

SEMICONDUCTOR DEVICE AND METHOD FOR FORMING THE SAME

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of priority of US Provisional Application No. 61/677,890 filed July 31, 2012, the contents of which being hereby incorporated by reference in its entirety for all purposes.

TECHNICAL FIELD

[0002] Various aspects of this disclosure relate to semiconductor devices and methods for fabricating such devices. More particularly, the various aspects of this disclosure relates to non thermo-compression wafer bonding.

BACKGROUND

[0003] Three-dimensional (3D) integrated circuit (IC) has been identified as a promising technology for complimentary metal-oxide-semiconductor (CMOS) scaling to ensure continuous performance enhancement ("More Moore") and functional diversification ("More than Moore") in a small footprint. A 3D IC is formed by stacking wafers and/or chips and interconnecting them vertically so that they function as a single device. One stacking approach is to bond the chips using patterned metallic medium. Metal bonding is preferred as it allows simultaneous formation of electrical, mechanical and hermetic bonds.

[0004] Metal bonding using copper/tin micro-bump has been developed and applied as the means for chip stacking. However, copper/tin micro-bump has poor

electrical, mechanical and thermal properties due to the formation of inter-metallic compounds. At the same time, copper/tin micro-bump does not scale readily to pitch of less than about 10 μm due to process limitations.

[0005] In response to the demand of reducing pitch size and increasing interconnect density, the technology of stacking of the wafers and/or chips has transformed from eutectic type bonding to diffusion type bonding. Copper to copper bonding with no solder layer, such as tin, is one of the examples of the diffusion type bonding. The copper to copper bonding provides an alternative for future scaling to pitch of less than about 10 μm with improved physical properties. Despite the technical merits, a number of process challenges must be overcome for the copper to copper bonding to be a viable solution, including surface oxidation control, surface uniformity and low temperature bonding of less than about 300 °C and high throughput.

[0006] The copper to copper (Cu-Cu) diffusion type bonding requires parallel thermo-compression application of heat and pressure for Cu-Cu bond formation. Bonding is typically performed at temperature range of 300 °C to 400 °C and contact pressure of several MPa for duration of 30 minutes or longer in vacuum. The fundamental mechanism of the thermo-compression bonding is inter-diffusion of atom and grain growth of two contacted copper surfaces. However, oxidation of the non-inert copper surface forms a barrier for inter-diffusion, hampering the bond formation at low temperature and resulting in poor physical (e.g., particularly electrical) properties due to the incorporation of oxygen in the copper layer. One method of improving thermo-compression bonding is to remove the surface oxide with energetic ions and bond the copper surfaces in ultra-high vacuum (UHV) ambient. However, this method is less

manufacturing worthy. Another method is to remove the surface oxide with wet cleaning such as acetic acid. Nevertheless, prolonged immersion in acetic acid is found to etch the copper layer, making the process control more challenging. Other methods include a non-UHV and non-corrosive method to passivate the copper surface with self-assembled monolayer in order to achieve lower temperature bonding.

[0007] Harsh conditions of the thermo-compression process, such as high temperature and contact pressure, often results in poor alignment accuracy, thermo-mechanical stress and structural integrity of the bonded structure. In addition, the thermo-compression process is diffusion limited, hence bonding does not form instantaneously, negatively impacting the throughput. Moreover, only a pair of wafers can be bonded at any time in a single machine, resulting in higher manufacturing costs.

[0008] Therefore, there is a need to improve wafer bonding technology to facilitate 3D ICs.

SUMMARY

[0009] In various embodiments, a method for forming a semiconductor device may be provided. The method may include providing a substrate. The method may also include forming a metal layer on the substrate. In addition, the method may include forming a first capping layer on the metal layer. The method may also include attaching the first capping layer to a second capping layer.

[0010] In various embodiments, a semiconductor device may be provided. The semiconductor device may include a substrate. The semiconductor device may also

include a metal layer disposed on the substrate. The semiconductor device may also include a capping layer disposed on the metal layer.

[0011] These and other advantages and features of the embodiments herein disclosed, will become apparent through reference to the following description and the accompanying drawings. Furthermore, it is to be understood that the features of the various embodiments described herein are not mutually exclusive and can exist in various combinations and permutations.

BRIEF DESCRIPTION OF THE DRAWINGS

[0012] In the drawings, like reference characters generally refer to the same parts throughout the different views. Also, the drawings are not necessarily to scale, emphasis instead generally being placed upon illustrating the principles of various embodiments. In the following description, various embodiments of the present disclosure are described with reference to the following:

[0013] FIG. 1A shows a schematic illustrating a cross-sectional view of a semiconductor device according to various embodiments;

[0014] FIG. 1B shows a schematic illustrating a cross-sectional view of a semiconductor device according to various embodiments;

[0015] FIG. 2 shows a schematic illustrating a cross-sectional view of a semiconductor device according to various embodiments;

[0016] FIG. 3 shows a schematic illustrating a method for forming a semiconductor device according to various embodiments;

[0017] FIGS. 4A to 4F show a method for forming a semiconductor device according to various embodiments;

[0018] FIGS. 5A to 5E show a method for forming a semiconductor device according to various embodiments;

[0019] FIG. 6 shows voids at bonding interface;

[0020] FIG. 7 shows a void-free bonding interface;

[0021] FIG. 8 shows a graph illustrating an electrical performance comparison of bonded wafers; and

[0022] FIG. 9 shows a graph illustrating a mechanical performance comparison of bonded wafers.

DETAILED DESCRIPTION

[0023] The following detailed description refers to the accompanying drawings that show, by way of illustration, specific details and embodiments in which the invention may be practiced. These embodiments are described in sufficient detail to enable those skilled in the art to practice the invention. Other embodiments may be utilized and structural, and logical changes may be made without departing from the scope of the invention. The various embodiments are not necessarily mutually exclusive, as some embodiments can be combined with one or more other embodiments to form new embodiments.

[0024] In some instances, well-know features are omitted or simplified to clarify the description of the exemplary embodiments of the present invention, and to thereby better explain the present invention. Furthermore, in order that the invention may be

readily understood and put into practical effect, particular embodiments will now be described by way of examples and not limitations, and with reference to the figures.

[0025] It should be understood that the terms “top”, “bottom”, etc., when used in the following description are used for convenience and to aid understanding of relative positions or directions, and not intended to limit the orientation of any device or any part of any device.

[0026] FIG. 1A shows a cross-sectional view of a semiconductor device 100a according to various embodiments. The semiconductor device may include a substrate. The semiconductor device may also include a metal layer disposed on the substrate. The semiconductor device may also include a capping layer disposed on the metal layer.

[0027] In various embodiments, the semiconductor device may include a barrier layer. The barrier layer may be disposed in between the metal layer and substrate.

[0028] In various embodiments, the capping layer may include a metal nitride. In various embodiments, the capping layer may include a metal nitride selected from a group which includes titanium nitride, tantalum nitride, copper nitride and a combination thereof. In various embodiments, the capping layer may include titanium nitride, and/or tantalum nitride, and/or copper nitride, and/or a combination thereof. In various embodiments, the metal layer may include copper, the capping layer may include titanium nitride, and the barrier layer may include titanium.

[0029] In various embodiments, the substrate may include circuit components prepared thereon. The circuit components may include transistors, capacitors as well as other components.

[0030] In various embodiments, the metal layer is an uppermost metal layer of an uppermost metal level.

[0031] In various embodiments, the metal layer may be formed in a dielectric layer using a damascene process, dual-damascene process or a combination thereof. The dielectric layer may be an interlayer dielectric (ILD) layer.

[0032] In various embodiments, the capping layer may include a plasma-activated capping layer.

[0033] The device, for example, may be a semiconductor device, such as a semiconductor die of an integrated circuit (IC). The IC may be any type of IC. For example, the IC may be memory devices such as a dynamic random access memory (DRAM), a static random access memory (SRAM) and various types of non-volatile memories including programmable read only memories (PROM) and flash memories, logic devices, communication devices, optoelectronic devices, digital signal processors (DSPs), microcontrollers, system-on-chips (SOCs) as well as other types of devices or a combination thereof. Other types of ICs or devices may be provided. The ICs can be incorporated into various products, such as phones, computers, personal digital assistants, tablets, cameras or other types of suitable products.

[0034] The device includes a substrate 102. The substrate, for example, may be a semiconductor substrate, such as a silicon substrate. Other types of substrates, for example, SiGe, SiGeC or SiC, may be provided. In other embodiments, the substrate 102 may be a crystalline-on-insulator (COI), such as a silicon-on-insulator (SOI) substrate. Other types of COI substrates may be provided. The substrate 102, for example, may be

a wafer which includes a plurality of devices. The wafer may be diced to singulate the devices.

[0035] A conductive layer 104 may be disposed on the substrate 102. The conductive layer 104, for example, is a metal layer. In one embodiment, the metal layer 104 is disposed on a first surface of the substrate 102. The first surface, for example, may be the top substrate surface. The top substrate surface may be a bare substrate surface. In other embodiments, the top substrate surface may be a processed substrate surface. The substrate 102 may be at any stage of processing. For example, the substrate 102 may be after pre-metal dielectric (PMD) formation, such as prior to forming a first metal layer or after formation of a final passivation layer, such as after formation of a final metal layer.

[0036] In one embodiment, the metal layer 104 may be a copper layer. Other metal material may be provided as the metal layer. The thickness of the metal layer 104, for example, may be about 100 nm. Other suitable thicknesses may be provided. The thickness, for example, may depend on the technology node. The metal layer 104, for example, may be a metal layer of an uppermost metal level. Other configurations of the metal layer may also be useful.

[0037] A capping layer 106 is disposed on the metal layer 104. In one embodiment, the capping layer 106 may be a titanium nitride (TiN) layer. In various embodiments, the capping layer may include other metal nitrides, such as tantalum nitride or copper nitride. The thickness of the capping layer, for example, may be about 10 nm. Other suitable thicknesses may be provided. The thickness, for example, may depend on the type of technology node and process capability.

[0038] The capping layer 106 prevents surface oxidation and functions as a fusion bonding layer. When activated by a plasma process, the surface of the capping layer becomes hydrophilic which is beneficial for forming an instantaneous fusion bonding with another activated capping layer at molecular level in an ambient condition. The hydrophilicity of the surface of the activated capping layer may be measured, for example, using water droplet contact angle measurement in a goniometer. In one embodiment, the surface of the activated capping layer has a low contact angle of less than about 10°.

[0039] In one embodiment, a barrier layer 112 may be disposed in between the metal layer 104 and the substrate 102. FIG. 1B shows a cross-sectional view of an embodiment of a device 100b having a barrier layer. The barrier layer 112 serves to prevent out-diffusion of the conductive material of the metal layer 104. For example, the barrier layer 112 may be a metal diffusion barrier to prevent diffusion of metal from the metal layer 104. The barrier layer 112 may also improve the adhesion of the metal layer 104 to the substrate 102. The thickness of the barrier layer 112, for example, may be about 10 nm. Other suitable thicknesses may be provided. The thickness, for example, may depend on the technology node. In one embodiment, the barrier layer 112 may be refractory metals, refractory metal nitrides, refractory metal silicon nitrides or combinations thereof. For example, the barrier layer 112 may include a refractory material, such as TiN, TaN, Ti, Ta, TiSN, TaSN, Cr, Co, Ni, Pt, CoP, or a combination thereof. Other types of materials that can inhibit diffusion of the conductive material may be provided. In one embodiment, the barrier layer 112 may be a titanium (Ti) layer. In another embodiment, the barrier layer 112 may be a titanium nitride (TiN) layer. In

yet another embodiment, the barrier layer 112 maybe a tantalum (Ta) layer, tantalum nitride (TaN) layer or a combination thereof.

[0040] FIG. 2 shows a cross-sectional view of an embodiment of a device 200. The device, for example, may be a semiconductor device, such as a semiconductor die of an integrated circuit (IC). The IC may be any type of IC. The devices may include circuit components (not shown) formed in and/or on a substrate 202. The circuit components, for example, may include transistors, capacitors as well as other components. The circuit components may be interconnected by interconnects formed in metal or interconnect levels (M1 to M_x, x is the number of metal levels). Contacts may be used to couple interconnects of, for example, adjacent metal levels or interconnects to the circuit components. For example, the contact level that couples interconnects of circuit components of the first metal level M1 is referred to as contact via (CA), while other contact levels disposed between adjacent metal levels are via levels (V1 to V_{x-1}). Typically, a device may have about 3 to 10 (e.g., x = 3 to 10) metal levels. Devices with other number of metal levels may be provided.

[0041] For purposes of discussion, however, only two metal levels (e.g., M1 and M2) are illustrated and described. It will be appreciated that numerous metal levels, other than two metal levels as illustrated and described herein, may be used according to various embodiments.

[0042] Referring to FIG. 2, two metal levels M1 and M2 are disposed in a dielectric material 208. The dielectric material 208, in one embodiment, may be disposed on the substrate 202. The dielectric material 208, for example, may include a plurality of dielectric layers which serve as inter-level dielectric (ILD) layers. Etch stop layers (not

shown) may be provided between ILD layers. Other configuration of ILD or dielectric layers may also be useful. In one embodiment, the dielectric material may be silicon dioxide. Other types of materials may be provided, such as borophosphosilicate glass (BPSG), phosphosilicate (PSG), high density plasma (HDP) oxide, high aspect ratio process (HARP) oxide, tetraethylorthosilicate (TEOS), carbon-doped oxide (CDO) or a combination therefore.

[0043] Each of the metal level includes contact in a contact level and interconnects in an interconnect level. In one embodiment, the contact and interconnects are conductive layers, such as metal layers. The metal layers, in one embodiment, may be copper layers. Providing other conductive material as the metal layers may also be useful. The thickness of the metal layer, for example, may be about 100 nm. Other suitable thicknesses may also be useful. The thickness, for example, may depend on the technology node.

[0044] The metal layers, for example, may be formed by damascene or dual damascene techniques. Damascene technique includes forming openings in a dielectric layer which are filled with a conductive material. Excess conductive material may be removed by, for example, chemical mechanical polishing. This forms contacts in the contact level or interconnects in the interconnect level. Dual damascene technique includes forming in the dielectric layer openings corresponding to contact openings and trench openings corresponding to interconnects, which are filled by a conductive material. Excess conductive material may be removed by, for example, chemical mechanical polishing. This forms contacts in the contact level and interconnect in the interconnect level in a single process.

[0045] In some embodiments, the contacts and interconnects may be formed by a combination of damascene and dual damascene techniques. For example, the CA and M1 levels are formed using damascene techniques while the other level is formed using dual damascene techniques. Other techniques or combination of techniques may be employed to form the contact and interconnect levels.

[0046] For a dual damascene technique, the conductive material of the contacts and interconnects are the same. For example, the conductive material of the contacts and interconnects of an uppermost metal level M2 may be copper (Cu). Damascene technique allows for the use of different conductive materials for contacts and interconnects. For example, in the first (or lowermost) metal level M1, the contact CA may be tungsten (W) and the interconnects may be copper (Cu). Other configuration of contact and interconnect materials may be provided.

[0047] Interconnects at the uppermost metal level M2 is an uppermost metal layer 204. In one embodiment, the metal layer 204 may be a copper layer. The metal layer, in one embodiment, has a top surface above a top surface of the ILD layer 208. The height of the top surface of the metal layer from the top surface of the ILD layer may be about 400 nm. Other heights may be provided. In one embodiment, the exposed top surface of the metal layer 204 is provided for attaching the substrate 202 to another substrate for realizing 3D ICs.

[0048] An etch stop layer 210 may be disposed on the ILD layer 208, leaving the metal layer 204 uncovered. In one embodiment, the etch stop layer 210 may be a silicon nitride (SiN) layer. Providing etch stop layer of other materials, such as silicon carbide (SiC) or nitrogen-doped silicon carbide (NBLOK) may also be useful. The thickness of

the etch stop layer 210, for example, may be about 10 nm. Other suitable thicknesses may be provided. The thickness, for example, may depend on the technology node. The etch stop layer 210, for example, has high selectivity with respect to the underlying ILD layer.

[0049] In one embodiment, the metal layer 204 is surrounded by a barrier layer (not shown). The barrier layer serves to prevent out-diffusion of the conductive material of the metal layer. For example, the barrier layer may be a copper diffusion barrier to prevent diffusion of copper from the metal layer. The barrier layer may also improve the adhesion of the metal layer to the ILD layer. The thickness of the barrier layer, for example, may be about 10 nm. Barrier layer with other suitable thicknesses may be provided. The thickness, for example, may depend on the technology node. In one embodiment, the barrier layer may be refractory metals, refractory metal nitrides, refractory metal silicon nitrides and combinations thereof. For example, the barrier layer may include a refractory material, such as TiN, TaN, Ti, Ta, TiSN, TaSN, Cr, Co, Ni, Pt, CoP, or a combination thereof. Other types of materials that can inhibit diffusion of the conductive material into the ILD layers and substrate may be provided. In one embodiment, the barrier layer may be a titanium (Ti) layer. In another embodiment, the barrier layer may be a titanium nitride (TiN) layer. In yet another embodiment, the barrier liner maybe a tantalum (Ta) layer, tantalum nitride (TaN) layer or a combination thereof.

[0050] A capping layer 206 is disposed on the exposed metal layer 204. In one embodiment, the capping layer 206 may be a titanium nitride (TiN) layer. In various embodiments, the capping layer may include other metal nitrides, such as tantalum nitride

or copper nitride. The thickness of the capping layer, for example, may be about 10 nm. Other suitable thicknesses may be provided. The thickness, for example, may depend on the technology node.

[0051] The capping layer 206 prevents surface oxidation and functions as a fusion bonding layer. When activated by a plasma process, the surface of the capping layer becomes hydrophilic which is beneficial for forming an instantaneous fusion bonding with another activated capping layer at molecular level in an ambient condition. The hydrophilicity of the surface of the activated capping layer may be measured, for example, using water droplet contact angle measurement in a goniometer. In one embodiment, the surface of the activated capping layer has a low contact angle of less than about 10°.

[0052] FIG. 3 shows a schematic 300 of a method for forming a device according to various embodiments. The method may include, in 302, providing a substrate. The method may also include, in 304, forming a metal layer on the substrate. The method may further include, in 306, forming a first capping layer on the metal layer. The method may also include, in 308, attaching the first capping layer to a second capping layer.

[0053] In various embodiments, the method may include forming a barrier layer in between the metal layer and substrate. The barrier layer serves to prevent out-diffusion of the conductive material of the metal layer. The barrier layer may also enhance the adhesion between the metal layer and substrate.

[0054] In various embodiments, the metal layer may include copper, the first and second capping layers may include the same metal nitride. In various embodiments, at least one of the first capping layer and second capping layer may include a metal nitride.

In various embodiments, at least one of the first capping layer and the second capping layer may include a metal nitride selected from a group which includes titanium nitride, tantalum nitride, copper nitride and a combination thereof. In various embodiments, at least one of the first capping layer and the second capping layer may include titanium nitride, and/or tantalum nitride, and/or copper nitride, and/or a combination thereof. In various embodiments, the first and second capping layers may include titanium nitride.

[0055] In various embodiments, the method may include activating the first and second capping layers. The first and second capping layers may be activated by, for example, a plasma process. The plasma may include oxygen, nitrogen, argon or a combination thereof. Bringing the activated first and second capping layers in close proximity initiates instantaneous fusion bonding. The fusion bonding is based on surface energy and does not depend on inter-diffusion which is time consuming.

[0056] In various embodiments, the method may include annealing the substrate after attaching the first and second capping layers. Annealing the substrate, post-bonding, enhances the bond strength between the first and second capping layers which are in an intimate contact. In one embodiment, the post-bonding annealing may be carried out in a batch process.

[0057] In various embodiments, the second capping layer may be disposed on a second substrate. The second capping layer disposed on the second substrate may be formed using a similar method of forming the first capping layer.

[0058] In various embodiments, attaching the first capping layer to the second capping layer may be performed at room temperature in an atmospheric ambient.

[0059] In various embodiments, the substrate may include circuit components prepared thereon. The circuit components may include transistors, capacitors as well as other components.

[0060] In various embodiments, the metal layer may be an uppermost metal layer of an uppermost metal layer. The metal layer may be disposed in a dielectric layer formed on the substrate. The dielectric layer may be an interlayer dielectric (ILD) layer.

[0061] In various embodiments, forming the metal layer may include forming the metal layer in a dielectric layer using a damascene process, dual-damascene process or a combination thereof.

[0062] FIGS. 4A to 4F illustrate a method for forming a device according to various embodiments. The device may be similar to that described in FIG. 1. Similar elements may not be described or described in detail.

[0063] FIG. 4A shows a schematic 400a illustrating a cross sectional view of a substrate 402. The substrate, for example, may be a semiconductor substrate, such as a silicon substrate.

[0064] FIG. 4B shows a schematic 400b illustrating a cross sectional view of a metal layer 404 formed on the substrate 402. The metal layer 404, in one embodiment, is a copper layer. The metal layer 404 may be formed by plating. For example, the metal layer 404 may be formed by electro or electro-less plating. A seed layer may be formed over the substrate 402 to facilitate forming the metal layer 404. Other techniques for forming the metal layer, such as chemical vapor deposition (CVD) or physical vapor deposition (PVD) may be provided. A planarizing process is performed after forming the metal layer 404. The planarizing process, in one embodiment, is a polishing process. For

example, the polishing process is a chemical mechanical polishing (CMP) process. The polishing process produces a planar or substantially planar top surface of the metal layer 404.

[0065] In one embodiment, a barrier layer (not shown) may be formed in between the substrate 402 and metal layer 404. The configuration of the barrier layer, for example, may be similar to the configuration of the barrier layer 112 as shown in FIG. 1B. As previously described, the substrate may be at any stage of processing. In one embodiment, the barrier layer may be a titanium (Ti) layer. Other types of barrier layer, such as TiN, Ta, TaN, TiSN, TaSN, Cr, Co, Ni, Pt, CoP may be provided. The barrier layer may be formed on the substrate 402 by, for example, physical vapor deposition (PVD) such as sputtering. Other techniques for forming the barrier layer, such as chemical vapor deposition (CVD) may be provided. A planarizing process is performed after forming the barrier layer. The planarizing process, in one embodiment, is a polishing process. For example, the polishing process is a chemical mechanical polishing (CMP) process. The polishing process produces a planar or substantially planar top surface of the barrier layer.

[0066] FIG. 4C shows a schematic 400c illustrating a cross sectional view of a capping layer 406 formed on the metal layer 404. The capping layer 406, in one embodiment, is a titanium nitride (TiN) layer. In various embodiments, the capping layer may include other metal nitrides, such as tantalum nitride or copper nitride. The capping layer 406 may be formed by physical vapor deposition (PVD), such as sputtering. Other techniques for forming the capping layer, such as chemical vapor deposition (CVD), atomic layer deposition (ALD) or in-situ process, may be provided.

[0067] FIG. 4D shows a schematic 400d illustrating a cross sectional view of the capping layer 406 subjected to an activation process. In one embodiment, the capping layer is activated in plasma 412. Other activation process may be provided. The activation using plasma produces an activated capping layer 416 with a hydrophilic surface having a contact angle of, for example, less than about 10° . The plasma activation may be carried out at low pressure in a vacuum of about 0.1 Pa to 100 Pa for duration of about 15 s. In one embodiment, the source gases of the plasma may include oxygen. In another embodiment, the source gases of the plasma may include nitrogen or argon. In yet another embodiment, the source gases of the plasma may include oxygen and nitrogen. Other types of source gases may be provided.

[0068] In one embodiment, wet cleaning is performed to the activated capping layer 416. The wet cleaning, in one embodiment, includes rinsing the substrate in de-ionized water for duration of about 1 min.

[0069] FIG. 4E shows a schematic 400e of a cross sectional view of attaching the activated capping layer 416 to a second activated capping layer 426. Bringing the activated capping layers (416 and 426) in close proximity initiates instantaneous fusion bonding. As shown in FIG. 4E, a gap 450 demonstrates the bonding process just prior to an intimate contact of the activated capping layers (416 and 426). In one embodiment, the second activated capping layer 426 may be formed on a second metal layer 424 which is formed on a second substrate 422 (e.g., donor wafer) using a similar method of forming the activated capping layer 416 on the substrate 402 (e.g., substrate wafer). The second metal layer 424, for example, may be a copper layer. In one embodiment, the second capping layer 426 may include the same metal nitride as the first capping layer 416. In

various embodiments, the first capping layer and the second capping layer may include different metal nitrides. The second capping layer 426, for example, may include titanium nitride. In various embodiments, the second capping layer 426 may include other metal nitrides, such as tantalum nitride or copper nitride. In one embodiment, attaching of the capping layers (416 and 426) may be performed at room temperature in an atmospheric ambient.

[0070] FIG. 4F shows a schematic 400f of a cross sectional view of the attached activated capping layers (416 and 426) in an intimate contact. The bonding strength of the attached activated capping layers that are in an intimate contact may be enhanced with a post-bonding annealing. In one embodiment, the post-bonding annealing may be carried out in a batch process. For example, the post-bonding annealing is performed at about 300 °C for duration of about 1 hour.

[0071] Various processes may be performed to complete the method for forming the device. This for example, includes forming connections to the components of the device, dicing and packaging. Other processes may also be included, depending on the stage of processing or type of device.

[0072] The present disclosure provides a method of indirect copper to copper bonding which can be carried out at room temperature in an atmospheric ambient. Furthermore, sufficiently low bonding pressure is needed to initiate contact and such initiation can occur at a substantially shorter duration (e.g., less than about 1 min). These features are in contrast with features of the thermo-compression (TC) method which requires at least several MPa bonding pressure, a longer initiation duration of more than 30 min and bonding temperature of about 300 °C to 400 °C.

[0073] FIGS. 5A to 5E illustrate a method for forming a device according to various embodiments. The device may be similar to that described in FIG. 2. Similar elements may not be described or described in detail.

[0074] FIG. 5A shows a schematic 500a of a cross sectional view of a substrate 502 having various features or elements formed thereon, including a metal layer 504. The substrate 502, for example, may be a semiconductor substrate, such as a silicon substrate. The devices may include circuit components (not shown) formed in and/or on the substrate 502. The circuit components, for example, may include transistors, capacitors as well as other components. The substrate 502 may be at any stage of processing. For example, the substrate may be after formation of an uppermost metal level. For purposes of discussion, however, only two metal levels are illustrated and described. It will be appreciated that numerous metal levels, other than two metal levels as illustrated and described herein, may be used according to various embodiments.

[0075] The device is prepared with two metal levels (M1 and M2) in an inter-level dielectric (ILD) layer 508. In one embodiment, interconnects at the uppermost metal level (M2) is a metal layer 504. The metal layer 504, for example, may be a copper layer. The metal levels may be formed using damascene and/or dual damascene techniques which are described previously with respect to FIG. 2. Other techniques of forming the metal levels may be provided.

[0076] The metal layer 504, in one embodiment, has a top surface above a top surface of the ILD layer 508. The height of the top surface of the metal layer from the top surface of the ILD layer may be about 400 nm. Other heights may be provided. In

one embodiment, the exposed top surface of the metal layer 504 is provided for attaching the substrate 502 to another substrate for realizing 3D ICs.

[0077] An etch stop layer 510 may be formed on the ILD layer 508, leaving the metal layer 504 uncovered. In one embodiment, the etch stop layer 510 may be a silicon nitride (SiN) layer. Other materials, such as silicon carbide (SiC) or NBLOK may be provided. The thickness of the etch stop layer 510, for example, may be about 10 nm. Other thicknesses may be provided. The etch stop layer 510, for example, has high selectivity with respect to the underlying ILD layer 508. In one embodiment, the etch stop layer 510 may be formed by chemical vapour deposition (CVD). Other methods of forming the etch stop layer 510 may be provided.

[0078] In one embodiment, the metal layer 504 is surrounded by a barrier layer (not shown). The barrier layer serves to prevent out-diffusion of the conductive material of the metal layer 504. The barrier layer may also improve the adhesion of the metal layer 504 to the ILD layer 508. The thickness of the barrier layer, for example, may be about 10 nm. Other thicknesses may be provided. In one embodiment, the barrier layer may be a titanium (Ti) layer. Other materials, such as titanium nitride (TiN) or tantalum (Ta) may be provided. In one embodiment, the barrier layer may be formed by physical vapour deposition (PVD), such as sputtering. Other methods of forming the barrier layer may be provided.

[0079] FIG. 5B shows a schematic 500b of a cross sectional view of a substrate 502 having various features or elements formed thereon, including a capping layer 506 formed on the metal layer 504. The capping layer 506, in one embodiment, is a titanium nitride (TiN) layer. In various embodiments, the capping layer may include other metal

nitrides, such as tantalum nitride or copper nitride. The capping layer 506 may be formed by physical vapor deposition (PVD), such as sputtering. Other techniques for forming the capping layer, such as chemical vapor deposition (CVD), atomic layer deposition (ALD) or in-situ process may be provided.

[0080] FIG. 5C shows a schematic 500c illustrating a cross sectional view of the capping layer 506 subjected to an activation process. In one embodiment, the capping layer is activated in plasma 512. Other activation process may be provided. The activation using plasma produces an activated capping layer 516 with a hydrophilic surface having a contact angle of, for example, less than about 10° . The plasma activation may be carried out at low pressure in a vacuum of about 0.1 Pa to 100 Pa for duration of about 15 s. In one embodiment, the source gases of the plasma may include oxygen. In another embodiment, the source gases of the plasma may include nitrogen or argon. In yet another embodiment, the source gases of the plasma may include oxygen and nitrogen. Other types of source gases may be provided.

[0081] In one embodiment, wet cleaning is performed on the activated capping layer 516. The wet cleaning, in one embodiment, includes rinsing the substrate in de-ionized water for duration of about 1 min.

[0082] FIG. 5D shows a schematic 500d of a cross sectional view of attaching the activated capping layer 516 to a second activated capping layer 526. Bringing the activated capping layers (516 and 526) in close proximity initiates instantaneous fusion bonding. As shown in FIG. 5D, a gap 550 demonstrates the bonding process just prior to an intimate contact of the activated capping layers (516 and 526). In one embodiment, the second activated capping layer 526 may be formed on a second metal layer 524 which

is formed on a second substrate 522 (e.g., donor wafer) using a similar method of forming the activated capping layer 516 on the substrate 502 (e.g., substrate wafer). The second metal layer 524, for example, may be a copper layer. In one embodiment, the second capping layer 526 may include the same metal nitride as the first capping layer 516. In various embodiments, the first capping layer and the second capping layer may include different metal nitrides. The second capping layer 526, for example, may include titanium nitride. In various embodiments, the second capping layer may include other metal nitrides, such as tantalum nitride or copper nitride. In one embodiment, attaching of the capping layers (516 and 526) may be performed at room temperature in an atmospheric ambient.

[0083] FIG. 5E shows a schematic 500e of a cross sectional view of the attached activated capping layers (516 and 526) in an intimate contact. The bonding strength of the attached activated capping layers that are in an intimate contact may be enhanced with a post-bonding annealing. For example, the post-bonding annealing is performed at about 300 °C for duration of about 1 hour. In one embodiment, the post-bonding annealing may be carried out in a batch process.

[0084] Various processes may be performed to complete the method for forming the device. This for example, includes forming connections to the components of the device, dicing and packaging. Other processes may also be included, depending on the stage of processing or type of device.

[0085] The present disclosure provides a method of indirect copper to copper bonding which can be carried out at room temperature in an atmospheric ambient. Furthermore, sufficiently low bonding pressure is needed to initiate contact and such

initiation can occur at a substantially shorter duration (e.g., less than about 1 min). These features are in contrast with features of the thermo-compression (TC) method which requires at least several MPa bonding pressure, a longer initiation duration of more than 30 min and bonding temperature of about 300 °C to 400 °C.

[0086] The following description provides some experimental results of the present disclosure according to various embodiments and thermo-compression method.

[0087] FIG. 6 is a transmission electron microscopy (TEM) picture 600 showing the interface of the bonded copper layers (604 and 624) using thermo-compression method. The presence of voids 630 at the bonding interface 632 is observed. As described above, the oxidation of copper forms a barrier for effective inter-diffusion. As a result, voids are formed at the bonding interface. The presence of the interfacial voids can adversely degrade the mechanical integrity and electrical reliability of the bonded layer, resulting in pre-mature failure of the devices.

[0088] FIG. 7 is a TEM picture 700 showing the interface of the bonded capping layers (716 and 726) using the method described according to various embodiments. As shown, the present indirect copper to copper bonding facilitated by capping layers (716 and 726) formed on the copper layers (704 and 724) produces a seamless (void-free) bonding interface 732.

[0089] Fig. 8 is a graph 800 showing the comparison of electrical performance of bonded wafers using the method according to various embodiments and thermo-compression method. The wafers bonded using the method according to various embodiments exhibit no differences in terms of electrical properties as compared to the wafers bonded using the thermo-compression method.

[0090] FIG. 9 is a graph 900 showing the comparison of mechanical performance of bonded wafers using the method according to various embodiments and thermo-compression method. The wafers bonded using the method according to various embodiments exhibit better mechanical performance in terms of toughness as compared to the wafers bonded using the thermo-compression method.

[0091] As described above, the various embodiments provide a low temperature method of stacking wafers at atmospheric ambient with no requirement on vacuum chamber. In one embodiment, the method may be carried out at room temperature without any *in-situ* heating which leads to lower thermal budget and ease of manufacturing. Furthermore, lower temperature also ensures better wafer to wafer alignment accuracy as a result of lower thermo-mechanical stress. Such an improvement in wafer alignment is always desired as finer pitch can be bonded and larger wafer size can be scaled easily.

[0092] According to various embodiments, two substrate surfaces are brought into contact after surface activation. The surface energy initiates an instantaneous fusion bonding based on molecular force. The bonding wave propagates across the entire substrate. Therefore, there is no need for excessive external contact force to initiate the bonding. Furthermore, the instantaneous initiation of the bonding enables the bonding process to be completed in less than about 1 minute. In the absence of high external force, the integrity of the wafer and the delicate structures on it are preserved. In contrast, thermo-compression method requires that high external contact force must be applied in parallel with high temperature to cause the two substrate surfaces to initiate bonding. In addition, the bonding process of such thermo-compression method can be

lengthy, for example, it may require more than 30 minutes for ensuring sufficient diffusion. Therefore, the various embodiments as described increase the throughput to at least 10 times as compared to the thermo-compression method.

[0093] In wafer to wafer bonding, chips of the same size on both wafers are often required, avoiding the wastage of the wafer area. This inflexibility imposes constraints on circuit design. Furthermore, the yield is compounded. For example, one bad die on a wafer will result in the failure of the whole bonded wafer stack. Die to wafer bonding, on the other hand, offers higher degree of flexibility and only known good dies (KGD) are bonded. Therefore, with the at least 10 times increase in throughput, the various embodiments as described provide an effective platform for the time-consuming die to wafer bonding.

[0094] The application of the various embodiments of the non thermo-compression bonding method using the capping layer to facilitate copper to copper bonding can be readily detected using existing failure analysis widely used in the semiconductor industry. Transmission electron microscopy (TEM) and in-situ energy dispersive X-ray spectroscopy (EDX) are able to verify the use of titanium nitride (TiN) or other metal nitride as the capping layer for bonding.

[0095] The various embodiments as described enable applications for 3D IC wafer to wafer and chip to wafer stacking. In particular, the various embodiments may be used in, but not limited to, 2.5D IC based on silicon interposer, high density memory stack, high bandwidth processor/memory stacking, heterogeneous stacking such as sensor/electronics (e.g., image sensor/readout circuit, microelectromechanical system (MEMS) sensor/readout circuit). In addition, the various embodiments may also

potentially applicable for other microsystem packaging applications such as MEMS and light-emitting diode (LED).

[0096] The various embodiments may be used by integrated device manufacturers (IDMs), foundries, outsourced assembly and test (OSAT) companies, and intellectual property companies. In addition, the various embodiments may also be used by fab-less companies in the mobile and high performance space.

[0097] A market survey report published in 2010 shows that the global 3D IC market is expected to grow from \$2.21 billion in 2009 to \$6.55 billion in 2016 at a Compound Annual Growth Rate (CAGR) of 16.9 % from 2011 to 2016. It is also estimated that 3D IC bonding represents a significant fraction of total 3D IC cost of about 41 %. This has called for bonding method with higher throughput and the various embodiments as described can serve this purpose well.

[0098] The disclosure may be embodied in other specific forms without departing from the spirit or essential characteristics thereof. The foregoing embodiments, therefore, are to be considered in all respects illustrative rather than limiting the disclosure described herein. Scope of the disclosure is thus indicated by the appended claims, rather than by the foregoing description, and all changes that come within the meaning and range of equivalency of the claims are intended to be embraced therein.

CLAIMS

1. A method for forming a semiconductor device comprising:
providing a substrate;
forming a metal layer on the substrate;
forming a first bonding layer on the metal layer; and
attaching the first bonding layer to a second bonding layer.
wherein at least one of the first bonding layer and the second bonding layer comprises a metal nitride selected from a group consisting of titanium nitride, tantalum nitride, copper nitride and a combination thereof.
2. The method of claim 1, further comprising:
forming a barrier layer in between the metal layer and substrate.
3. The method of claim 2,
wherein the metal layer comprises copper.
4. The method of claim 1, wherein the first and second bonding layers comprise the same metal nitride.
5. The method of claim 4, wherein the first and second bonding layers comprise titanium nitride.

6. The method of claim 5, further comprising activating the first and second bonding layers.
7. The method of claim 6, wherein activating the first and second bonding layers comprises activating the first and second bonding layers in plasma.
8. The method of claim 7, wherein the plasma comprises oxygen, nitrogen, argon or a combination thereof.
9. The method of claim 8, further comprising:
annealing the substrate after attaching the first and second bonding layers.
10. The method of claim 9, wherein the second bonding layer is disposed on a second substrate.
11. The method of claim 1, wherein attaching the first bonding layer to the second bonding layer is performed at room temperature in an atmospheric ambient.
12. The method of claim 1, wherein the substrate comprises circuit components prepared thereon.

13. The method of claim 1, wherein the metal layer is an uppermost metal layer of an uppermost metal level.
14. The method of claim 13, wherein forming the metal layer comprises forming the metal layer in a dielectric layer using a damascene process, dual-damascene process or a combination thereof.
15. The method of claim 14, further comprising:
forming a barrier layer in between the metal layer and the dielectric layer.
16. A semiconductor device comprising:
a substrate;
a metal layer disposed on the substrate; and
a bonding layer disposed on the metal layer.
wherein the bonding layer comprises a metal nitride selected from a group consisting of titanium nitride, tantalum nitride, copper nitride and a combination thereof
17. The semiconductor device of claim 16, further comprising:
a barrier layer disposed in between the metal layer and substrate.
18. The semiconductor device of claim 16,
wherein the metal layer comprises copper; and
wherein the bonding layer comprises titanium nitride.

19. The semiconductor device of claim 16, wherein the substrate comprises circuit components prepared thereon.
20. The semiconductor device of claim 16, wherein the metal layer is an uppermost metal layer of an uppermost metal level.
21. The semiconductor device of claim 20, wherein the metal layer is formed in a dielectric layer using a damascene process, dual-damascene process or a combination thereof.
22. The semiconductor device of claim 16, wherein the bonding layer comprises a plasma-activated bonding layer.