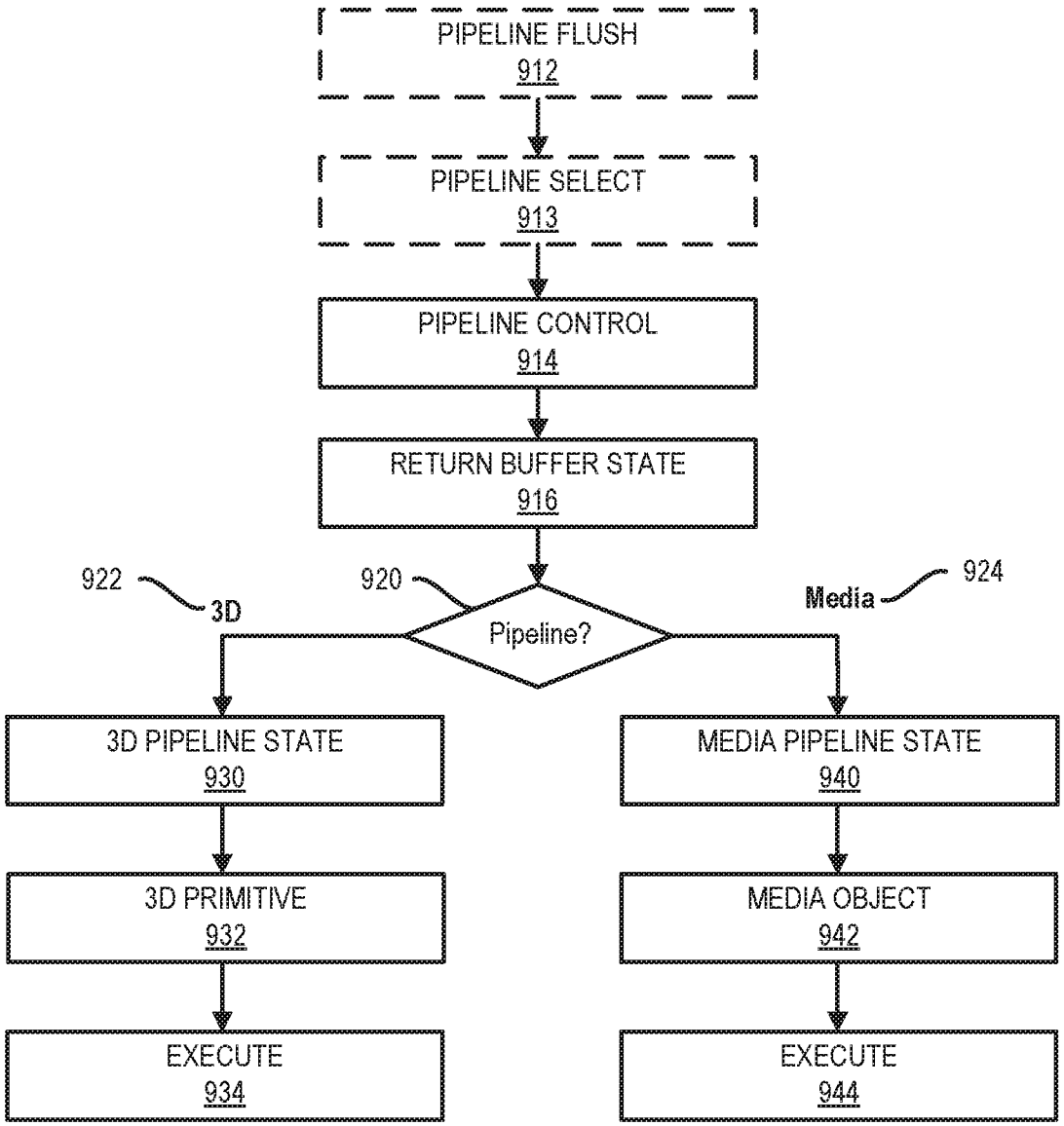

GRAPHICS PROCESSOR
600
MEDIA PIPELINE
830
DISPLAY ENGINE
840
COMMAND STREAMER
803
820
802
VIDEO FRONT-END
834
MEDIA ENGINE
837
2D ENGINE
841
DISPLAY CONTROLLER
843
VERTEX FETCHER
805
831
850
VERTEX SHADER
807
HULL SHADER
811
TESSELLATOR
813
DOMAIN SHADER
817
GEOMETRY SHADER
819
STREAM OUT
823
CLIP/ SETUP
829
RING INTERCONNECT
THREAD DISPATCHER
L1 CACHE
851
EXECUTION UNITS
852A
EXECUTION UNITS
852B
SAMPLER
854
DATA PORT
856
TEXTURE CACHE
858
RASTER/ DEPTH
873
L3 CACHE
875
PIXEL OPS
877
RENDER CACHE
878
DEPTH CACHE
879
RENDER ENGINE
870

FIG. 8

FIG. 9A GRAPHICS PROCESSOR COMMAND FORMAT
900

FIG. 9B GRAPHICS PROCESSOR COMMAND SEQUENCE
910



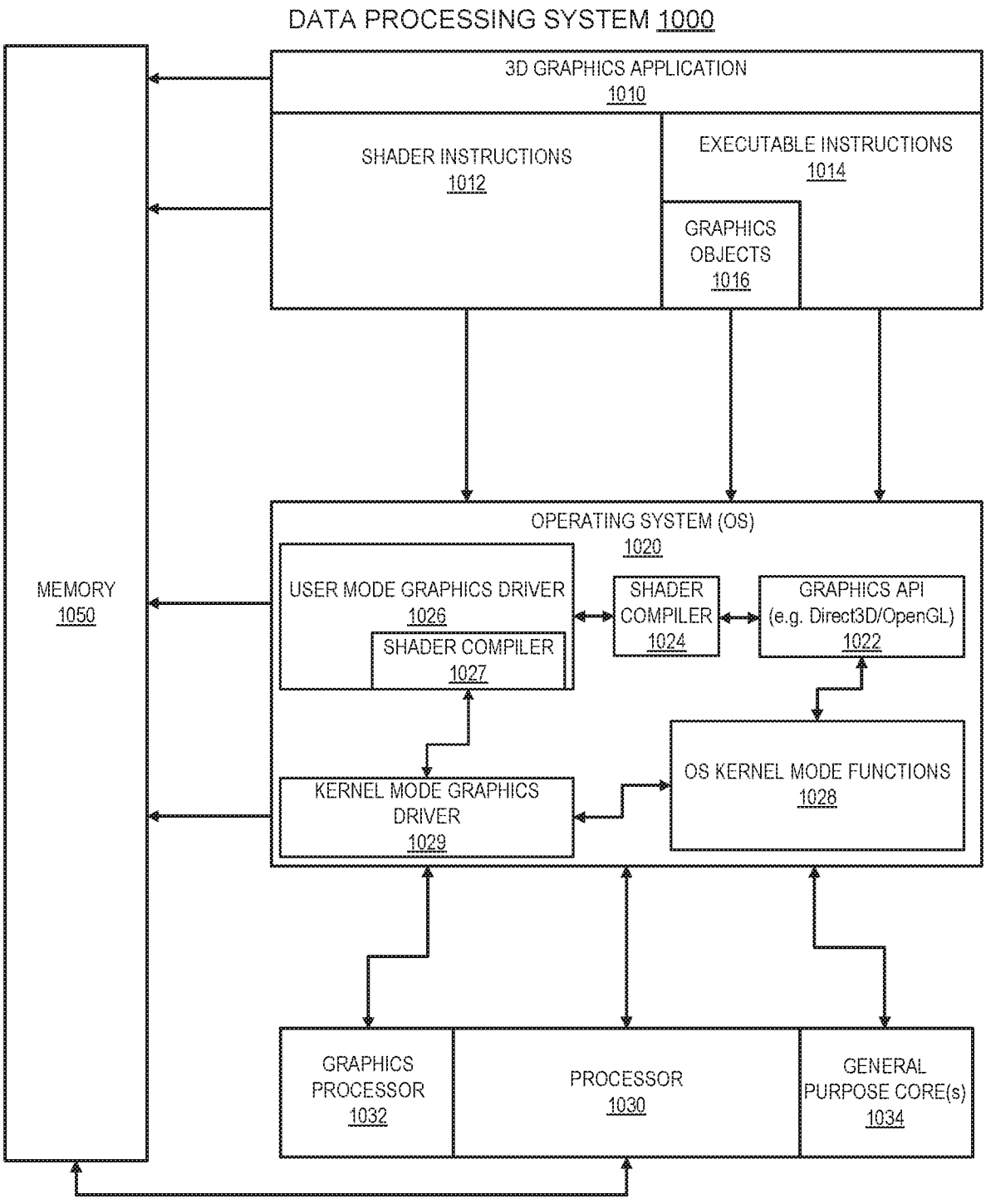

DATA PROCESSING SYSTEM 1000
3D GRAPHICS APPLICATION
1010
SHADER INSTRUCTIONS
1012
EXECUTABLE INSTRUCTIONS
1014
GRAPHICS OBJECTS
1016
MEMORY
1050
OPERATING SYSTEM (OS)
1020
USER MODE GRAPHICS DRIVER
1026
SHADER COMPILER
1027
SHADER COMPILER
1024
GRAPHICS API
(e.g. Direct3D/OpenGL)
1022
KERNEL MODE GRAPHICS DRIVER
1029
OS KERNEL MODE FUNCTIONS
1028
GRAPHICS PROCESSOR
1032
PROCESSOR
1030
GENERAL PURPOSE CORE(s)
1034

FIG. 10

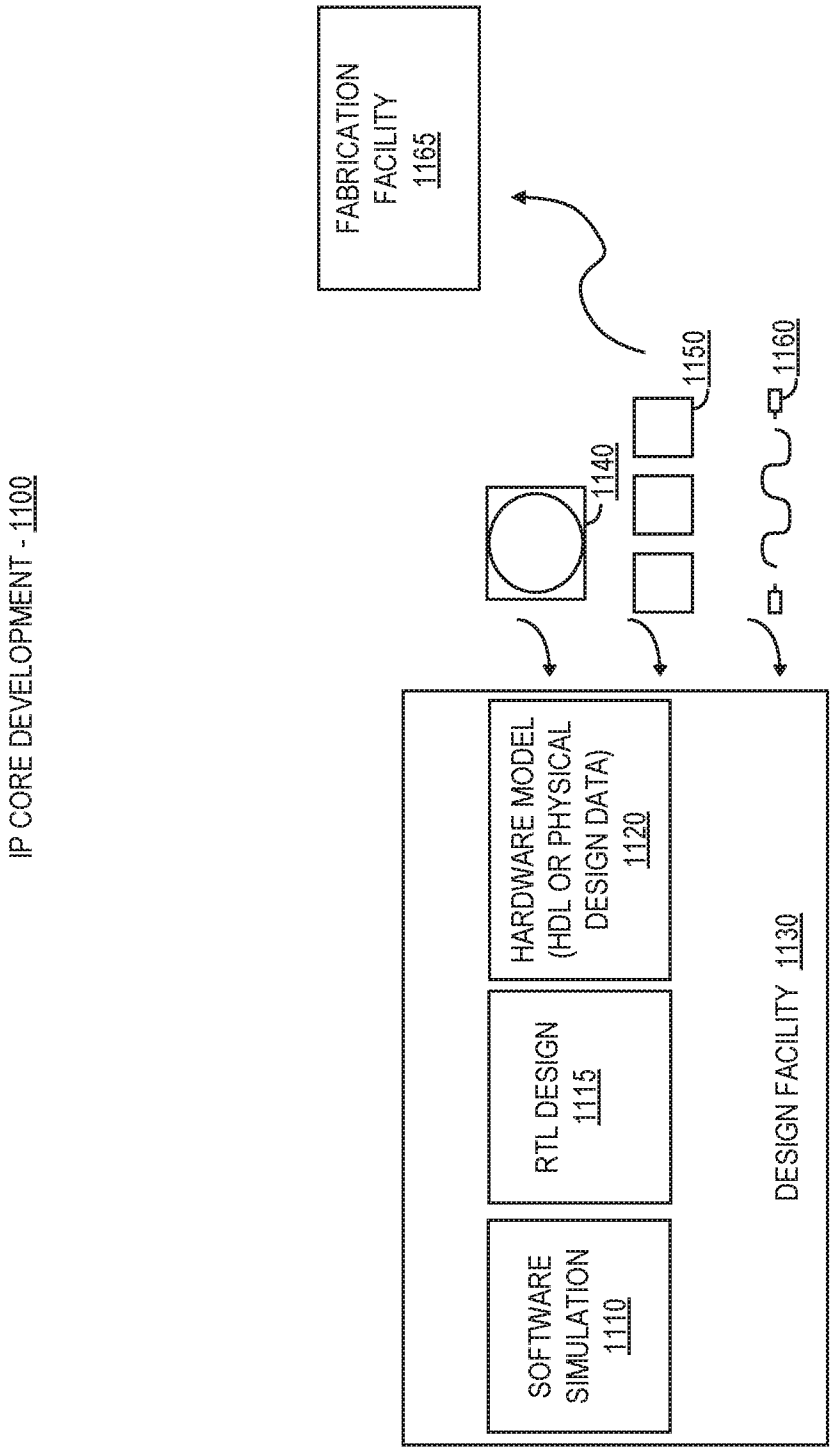

IP CORE DEVELOPMENT - 1100
FABRICATION FACILITY 1165
1140
1150
1160
HARDWARE MODEL (HDL OR PHYSICAL DESIGN DATA) 1120
RTL DESIGN 1115
SOFTWARE SIMULATION 1110
DESIGN FACILITY 1130

FIG. 11

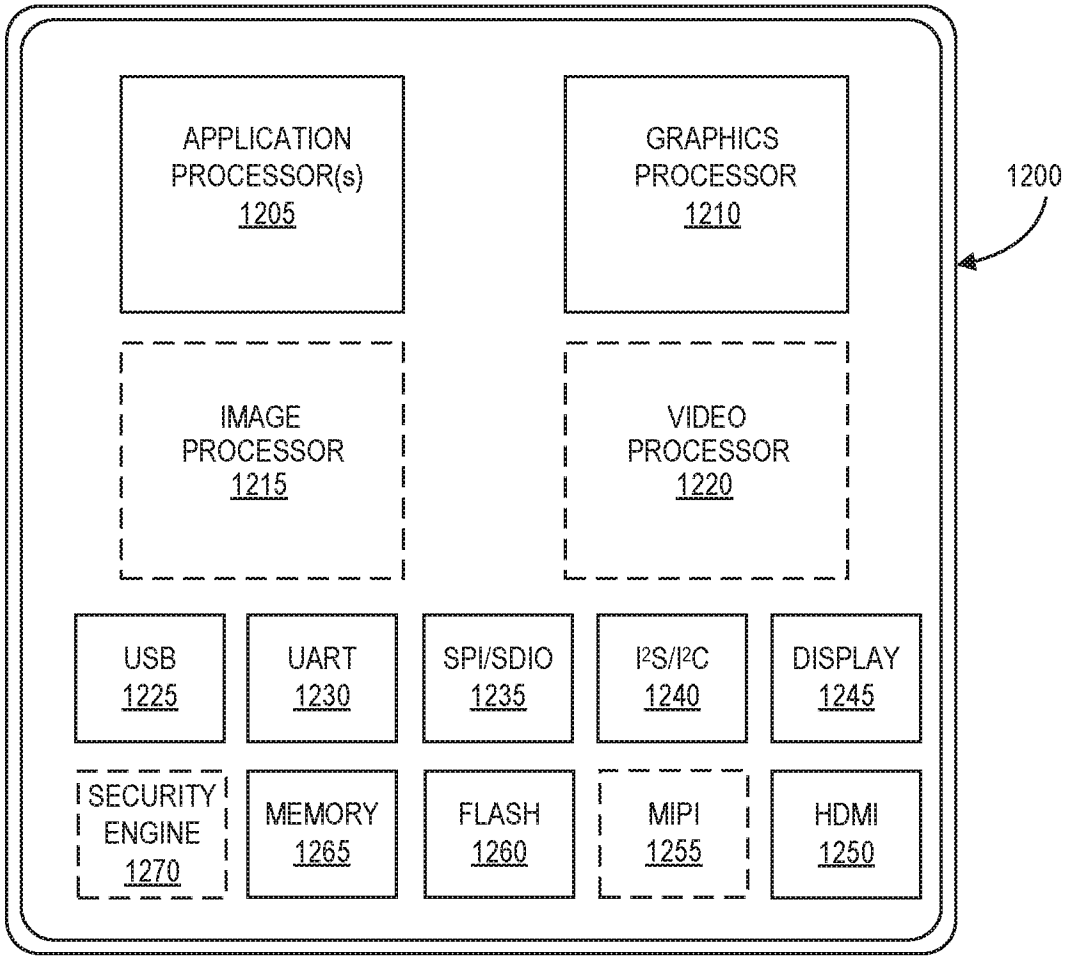

APPLICATION PROCESSOR(s) 1205
GRAPHICS PROCESSOR 1210
1200
IMAGE PROCESSOR 1215
VIDEO PROCESSOR 1220
USB 1225
UART 1230
SPI/SDIO 1235
I²S/I²C 1240
DISPLAY 1245
SECURITY ENGINE 1270
MEMORY 1265
FLASH 1260
MIPI 1255
HDMI 1250

FIG. 12

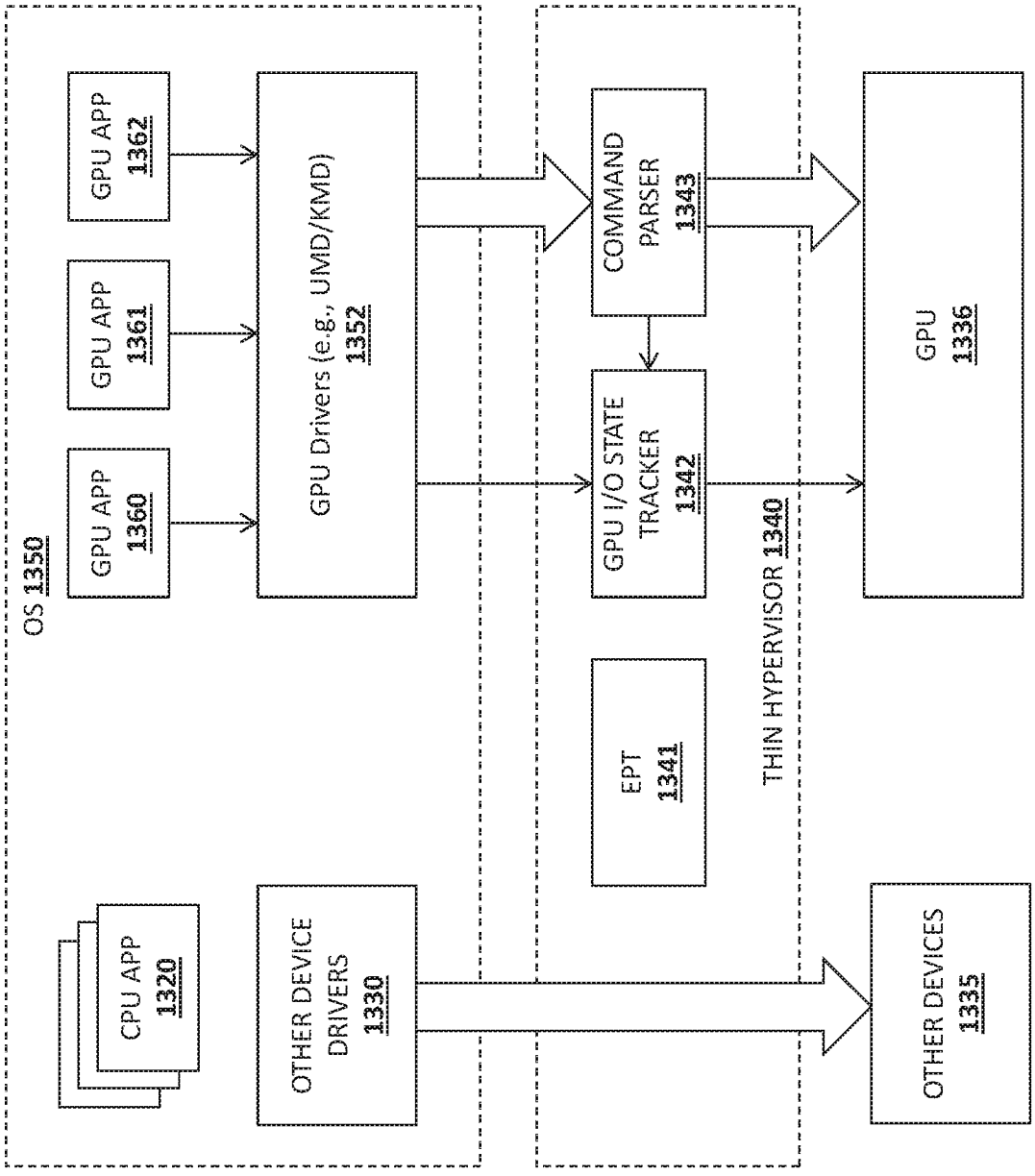

OS 1350
GPU APP 1362
GPU APP 1361
GPU APP 1360
GPU Drivers (e.g., UMD/KMD) 1352
COMMAND PARSER 1343
GPU I/O STATE TRACKER 1342
EPT 1341
THIN HYPERVISOR 1340
GPU 1336
CPU APP 1320
OTHER DEVICE DRIVERS 1330
OTHER DEVICES 1335

FIG. 13

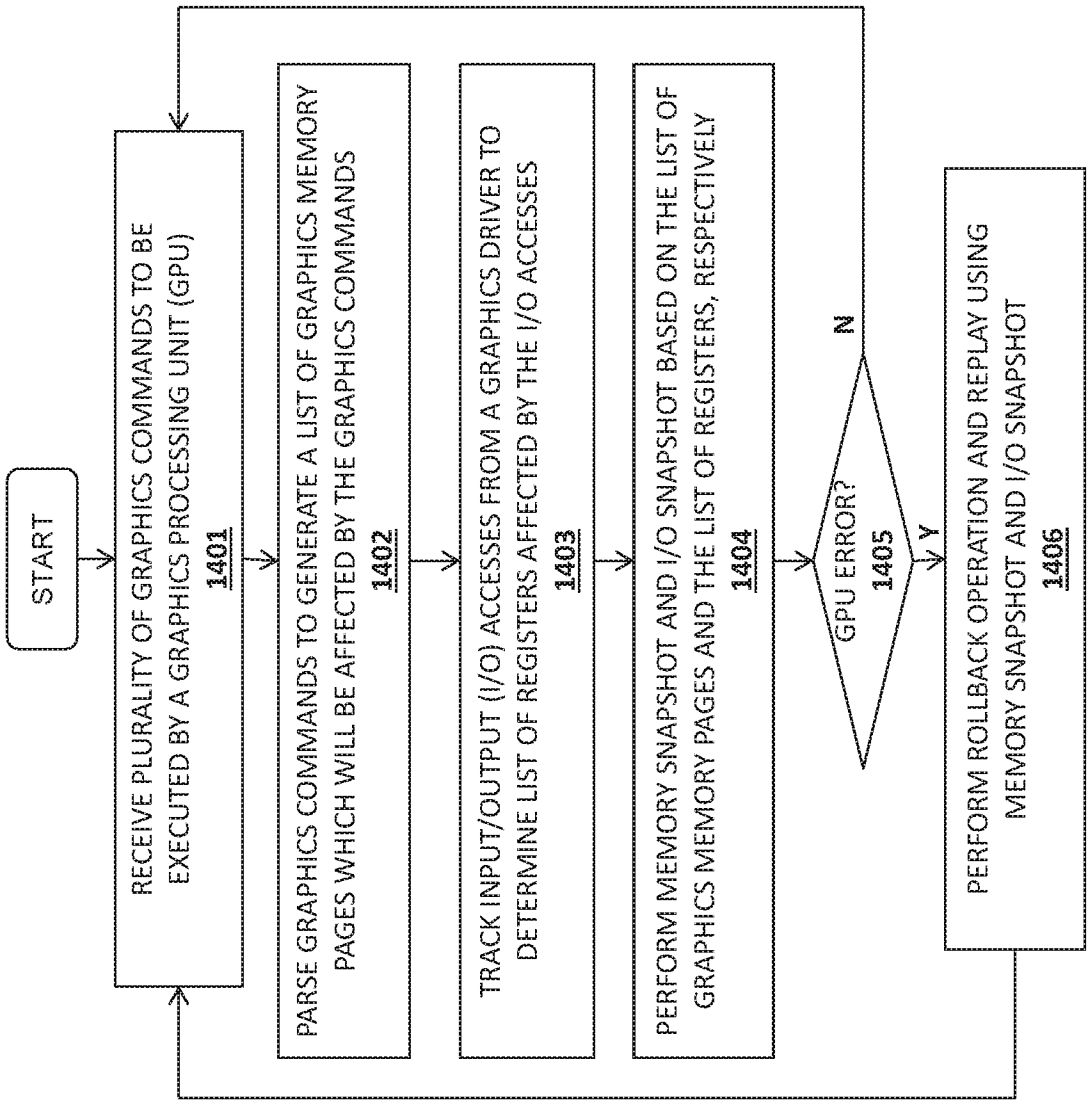

START
RECEIVE PLURALITY OF GRAPHICS COMMANDS TO BE EXECUTED BY A GRAPHICS PROCESSING UNIT (GPU)
1401
PARSE GRAPHICS COMMANDS TO GENERATE A LIST OF GRAPHICS MEMORY PAGES WHICH WILL BE AFFECTED BY THE GRAPHICS COMMANDS
1402
TRACK INPUT/OUTPUT (I/O) ACCESSES FROM A GRAPHICS DRIVER TO DETERMINE LIST OF REGISTERS AFFECTED BY THE I/O ACCESSES
1403
PERFORM MEMORY SNAPSHOT AND I/O SNAPSHOT BASED ON THE LIST OF GRAPHICS MEMORY PAGES AND THE LIST OF REGISTERS, RESPECTIVELY
1404
GPU ERROR?
1405
N
Y
PERFORM ROLLBACK OPERATION AND REPLAY USING MEMORY SNAPSHOT AND I/O SNAPSHOT
1406

FIG. 14

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2016/078282

A. CLASSIFICATION OF SUBJECT MATTER

G06F 11/14(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

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

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

WPI, EPODOC, CNPAT, CNKI: CPU, GPU, rollback, register, IO, I/O, snapshot, error, crash, memory

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2015074458 A1 (INMAGE SYSTEMS, INC.) 12 March 2015 (2015-03-12) description, paragraphs [263]-[330]	1-24
Y	CN 102023863 A (ZTE CORP.) 20 April 2011 (2011-04-20) description, pages 4-5	1-24
A	CN 101317160 A (INTEL CORP.) 03 December 2008 (2008-12-03) the whole document	1-24
A	WO 2014062191 A1 (HEWLETT-PACKARD DEVELOPMENT COMPANY, L.P.) 24 April 2014 (2014-04-24) the whole document	1-24

☐ 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 another 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
12 December 2016	**29 December 2016**
Name and mailing address of the ISA/CN **STATE INTELLECTUAL PROPERTY OFFICE OF THE P.R.CHINA** **6, Xitucheng Rd., Jimen Bridge, Haidian District, Beijing 100088** **China** Facsimile No. **(86-10)62019451**	Authorized officer **WU,Xiaodong** Telephone No. **(86-10)62413717**

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2016/078282

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
US	2015074458	A1	12 March 2015	US	9098455	B2	04 August 2015
				US	2009313503	A1	17 December 2009
				US	8949395	B2	03 February 2015
				US	2006010227	A1	12 January 2006
				US	2006031468	A1	09 February 2006
				US	2007271428	A1	22 November 2007
				US	2007282921	A1	06 December 2007
				US	2008294843	A1	27 November 2008
				US	2010169591	A1	01 July 2010
				US	2010169587	A1	01 July 2010
				US	2010169452	A1	01 July 2010
				US	2010169283	A1	01 July 2010
				US	2010169282	A1	01 July 2010
				US	2010169281	A1	01 July 2010
				US	2011184918	A1	28 July 2011
CN	102023863	A	20 April 2011	CN	102023863	B	22 July 2015
CN	101317160	A	03 December 2008	WO	2007078891	A1	12 July 2007
				CN	101317160	B	22 August 2012
				EP	1966697	A1	10 September 2008
				EP	1966697	B1	26 May 2010
				AT	469394	T	15 June 2010
				US	2007162520	A1	12 July 2007
				US	7730286	B2	01 June 2010
				DE	602006014596	E	08 July 2010
WO	2014062191	A1	24 April 2014	EP	2909724	A4	20 July 2016
				US	2015261463	A1	17 September 2015
				EP	2909724	A1	26 August 2015
				CN	104854566	A	19 August 2015