



(51) International Patent Classification:

H01L 25/065 (2006.01) *H01L 23/532* (2006.01)
H01L 25/07 (2006.01) *H01L 23/00* (2006.01)

(21) International Application Number:

PCT/US2017/025663

(22) International Filing Date:

01 April 2017 (01.04.2017)

(25) Filing Language:

English

(26) Publication Language:

English

(71) Applicant: **INTEL CORPORATION** [US/US]; 2200 Mission College Boulevard, Santa Clara, California 94054 (US).

(72) Inventors; and

(71) Applicants: **HOSSAIN, MD Altaf** [US/US]; 6418 NW McMullen Avenue, Portland, Oregon 97229 (US). **SOMASEKHAR, Dinesh** [IN/US]; 621 NW 112th Avenue,

Portland, Oregon 97229 (US). **SUBBAREDDY, Dheeraj** [IN/US]; 5885 NW 170th Avenue, Portland, Oregon 97229 (US).

(74) Agent: **BRASK, Justin, K.** et al.; Schwabe, Williamson & Wyatt, 1211 SW 5th Ave., Suite 1900, Portland, OR 97204 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, 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, KH, KN, KP, KR, KW, 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.

(54) Title: INNOVATIVE WAY TO DESIGN SILICON TO OVERCOME RETICLE LIMIT

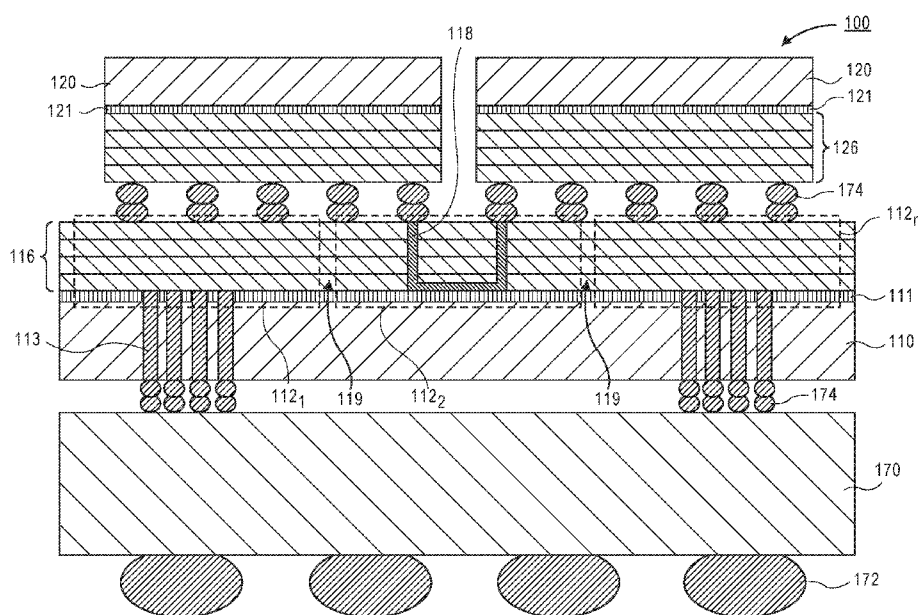


FIG. 1

(57) Abstract: Embodiments of the invention include a stacked die system and methods for forming such systems. In an embodiment, the stacked die system may include a first die. The first die may include a device layer and a plurality of routing layers formed over the device layer. The plurality of routing layers may be segmented into a plurality of sub regions. In an embodiment no conductive traces in the plurality of routing layers pass over a boundary between any of the plurality of sub regions. In an embodiment, the stacked die system may also include a plurality of second dies stacked over the first die. According to an embodiment, at least a two of the second dies are communicatively coupled to each other by a die to die interconnect formed entirely within a single sub region in the first die.



(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))*

INNOVATIVE WAY TO DESIGN SILICON TO OVERCOME RETICLE LIMIT

FIELD OF THE INVENTION

Embodiments of the invention are in the field of stacked semiconductor dies and, in particular, stacked semiconductor dies where the bottom die is larger than the reticle limit.

BACKGROUND OF THE INVENTION

High performance computing applications, such as servers, are driving an increase in the physical size of the die. However, the maximum size of the die may be limited by processing design rules. For example, the reticle limit used to fabricate the die cannot be exceeded without additional processing, which increases the overall cost of the die. Currently, when the reticle limit is not large enough to accommodate the entire size of the die, reticle stitching may be used. The use of stitching is not desirable for high volume manufacturing due to increases in cost of each die. The additional cost is due to the need for two masks for every layer in order to form the interconnects across the normal reticle boundaries. In addition to the cost of an additional mask layer, the addition of a mask layer increases the probability of overlay errors and may reduce the yield.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a cross-sectional illustration of a system that includes a plurality of stacked dies where the lower die is unstitched and larger than the reticle limit, according to an embodiment of the invention.

Figure 2 is a perspective illustration of a system that includes a plurality of stacked dies where a die to die interconnect line formed entirely in one sub region of the bottom die is shown, according to an embodiment of the invention.

Figure 3A is a cross-sectional illustration of a first die after the front end of line processing has been completed, according to an embodiment of the invention.

Figure 3B is a cross-sectional illustration of the first die after through silicon vias are formed in the first die, according to an embodiment of the invention.

Figure 3C is a cross-sectional illustration of the first die after the thickness of the semiconductor substrate is reduced to expose the through silicon vias, according to an embodiment of the invention.

Figure 3D is a cross-sectional illustration of the first die after a plurality of interconnect layers have been formed over the device layer, according to an embodiment of the invention.

Figure 3E is a cross-sectional illustration after the first die is mounted to a packaging substrate, according to an embodiment of the invention.

Figure 3F is a cross-sectional illustration after a plurality of second dies are mounted onto the first die, according to an embodiment of the invention.

5 Figure 4 is a schematic of a computing device built in accordance with an embodiment of the invention.

DETAILED DESCRIPTION OF THE INVENTION

Described herein are systems with a plurality of stacked dies where the bottom die is
10 larger than the reticle limit and includes no stitching and methods of forming such stacked die systems. In the following description, various aspects of the illustrative implementations will be described using terms commonly employed by those skilled in the art to convey the substance of their work to others skilled in the art. However, it will be apparent to those skilled in the art that the present invention may be practiced with only some of the described aspects. For purposes of
15 explanation, specific numbers, materials and configurations are set forth in order to provide a thorough understanding of the illustrative implementations. However, it will be apparent to one skilled in the art that the present invention may be practiced without the specific details. In other instances, well-known features are omitted or simplified in order not to obscure the illustrative implementations.

20 Various operations will be described as multiple discrete operations, in turn, in a manner that is most helpful in understanding the present invention, however, the order of description should not be construed to imply that these operations are necessarily order dependent. In particular, these operations need not be performed in the order of presentation.

As noted above, the increased processing demands in high performance computing
25 systems, such as servers, may require the growth of dies to sizes beyond the typical size of available reticles. In stacked die systems, increasing the size of the bottom die beyond the reticle limit may allow for more dies to be stacked on the bottom die. However, die stitching is currently needed to fabricate dies that are larger than the reticle limit. As noted above, the die stitching process increases the overall cost of each device and reduces throughput. Accordingly,
30 embodiment of the invention include stacked dies where the bottom die is larger than the reticle limit but does not require stitching. Instead, embodiments of the invention allow for larger bottom dies without breaking the design rules by utilizing an interconnect that spans both the top and bottom die.

According to an embodiment, the bottom die may be an older generation process

technology and the top die (or dies) may utilize advanced process technology. Accordingly, the bottom die may be used for lower speed processing operations and the top dies may be used for faster processing operations, such as processing cores. As such, increasing the size of the bottom die provides additional space to stack a plurality of high performance cores in a single system.

5 Therefore, embodiments of the invention allow for improved performance due to the increased number of processing cores that may be included in a single system.

Referring now to Figure 1, a cross-sectional illustration of a stacked die system 100 is shown according to an embodiment of the invention. According to an embodiment, the stacked die system 100 may include a first die 110 that is mounted onto a packaging substrate 170. The
10 first die 110 may be mounted onto the packaging substrate 170 with any suitable interconnect 174, such as solder bumps, controlled collapse chip connection (C4) bumps, or the like. The packaging substrate 170 may be any suitable packaging substrate known in the art, such as a cored or coreless packaging substrate. In the illustrated embodiment, the packaging substrate 170 may also include a plurality of solder bumps 172 for connecting the packaging substrate to a
15 mother board or the like (not shown).

In an embodiment, the first die 110 may include any semiconductor material. For example the semiconductor material may be a crystalline substrate formed using a bulk silicon or a silicon-on-insulator substructure. In other implementations, the semiconductor material may be formed using alternate materials, which may or may not be combined with silicon, that include
20 but are not limited to germanium, indium antimonide, lead telluride, indium arsenide, indium phosphide, gallium arsenide, indium gallium arsenide, gallium antimonide, or other combinations of group III-V or group IV materials. Although a few examples of materials from which the first die 110 may be formed are described here, any material that may serve as a foundation upon which a semiconductor device may be built falls within the scope of the present
25 invention.

In an embodiment the first die 110 may include an active device layer 111 that includes integrated circuitry. In some embodiments, the active device layer 111 of the first die 110 may include integrated circuitry that is used for processing operations that require lower processing speeds compared to processing operations that will be implemented in the plurality of second
30 dies 120, described in greater detail below. In an embodiment, a plurality of vias 113 may communicatively couple the active device layer 111 to the bumps 174 on the opposite side of the first die 110. For example, the vias 113 may be through silicon vias (TSVs) or the like.

In an embodiment, a plurality of routing layers 116 may be formed over the active device layer 111. For example, the routing layers 116 may include a plurality of conductive routing

lines and vias (not shown) formed within a plurality of dielectric layers, as is known in the art. According to an embodiment, the plurality of routing layers may be formed with typical lithographic and deposition processes. However, in embodiments of the invention the footprint of the first die 110 is larger than the reticle limit used to pattern the routing layers 116. As such, 5 embodiments include a plurality of routing layers that are separated into a plurality of sub regions 112₁-112_n. In an embodiment each sub region 112 may be substantially the same size as the reticle limit.

Accordingly, multiple exposures made while stepping a reticle over the surface of the first die 110 may allow for the first die 110 to be larger than the reticle limit. In an embodiment, each 10 of the sub regions 112 in the first die 110 are not stitched together with additional masking and patterning. As such, conductive traces in the routing layers 116 may not pass over the boundaries 119 between the sub regions 112. For example, the die to die interconnect 118 is formed entirely within the sub region 112₂ and communicatively couples the two second dies 120 mounted over the first die 110. For clarity, the only conductive trace illustrated in the routing layers 116 in 15 Figure 1 is the die to die interconnect 118, though it is to be appreciated that many other conductive traces may be included in the routing layers 116. Additionally, while the die to die interconnect line 118 is illustrated as two vias that connect to a trace on the bottommost routing layer of the routing layers 116, it is to be appreciated that the vias and routing layer may be formed in any of the layers.

20 The ability to form first dies 110 that have a larger footprint than the reticle limit provides a larger surface on which to mount second dies 120. In an embodiment, the second dies 120 may be fabricated with different processing technology than the first dies 110. In an embodiment, the second dies 120 may be an advanced processing technology relative to the first die 110. For example, the second dies 120 may be high performance dies, such as processing cores or the like. 25 In an embodiment, the second dies 120 may include a semiconductor substrate and a device layer 121.

According to an embodiment, the semiconductor substrate of the second dies 120 may be the same semiconductor material as the first semiconductor die 110 or the second dies 120 may include a different semiconductor material than the first die 110. For example, the first 30 semiconductor die 110 may be silicon and the second semiconductor dies 120 may be a III-V semiconductor. In an embodiment, each of the second semiconductor dies 120 may be substantially similar to each other. Alternative embodiments may include second semiconductor dies 120 that are not all substantially similar to each other.

Even though the sub regions 112 of the first die 110 are not stitched together, embodiments

of the invention allow for the second dies 120 to be communicatively coupled. In an embodiment, the second dies 120 may be communicatively coupled by the die to die interconnect 118 in the first die 110. For example, bumps 172 on each of the second dies 120 may be attached to portions of the die to die interconnect 118. Accordingly, signals may pass over the reticle boundary 119 in the first die 110, even though no stitching is present between the sub regions 112.

Referring now to Figure 2, a perspective view of a stacked die system with a plurality of second dies 120_1 - 120_n stacked on a first die 110 is shown, according to an embodiment of the invention. Figure 2 more clearly illustrates size and arrangement of the sub regions 112_1 - 112_n of the first die 110. In the illustrated embodiment, the first die 110 includes nine sub regions 112 arranged in a three by three pattern. However, it is to be appreciated that the first die 110 may include any number of sub regions 112 and the sub regions 112 may be arranged in any desired pattern. In the illustrated embodiment, four second dies 120_1 - 120_n are stacked on the first die 110 in a two by two arrangement. However, it is to be appreciated that any number of second dies 120 may be used and they may be arranged in any desired pattern.

In an embodiment, the second dies 120 may be arranged on the first die 110 so that they cross over a boundary region 119 between sub regions 112 of the first die 110. Additionally, embodiments of the invention may include one or more of second dies 120 that have a larger footprint than a footprint of the sub regions 112. For example, as illustrated in Figure 2, the length L_2 of a side of a second die 120 may be greater than a length L_1 of a sub region 112. While the second dies 120 may be larger (area wise) than each sub region 112 of the first die 110, it is to be appreciated that the second dies 120 may not be larger than the reticle limit used to fabricate the second dies 120. This is possible because embodiments of the invention do not require the first die 110 and the second dies 120 to be processed on the same wafer. Accordingly, the first die 110 and the second dies 120 may be fabricated with different semiconductor processing equipment and/or with different processing technologies. This may be particularly beneficial when the first die 110 is used for low speed operations and, therefore, does not require advanced processing techniques that may be needed for second dies 120 that are used for high speed operations. Therefore, the decoupled processing allows for improved efficiencies because the more expensive advanced processing technologies may only be used to fabricate the higher performance second dies 120, whereas the less expensive processing technologies may be used to fabricate the lower performance first die 110.

In order to provide die-to-die communication between two of the second dies 120, embodiments of the invention may include die-to-die interconnects formed in a single sub region

112 of the first die 110. For example, a conductive trace and vias 118 (schematically represented as a dashed line) may be patterned in a single sub region 112 of the first die 110. As illustrated in Figure 2, second die 120₁ and second die 120₂ are communicatively coupled by conductive traces and vias 118. The conductive trace and vias 118 are formed entirely in sub region 112₂ of the first die, and provide a connection between second die 120₁ and second die 120₂. As such, the conductive trace and vias 118 may be formed without needing additional mask layers to stitch across the boundary regions 119 between sub regions 112. While a single die to die interconnect 118 is shown in sub region 112₂, it is to be appreciated that any number of the sub regions 112 may include die to die interconnects 118 and/or any number of the die to die interconnects may be formed in each sub region 112.

Referring now to Figures 3A-3F a process flow for forming a stacked die system where the first die is an unstitched die with sub portions that are larger than the reticle limit is shown, according to an embodiment of the invention. In Figure 3A, a cross-sectional illustration of first die 110 is shown. In an embodiment, a device layer 111 may be formed on the first die 110. For example, the device layer 111 may include transistors, diodes, or any other circuit elements (not shown), as is known in the art. The device layer 111 may be fabricated with typical front end of line (FEOL) semiconductor processing technologies.

Referring now to Figure 3B, a cross-sectional illustration of the first die 110 after a plurality of vias 113 are formed into the first die 110 is shown, according to an embodiment of the invention. In an embodiment, the vias 113 may be formed by etching a trench into the semiconductor substrate of the first die 110 and disposing a conductive material into the trenches. While not shown in Figure 3B, embodiments may also include forming a diffusion barrier between the conductive material and the semiconductor material in order to minimize diffusion of the conductive material into the semiconductor material.

Referring now to Figure 3C, a cross-sectional illustration of the first die 110 after the first die 110 has been thinned is shown, according to an embodiment of the invention. Thinning the first die 110 exposes a surface 114 of the vias 113. As such, the vias 113 may electrically couple opposing sides of the first die 110 and allow for the stacking of subsequent dies, as will be described in greater detail below.

Referring now to Figure 3D, a cross-sectional illustration of the first die 110 is shown after routing layers 116 are formed over the device layer 111, according to an embodiment of the invention. According to an embodiment, the footprint of the first die is greater than the reticle limit. Accordingly, the first die is fabricated with multiple exposures of sub regions 112. For example, for each layer of the routing layers 116, a reticle may be positioned over a sub region

112 and that sub region is exposed. The process may be repeated with each sub region being individually exposed with a reticle. In an embodiment, each sub region 112 may include conductive traces and vias for routing. However, for clarity and in order to not obscure embodiments of the invention, the routing lines other than the die to die interconnect 118 are omitted. In an embodiment, the die to die interconnect 118 may be used to communicatively couple two second dies that will be mounted over the first die 110. According to an embodiment, the die to die interconnect 118 is formed entirely within a single sub region 112₂. As such, no stitching is needed to form the die to die interconnect 118 since the interconnect 118 never crosses over a reticle boundary 119 between sub regions 112. While the die to die interconnect 118 is shown in a single sub region 112, it is to be appreciated that any number of the sub regions 112 may include die to die interconnects 118, and/or any number of die to die interconnects 118 may be formed in a single sub region 112. Additionally, while the die to die interconnect 118 is illustrated as two vias that connect to a trace on the bottommost layer of the routing layers 116, it is to be appreciated that the vias and routing layer may be formed in any layer of the routing layers 116. In some embodiments, the die to die interconnect 118 may be formed on the topmost layer of the routing layers 116, and therefore, may only require a conductive trace and the vias may be omitted.

Referring now to Figure 3E, a cross-sectional illustration of the first die 110 is shown after the first die 110 is mounted to a packaging substrate 170, according to an embodiment of the invention. In an embodiment, the first die 110 may be diced (i.e., singulated from a wafer that includes a plurality of first dies 110) prior to being mounted on the packaging substrate 170. In an embodiment, the first die 110 may be mounted to the packaging substrate 170 with a plurality of bumps 174 (e.g., solder bumps, C4 bumps, or the like). Embodiments may also include a packaging substrate 170 that includes a plurality of solder bumps 172 formed on a surface of the packaging substrate 170 facing away from the first die 110. For example, the solder bumps 172 may be used to mount the packaging substrate to another substrate, such as a printed circuit board (PCB), motherboard, or the like.

Referring now to Figure 3F, a cross-sectional illustration of the first die 110 is shown after a plurality of second dies 120 have been mounted to a surface of the first die 110, according to an embodiment of the invention. In an embodiment, the plurality of second dies 120 may be mounted to the first die with known interconnects 174, such as solder bumps, C4 bumps, or the like. In an embodiment, the plurality of second dies 120 may be higher performance processing dies compared to the first die 110. In some embodiments, each of the second dies 120 may be substantially similar to each other. Other embodiments may include second dies 120 that are not

all substantially similar to each other.

As illustrated in Figure 3F, the second dies 120 may be positioned over the first die 110 such that the footprint of the second die 120 crosses over the boundary region 119 between sub regions of the first die 110. In an embodiment, the die to die connection 118 may

5 communicatively couple two of the second dies 120. In order to eliminate the need for die stitching, the die to die connection 118 may be fabricated in a single sub region 112 of the first die 110. For example, in Figure 3F, the die to die interconnect 118 is entirely formed within the second sub region 112₂. Since portions of each second die 120 overlap at least a portion of the second sub region 112₂, the die to die interconnect 118 may be formed between the two second
10 dies 120 without needing a conductive trace that extends beyond the edge of the sub region 112₂. In the illustrated embodiment, solder bumps 174 coupled to each of the second dies 120 may be electrically coupled to the die to die interconnect 118 formed on the first die 110. While the innermost solder bumps 174 on each of the second dies 120 are shown as being connected to the die to die interconnect 118, it is to be appreciated that any of the solder bumps 174 may be
15 coupled to the die to die connection 118 so long as the bumps 174 are positioned over at least a portion of the same sub region 112, according to other embodiments.

Figure 4 illustrates a computing device 400 in accordance with one implementation of the invention. The computing device 400 houses a board 402. The board 402 may include a number of components, including but not limited to a processor 404 and at least one communication chip
20 406. The processor 404 is physically and electrically coupled to the board 402. In some implementations the at least one communication chip 406 is also physically and electrically coupled to the board 402. In further implementations, the communication chip 406 is part of the processor 404.

Depending on its applications, computing device 400 may include other components that
25 may or may not be physically and electrically coupled to the board 402. These other components include, but are not limited to, volatile memory (e.g., DRAM), non-volatile memory (e.g., ROM), flash memory, a graphics processor, a digital signal processor, a crypto processor, a chipset, an antenna, a display, a touchscreen display, a touchscreen controller, a battery, an audio codec, a video codec, a power amplifier, a global positioning system (GPS) device, a compass,
30 an accelerometer, a gyroscope, a speaker, a camera, and a mass storage device (such as hard disk drive, compact disk (CD), digital versatile disk (DVD), and so forth).

The communication chip 406 enables wireless communications for the transfer of data to and from the computing device 400. The term “wireless” and its derivatives may be used to describe circuits, devices, systems, methods, techniques, communications channels, etc., that

may communicate data through the use of modulated electromagnetic radiation through a non-solid medium. The term does not imply that the associated devices do not contain any wires, although in some embodiments they might not. The communication chip 406 may implement any of a number of wireless standards or protocols, including but not limited to Wi-Fi (IEEE 802.11 family), WiMAX (IEEE 802.16 family), IEEE 802.20, long term evolution (LTE), Ev-DO, HSPA+, HSDPA+, HSUPA+, EDGE, GSM, GPRS, CDMA, TDMA, DECT, Bluetooth, derivatives thereof, as well as any other wireless protocols that are designated as 3G, 4G, 5G, and beyond. The computing device 400 may include a plurality of communication chips 406. For instance, a first communication chip 406 may be dedicated to shorter range wireless communications such as Wi-Fi and Bluetooth and a second communication chip 406 may be dedicated to longer range wireless communications such as GPS, EDGE, GPRS, CDMA, WiMAX, LTE, Ev-DO, and others.

The processor 404 of the computing device 400 includes an integrated circuit die packaged within the processor 404. In some implementations of the invention, the integrated circuit die of the processor may include a stacked die system including an unstitched first die and a plurality of second dies stacked over the first die, in accordance with implementations of the invention. The term “processor” may refer to any device or portion of a device that processes electronic data from registers and/or memory to transform that electronic data into other electronic data that may be stored in registers and/or memory.

The communication chip 406 also includes an integrated circuit die packaged within the communication chip 406. In accordance with another implementation of the invention, the integrated circuit die of the communication chip may include a stacked die system including an unstitched first die and a plurality of second dies stacked over the first die, in accordance with implementations of the invention.

The above description of illustrated implementations of the invention, including what is described in the Abstract, is not intended to be exhaustive or to limit the invention to the precise forms disclosed. While specific implementations of, and examples for, the invention are described herein for illustrative purposes, various equivalent modifications are possible within the scope of the invention, as those skilled in the relevant art will recognize.

These modifications may be made to the invention in light of the above detailed description. The terms used in the following claims should not be construed to limit the invention to the specific implementations disclosed in the specification and the claims. Rather, the scope of the invention is to be determined entirely by the following claims, which are to be construed in accordance with established doctrines of claim interpretation.

Example 1: a stacked die system, comprising: first die, wherein the first die comprises: a device layer; and a plurality of routing layers formed over the device layer, wherein the plurality of routing layers are segmented into a plurality of sub regions, and wherein no conductive traces in the plurality of routing layers pass over a boundary between any of the plurality of sub regions; and a plurality of second dies stacked over the first die, wherein one or more of the second dies cross over the boundary between any of the plurality of sub regions.

Example 2: the stacked die system of Example 1, wherein at least two of the second dies are positioned over at least a portion of the same sub region of the first die.

Example 3: the stacked die system of Example 2, wherein the two second dies are communicatively coupled to each other by a die to die interconnect formed entirely within the sub region of the first die that the two second dies are positioned at least partially over.

Example 4: the stacked die system of Example 1, Example 2, or Example 3, wherein a footprint of each of the second dies is larger than a footprint of the sub regions of the first die.

Example 5: the stacked die system of Example 1, Example 2, Example 3, or Example 4, wherein second dies are mounted to the first die with a controlled collapse chip connection.

Example 6: the stacked die system of Example 1, Example 2, Example 3, Example 4, or Example 5, wherein the plurality of second dies are high speed processing dies relative to the first die.

Example 7: the stacked die system of Example 6, wherein the second dies are processing cores.

Example 8: the stacked die system of Example 1, Example 2, Example 3, Example 4, Example 5, or Example 6, wherein the plurality of second dies are formed on a different semiconductor material than the first die.

Example 9: the stacked die system of Example 8, wherein the plurality of second dies are formed on a III-V semiconductor substrate, and the first die is formed on a silicon substrate.

Example 10: the stacked die system of Example 1, Example 2, Example 3, Example 4, Example 5, Example 6, Example 7, Example 8, or Example 9, wherein the sub regions on the first die are arranged in a three by three pattern, and wherein there are four second dies.

Example 11: a method for forming a stacked die system, comprising: forming an active device layer on a first semiconductor die; forming a plurality of routing layers over a surface of the active device layer, wherein the routing layers are segmented into a plurality of sub regions, and wherein no conductive traces in the plurality of routing layers pass over a boundary between any of the plurality of sub regions; and mounting a plurality of second dies onto the first die.

Example 12: the method of Example 11, wherein the first die has a footprint that is larger

than a reticle limit of the processing used to form the plurality of routing layers.

Example 13: the method of Example 11 or Example 12, wherein forming the routing layers includes individually exposing each sub region with a reticle, and wherein no two sub regions are simultaneously exposed with a single reticle.

5 Example 14: the method of Example 11, Example 12, or Example 13, wherein at least two of the second dies are mounted over at least a portion of the same sub region of the first die.

Example 15: the method of Example 14, wherein the two second dies are communicatively coupled to each other by a die to die interconnect formed entirely within a single sub region of the first die.

10 Example 16: the method of Example 11, Example 12, Example 13, Example 14, or Example 15, further comprising: forming a plurality of vias through the first die, wherein the vias communicatively couple the device layer of the first die to a surface of the first die opposite to the device layer; and mounting the first die to a packaging substrate, wherein the device layer of the first die is formed on a surface of the first die facing away from a surface of the packaging
15 substrate.

Example 17: the method of Example 11, Example 12, Example 13, Example 14, Example 15, or Example 16, wherein the plurality of second dies are fabricated with advanced processing technology relative to the first die.

Example 18: the method of Example 17, wherein the second dies are cores.

20 Example 19: the method of Example 17 or Example 18, wherein a footprint of each of the second dies is greater than a footprint of each of the sub regions on the first die.

Example 20: the method of Example 11, Example 12, Example 13, Example 14, Example 15, Example 16, Example 17, Example 18, or Example 19, wherein the plurality of sub regions includes nine or more sub regions, and wherein the plurality of second dies includes four or more
25 second dies.

Example 21: a stacked die system, comprising: a first die, wherein the first die comprises: a device layer; a plurality of routing layers formed over the device layer, wherein the plurality of routing layers are segmented into a plurality of sub regions, and wherein no conductive traces in the plurality of routing layers pass over a boundary between any of the plurality of sub regions;
30 and a plurality of vias through the first die, wherein the vias electrically couple the device layer with a surface of the die opposite to the device layer; a plurality of second dies stacked over the first die, wherein one or more of the second dies cross over the boundary between any of the plurality of sub regions, and wherein at least a first of the second dies are communicatively coupled to at least a second of the second dies by a die to die interconnect formed entirely within

a single sub region in the first die; and a packaging substrate communicatively coupled to a the plurality of vias in the first die.

Example 22: the stacked die system of Example 21, wherein a footprint of each of the second dies is larger than a footprint of the sub regions of the first die.

5 Example 23: the stacked die system of Example 21 or Example 22, wherein the plurality of second dies are high speed processing dies relative to the first die.

Example 24: the stacked die system of Example 21, Example 22, or Example 23, wherein the second dies are processing cores.

10 Example 25: the stacked die system of Example 21, Example 22, Example 23, or Example 24, wherein the sub regions on the first die are arranged in a three by three pattern, and wherein there are four second dies arranged in a two by two pattern.

CLAIMS**What is claimed is:**

1. A stacked die system, comprising:
5 a first die, wherein the first die comprises:
 a device layer; and
 a plurality of routing layers formed over the device layer, wherein the
plurality of routing layers are segmented into a plurality of sub regions, and wherein no
conductive traces in the plurality of routing layers pass over a boundary between any of
10 the plurality of sub regions; and
 a plurality of second dies stacked over the first die, wherein one or more of the
second dies cross over the boundary between any of the plurality of sub regions.
2. The stacked die system of claim 1, wherein at least two of the second dies are positioned
15 over at least a portion of the same sub region of the first die.
3. The stacked die system of claim 2, wherein the two second dies are communicatively
coupled to each other by a die to die interconnect formed entirely within the sub region of
the first die that the two second dies are positioned at least partially over.
- 20 4. The stacked die system of claim 1, wherein a footprint of each of the second dies is larger
than a footprint of the sub regions of the first die.
5. The stacked die system of claim 1, wherein second dies are mounted to the first die with
25 a controlled collapse chip connection.
6. The stacked die system of claim 1, wherein the plurality of second dies are high speed
processing dies relative to the first die.

7. The stacked die system of claim 6, wherein the second dies are processing cores.
8. The stacked die system of claim 1, wherein the plurality of second dies are formed on a
5 different semiconductor material than the first die.
9. The stacked die system of claim 8, wherein the plurality of second dies are formed on a III-V semiconductor substrate, and the first die is formed on a silicon substrate.
10. The stacked die system of claim 1, wherein the sub regions on the first die are arranged in
10 a three by three pattern, and wherein there are four second dies.
11. A method for forming a stacked die system, comprising:
forming an active device layer on a first semiconductor die;
15 forming a plurality of routing layers over a surface of the active device layer,
wherein the routing layers are segmented into a plurality of sub regions, and wherein no
conductive traces in the plurality of routing layers pass over a boundary between any of
the plurality of sub regions; and
mounting a plurality of second dies onto the first die.
- 20
12. The method of claim 11, wherein the first die has a footprint that is larger than a reticle
limit of the processing used to form the plurality of routing layers.
13. The method of claim 12, wherein forming the routing layers includes individually
25 exposing each sub region with a reticle, and wherein no two sub regions are
simultaneously exposed with a single reticle.

14. The method of claim 11, wherein at least two of the second dies are mounted over at least a portion of the same sub region of the first die.

5

15. The method of claim 14, wherein the two second dies are communicatively coupled to each other by a die to die interconnect formed entirely within a single sub region of the first die.

- 10 16. The method of claim 11, further comprising:
forming a plurality of vias through the first die, wherein the vias communicatively couple the device layer of the first die to a surface of the first die opposite to the device layer; and
mounting the first die to a packaging substrate, wherein the device layer of the
15 first die is formed on a surface of the first die facing away from a surface of the packaging substrate.

17. The method of claim 11, wherein the plurality of second dies are fabricated with advanced processing technology relative to the first die.

20

18. The method of claim 17, wherein the second dies are cores.

19. The method of claim 17, wherein a footprint of each of the second dies is greater than a footprint of each of the sub regions on the first die.

25

20. The method of claim 11, wherein the plurality of sub regions includes nine or more sub

regions, and wherein the plurality of second dies includes four or more second dies.

21. A stacked die system, comprising:

a first die, wherein the first die comprises:

5 a device layer;

a plurality of routing layers formed over the device layer, wherein the plurality of routing layers are segmented into a plurality of sub regions, and wherein no conductive traces in the plurality of routing layers pass over a boundary between any of the plurality of sub regions; and

10 a plurality of vias through the first die, wherein the vias electrically couple the device layer with a surface of the die opposite to the device layer;

a plurality of second dies stacked over the first die, wherein one or more of the second dies cross over the boundary between any of the plurality of sub regions, and wherein at least a first of the second dies are communicatively coupled to at least a second of the second dies by a die to die interconnect formed entirely within a single sub region in the first die; and

15 a packaging substrate communicatively coupled to a the plurality of vias in the first die.

20 22. The stacked die system of claim 21, wherein a footprint of each of the second dies is larger than a footprint of the sub regions of the first die.

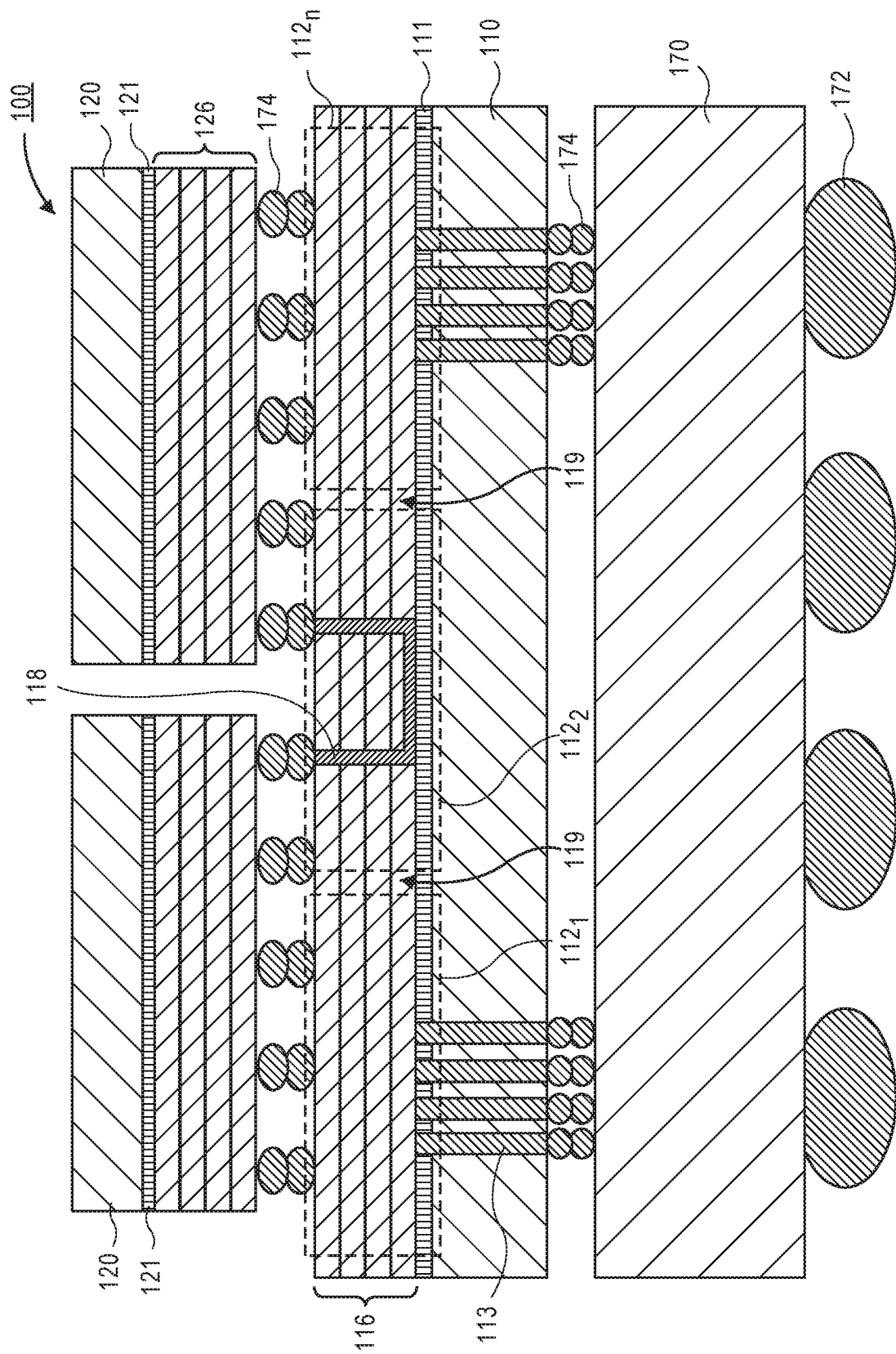
23. The stacked die system of claim 21, wherein the plurality of second dies are high speed processing dies relative to the first die.

25

24. The stacked die system of claim 23, wherein the second dies are processing cores.

25. The stacked die system of claim 21, wherein the sub regions on the first die are arranged

in a three by three pattern, and wherein there are four second dies arranged in a two by two pattern.



ॐ

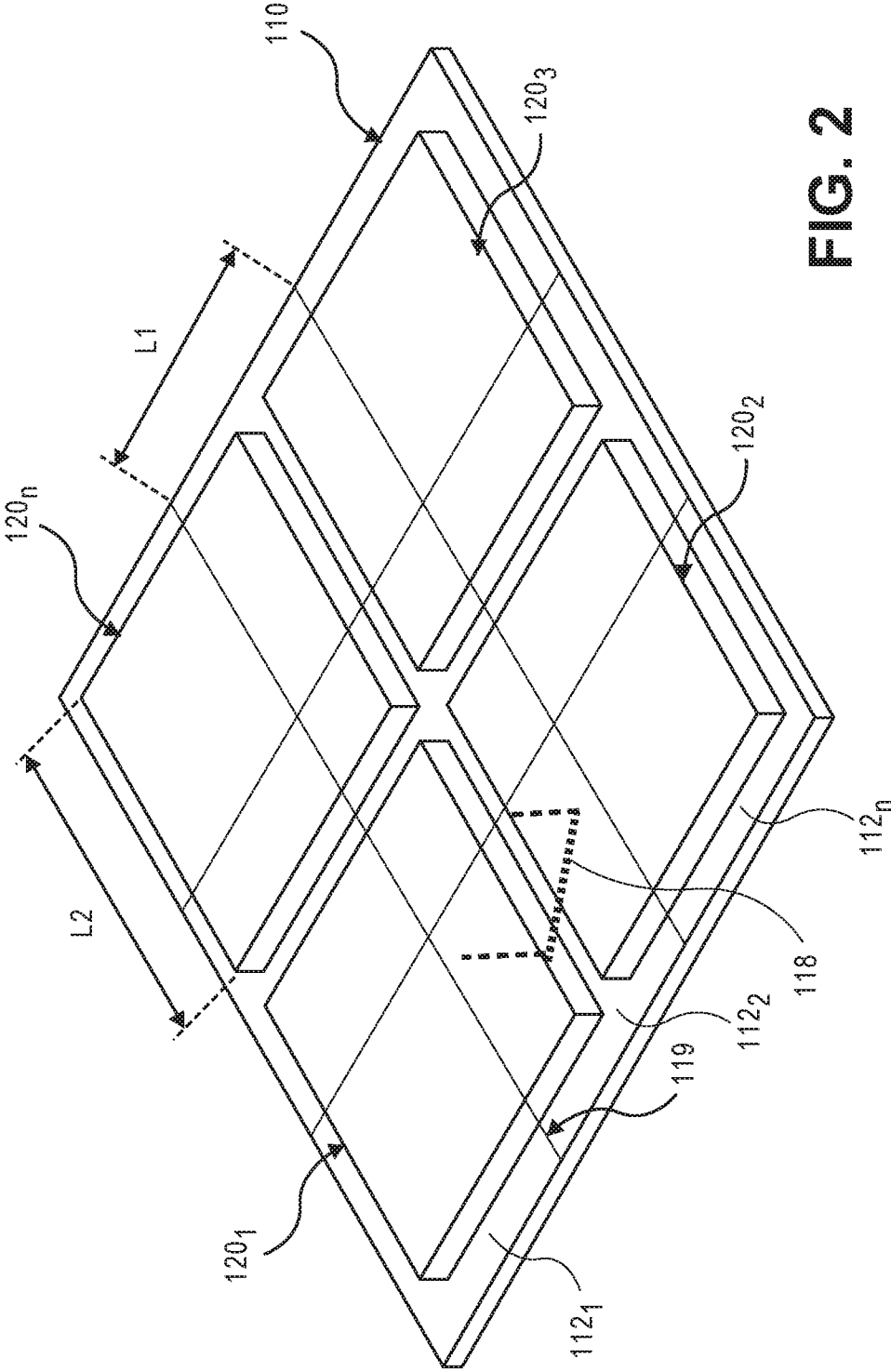


FIG. 2

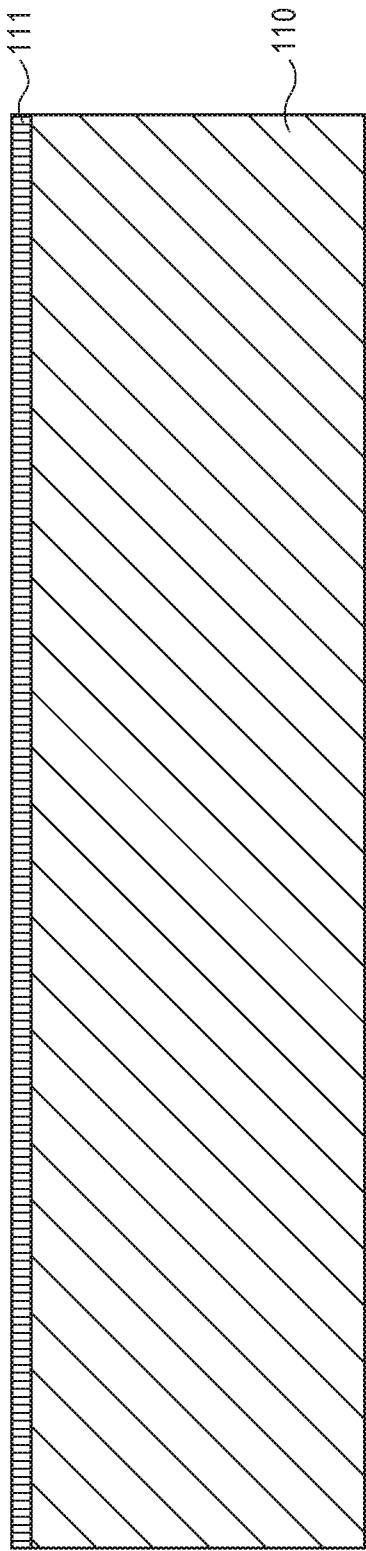


FIG. 3A

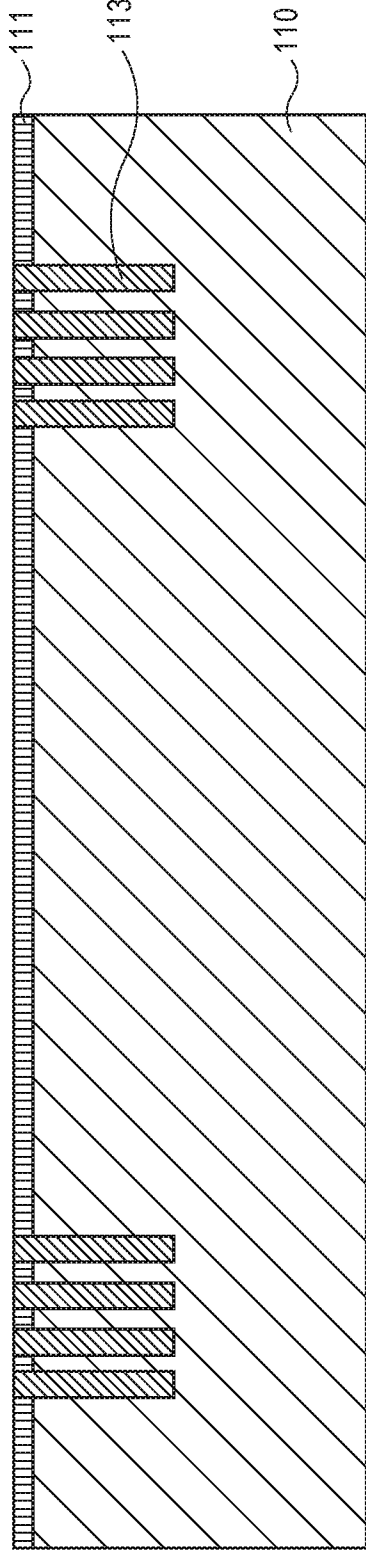


FIG. 3B

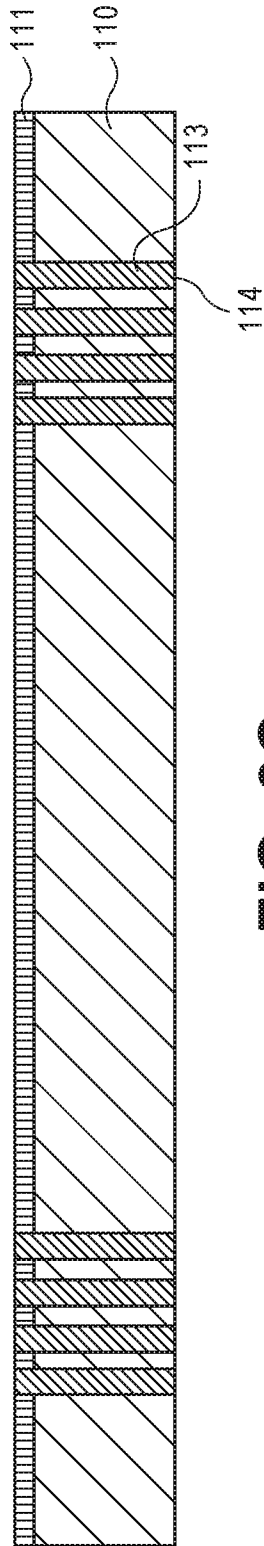


FIG. 3C

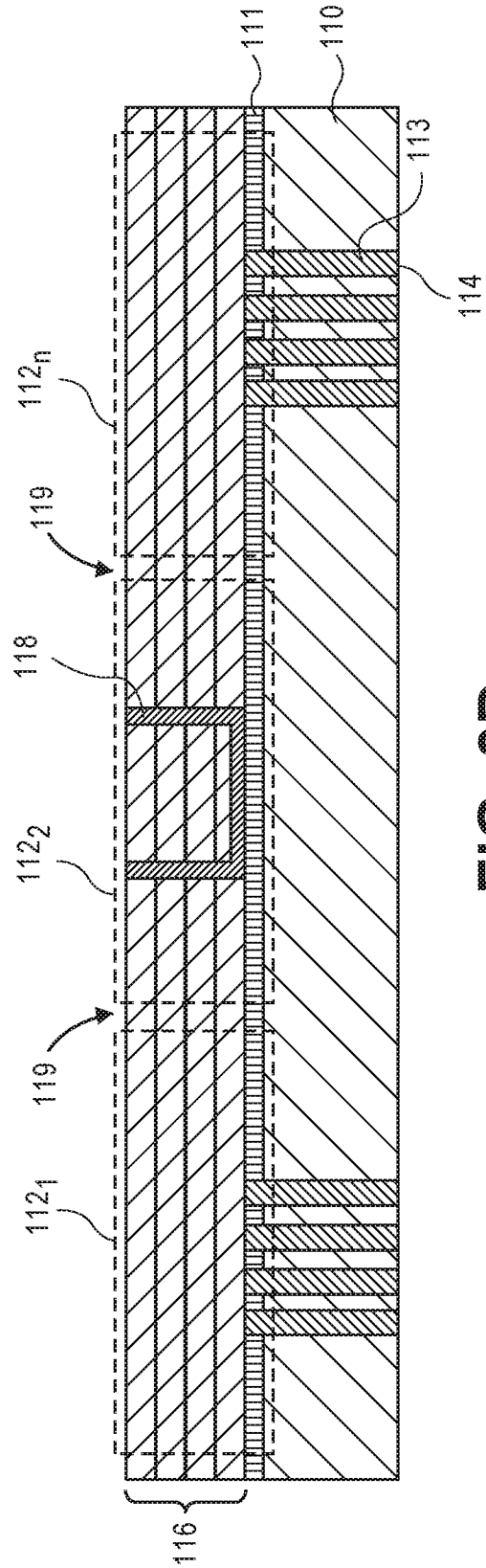


FIG. 3D

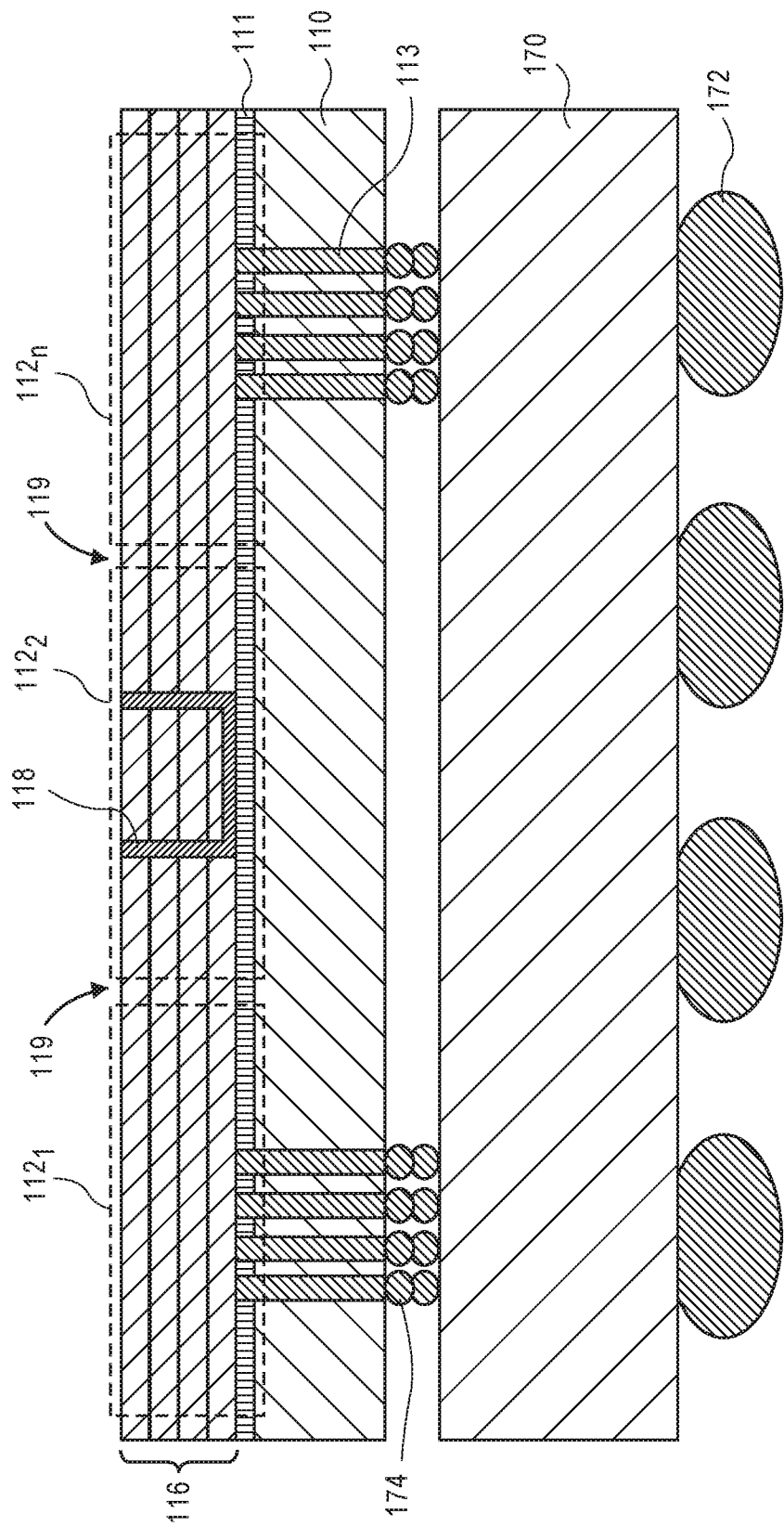


FIG. 3E

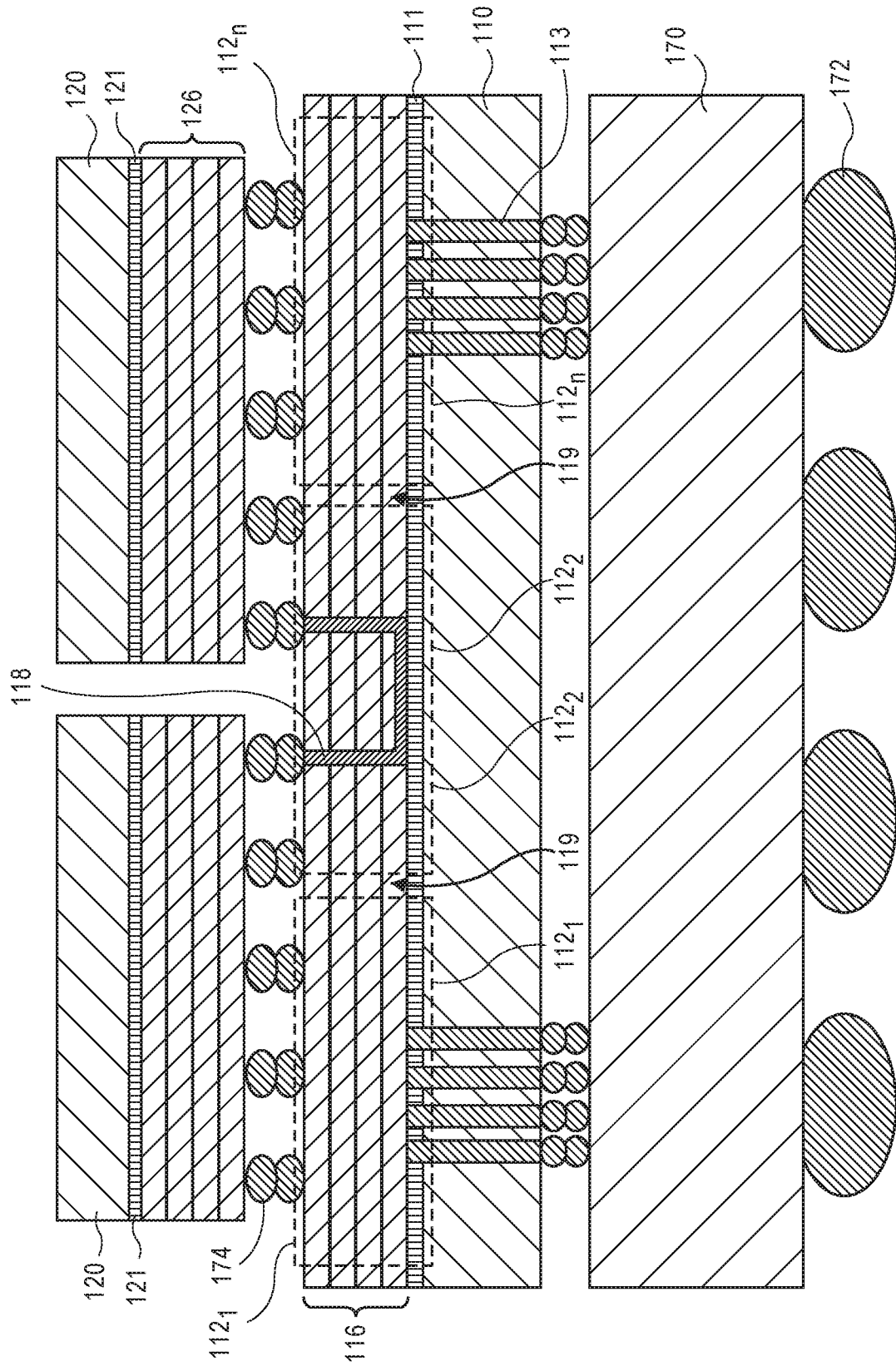
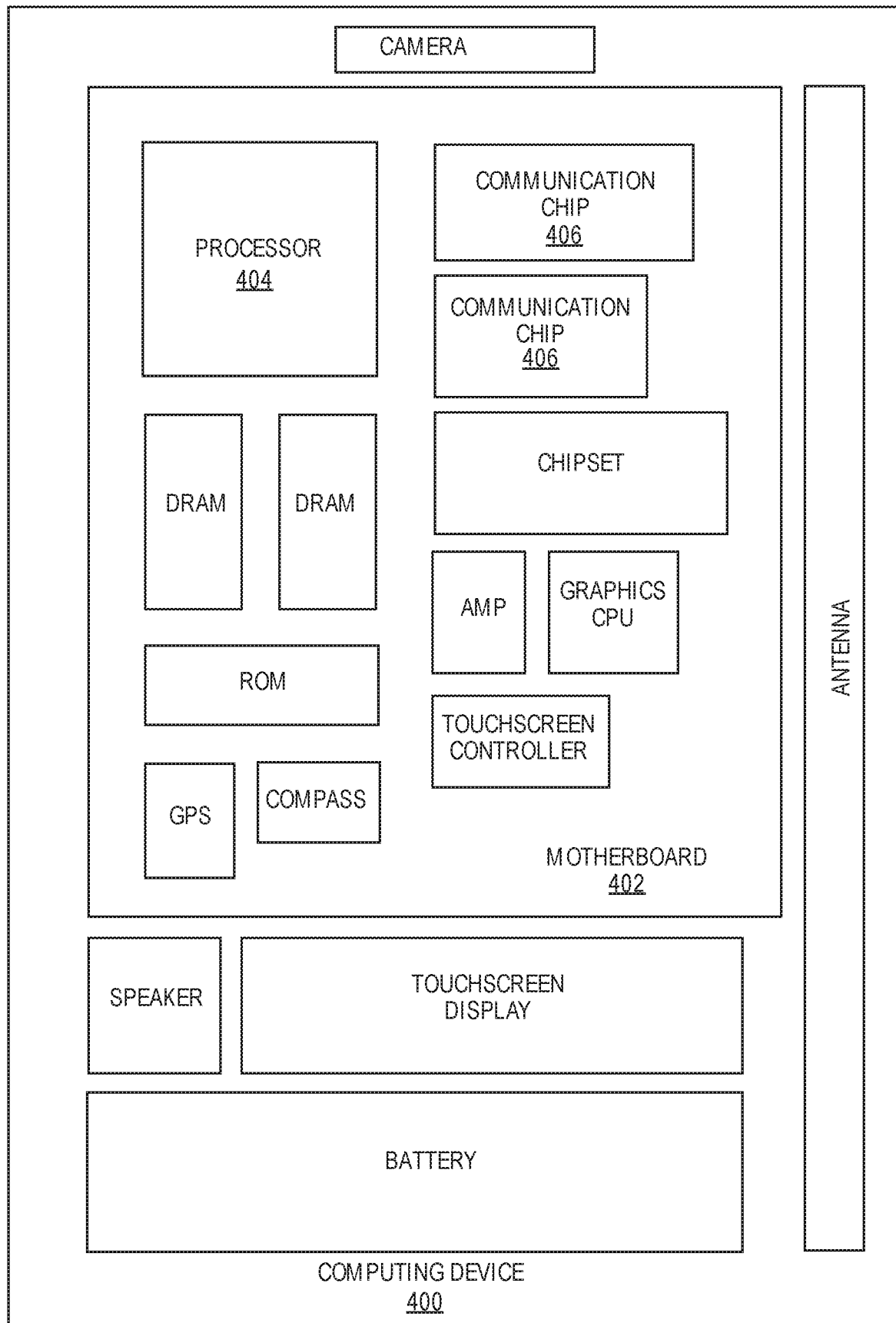


Figure 3

**FIG. 4**

A. CLASSIFICATION OF SUBJECT MATTER**H01L 25/065(2006.01)i, H01L 25/07(2006.01)i, H01L 23/532(2006.01)i, H01L 23/00(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)

H01L 25/065; H01L 23/52; H01L 21/768; H03K 19/173; H01L 23/00; H01L 23/48; H01L 23/538; H01L 25/07; H01L 23/532

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

Korean utility models and applications for utility models

Japanese utility models and applications for utility models

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

eKOMPASS(KIPO internal) & Keywords: second die, device layer, routing layer, segment, sub region, boundary, via, substrate

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 7619441 B1 (RAHMAN et al.) 17 November 2009 See column 7, line 19 - column 8, line 33 and figure 6.	1-25
Y	US 2015-0145141 A1 (INVENSAS CORPORATION) 28 May 2015 See paragraphs [0026]-[0047] and figures 1A-2B.	1-25
A	US 2011-0241208 A1 (MAHAJAN et al.) 06 October 2011 See paragraphs [0016]-[0027] and figures 1-2.	1-25
A	US 9368450 B1 (QUALCOMM INCORPORATED) 14 June 2016 See column 4, line 60 - column 9, line 42 and figures 2-3.	1-25
A	US 2009-0020881 A1 (KIM, SANG CHUL) 22 January 2009 See paragraphs [0028]-[0039] and figures 1-2.	1-25



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

28 November 2017 (28.11.2017)

Date of mailing of the international search report

28 November 2017 (28.11.2017)

Name and mailing address of the ISA/KR

International Application Division

Korean Intellectual Property Office

189 Cheongsa-ro, Seo-gu, Daejeon, 35208, Republic of Korea

Facsimile No. +82-42-481-8578

Authorized officer

CHOI, Sang Won

Telephone No. +82-42-481-8291



INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/US2017/025663

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 7619441 B1	17/11/2009	None	
US 2015-0145141 A1	28/05/2015	US 2015-0145116 A1 US 2015-0380377 A1 US 2016-0049390 A1 US 2016-0307832 A1 US 2017-0154875 A1 US 9263394 B2 US 9379074 B2 US 9583456 B2 US 9728527 B2	28/05/2015 31/12/2015 18/02/2016 20/10/2016 01/06/2017 16/02/2016 28/06/2016 28/02/2017 08/08/2017
US 2011-0241208 A1	06/10/2011	CN 101960589 A CN 101960589 B DE 112009000351 B4 DE 112009000351 T5 DE 112009005519 A5 GB 2470866 A GB 2470866 B JP 2011-515842 A KR 10-1182010 B1 KR 10-2010-0116689 A TW 200950034 A TW I425602 B US 2009-0244874 A1 US 8064224 B2 US 8441809 B2 WO 2009-146007 A2 WO 2009-146007 A3	26/01/2011 10/10/2012 17/07/2014 05/05/2011 31/12/2014 08/12/2010 03/10/2012 19/05/2011 11/09/2012 01/11/2010 01/12/2009 01/02/2014 01/10/2009 22/11/2011 14/05/2013 03/12/2009 21/01/2010
US 9368450 B1	14/06/2016	WO 2017-034641 A1	02/03/2017
US 2009-0020881 A1	22/01/2009	CN 101350344 A CN 101350344 B KR 10-0871794 B1 TW 200910568 A US 7846777 B2	21/01/2009 04/07/2012 02/12/2008 01/03/2009 07/12/2010