

EXHIBIT 803-A-11

Fan-out Mediatek Innovation Package (“Mediatek FOMIP”)

Fan-out Mediatek Innovation Package (“Mediatek FOMIP”) as described by the below references anticipates and/or renders obvious (alone or in combination with other references), including at least under Complainant’s apparent infringement theory, claims 1-3, 7-10 of U.S. Patent No. 11,841,803 (“the ’803 patent”) under 35 U.S.C. §§ 102 and/or 103.

- John Lau, Heterogeneous Integrations, Springer (2019) (“Lau”)
- Nan-Cheng Chen et. al., A Novel System in Package with Fan-out WLP for high speed SERDES application, IEEE 66th Electronic Components and Technology Conference (2016) (“Chen-2016”)
- Kiyoshi Oi et. al., Development of New 2.5D Package with Novel Integrated Organic Interposer Substrate with Ultra-fine Wiring and High Density Bumps, Electronic Components & Technology Conference (2014) (“Oi”)

NVIDIA has charted the portions of the references that correspond to each element, including at least under Complainant’s apparent interpretation of the claims as reflected in Complainant’s infringement contentions. Where the chart below states that the reference “discloses,” “teaches,” or “suggests” a claim limitation, such disclosure, teaching, or suggestion may be express, implied, or inherent. The cited evidence is merely illustrative, and NVIDIA reserves the right to cite alternative or additional evidence.

No part of this exemplary chart construes, or is intended to construe, the specification, file history, or claims of the asserted patents. In addition, nothing stated in this chart shall be treated as an admission or suggestion that NVIDIA agrees with Complainant regarding either the scope of any of the asserted claims or Complainant’s apparent interpretation of the scope of the asserted claims, or that NVIDIA’s accused products meet any limitations of the claims.

To the extent that Complainant contends that Mediatek FOMIP does not disclose one or more limitations of the asserted claims, it would have been obvious to combine the teachings of Mediatek FOMIP with (1) the knowledge of one of ordinary skill in the art, and/or (2) the teachings of any of the prior art set forth in NVIDIA’s other invalidity charts and cover pleading, to show all the limitations of the asserted claims, including for the reasons set forth in NVIDIA’s Invalidity Contention Cover Pleading, which are incorporated herein.

U.S. Patent No. 11,841,803	Fan-out Mediatek Innovation Package (“Mediatek FOMIP”)
1[Pre]: A system, comprising:	Including at least under Complainant’s apparent infringement theory, Mediatek FOMIP discloses and/or renders obvious this limitation. <i>See, e.g.,</i> <i>See evidence cited for ’803 patent claims 1-3, 7-10.</i> This limitation also would have been obvious to a person of ordinary skill in the art based on (1) Mediatek FOMIP alone; (2) Mediatek FOMIP in combination with the general knowledge of a person of ordinary skill in the art; and/or (3) Mediatek FOMIP in combination with the teachings with respect to this claim element as detailed in NVIDIA’s other invalidity charts and cover pleading.
1[A]: a central processing unit (CPU) communicably coupled to a first graphics processing unit (GPU) chiplet of a GPU chiplet array,	Including at least under Complainant’s apparent infringement theory, Mediatek FOMIP discloses and/or renders obvious this limitation. <i>See, e.g.,</i> <u>Lau</u> Lau, at Fig. 1.28

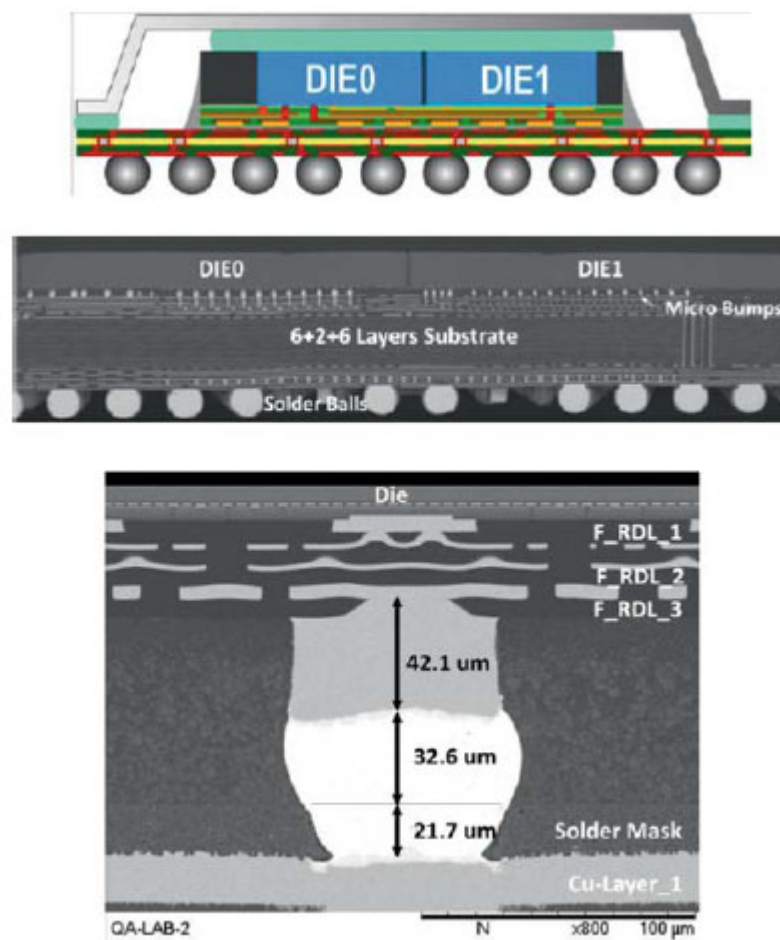
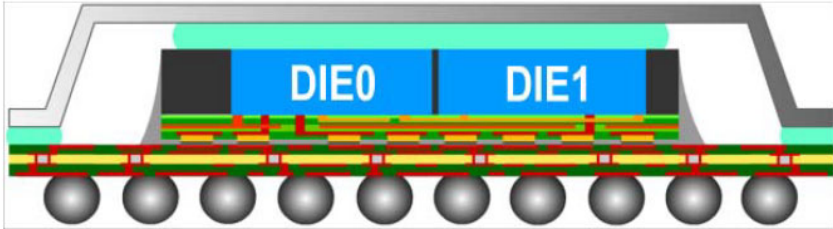
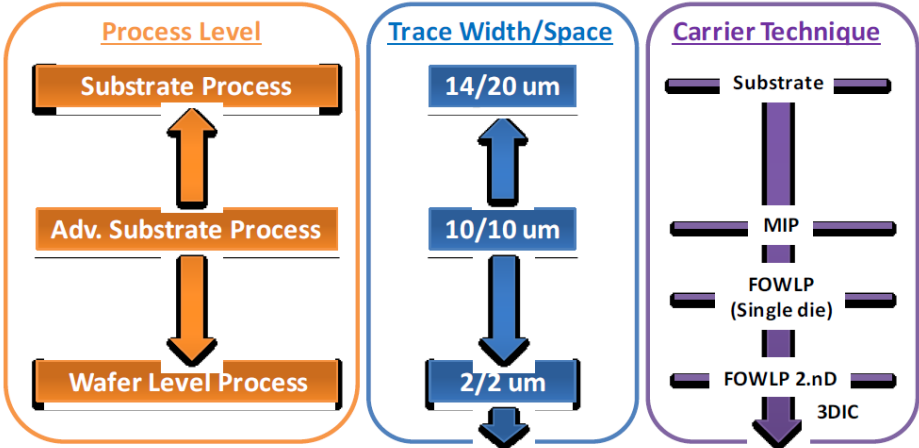


Fig. 1.28 Mediatek's fan-out RDL substrate

Lau, at 27 (“1.10.3 Mediatek’s RDLs by FOWLP

In 2016, Mediatek [133] proposed similar TSV-less interposer RDLs fabricated

U.S. Patent No. 11,841,803	Fan-out Mediatek Innovation Package (“Mediatek FOMIP”)
	with FOWLP technology as shown in Fig. 1.28. Instead of the C4 bump, they used a microbump (Cu pillar + solder cap) to connect the bottom RDL to the 6-2-6 package substrate.”) Chen-2016 Chen-2016, at Fig.1.  Fig.1 FOWLP process for multi-chips integration Chen-2016, at Fig.2.  Fig. 2 Capabilities of different PKG carriers Chen-2016, at Fig. 3.

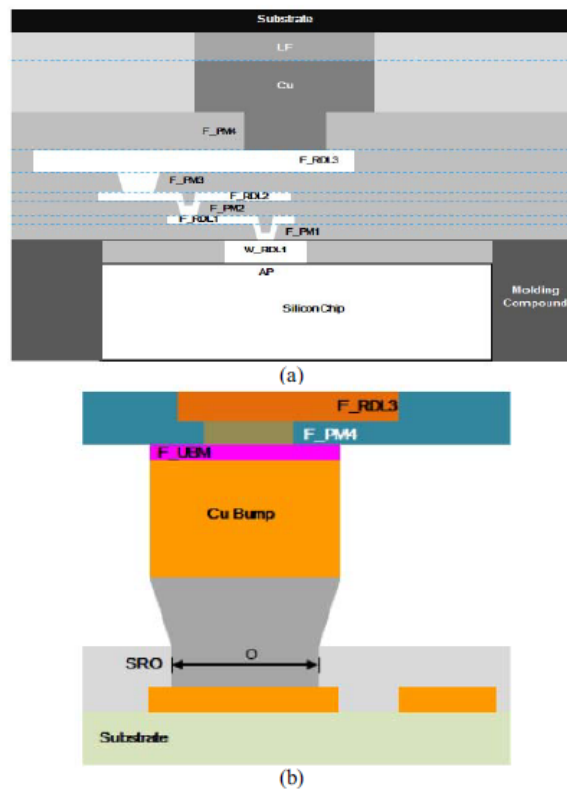


Fig.3 (a) FOWLP carrier structure, and (b) cu-pillar bump jointed to BGA substrate

Chen-2016, at Fig.4.

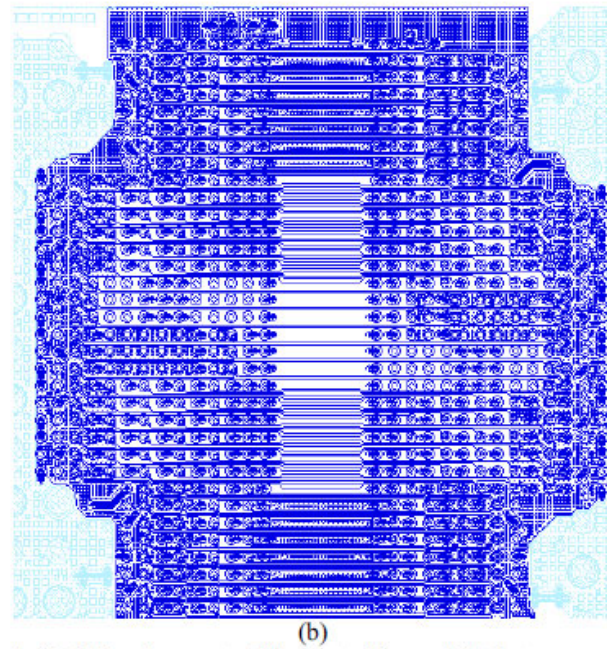
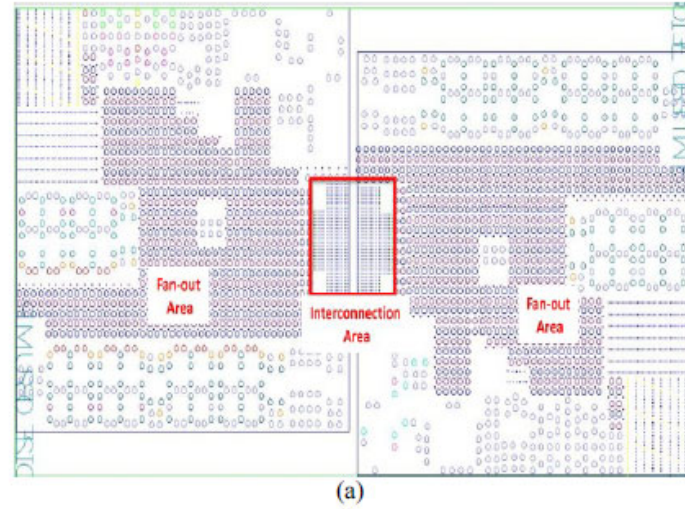


Fig.4 (a) Chip placement of the test chip, and (b) interconnection area on the FOWLP carrier.

Chen-2016, at Fig.5.

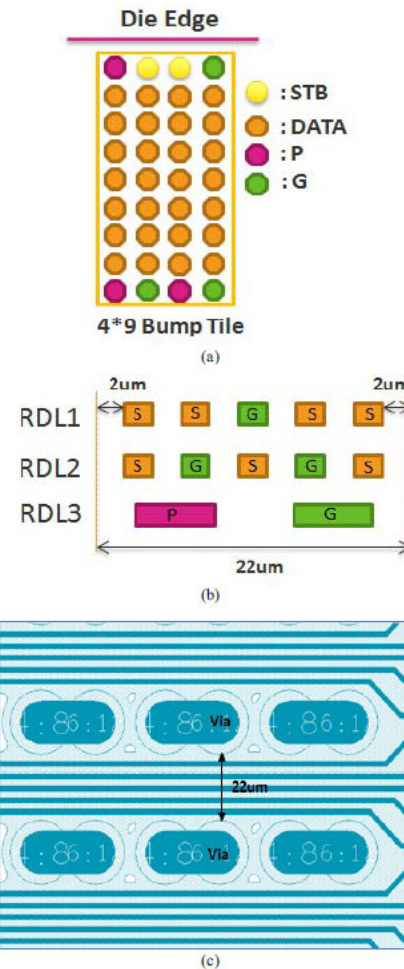


Fig.5 (a) Bump tile of interconnection bumps, (b) layout strategy of 4x9 bump matrix, and (c) an example of interconnection traces routing

Chen-2016, at Fig. 6.

U.S. Patent No. 11,841,803

Fan-out Mediatek Innovation Package (“Mediatek FOMIP”)

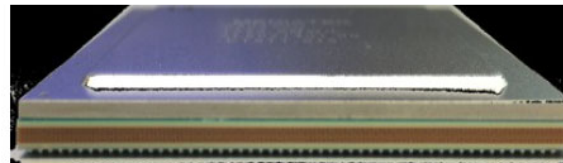
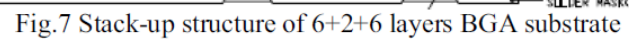


Fig.6 the package outline of the test chip

Chen-2016, at Fig. 7.



Chen-2016, at Fig. 8.

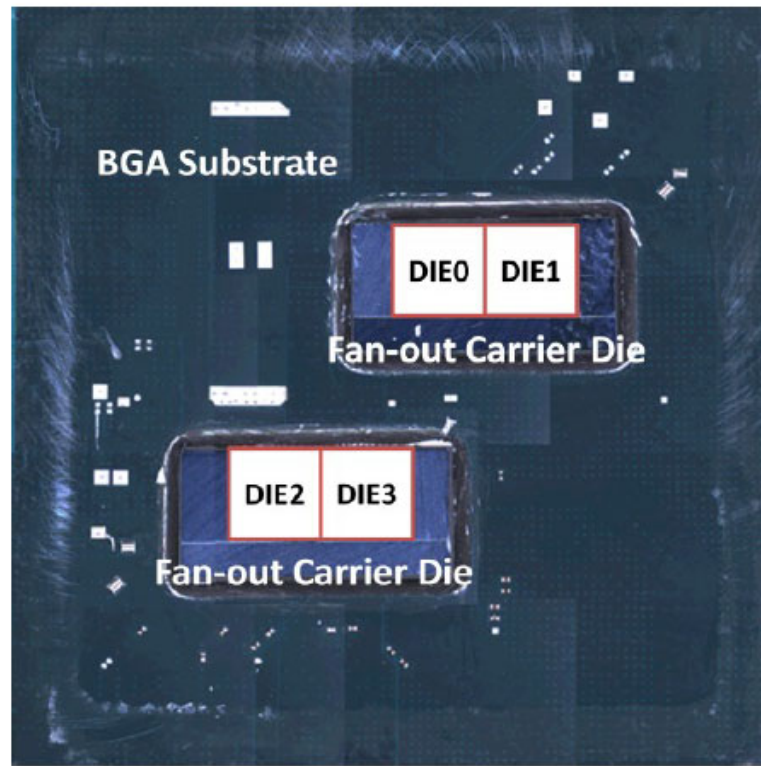
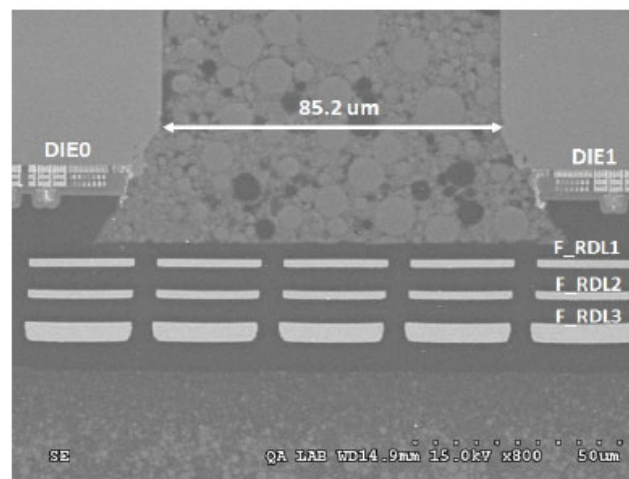
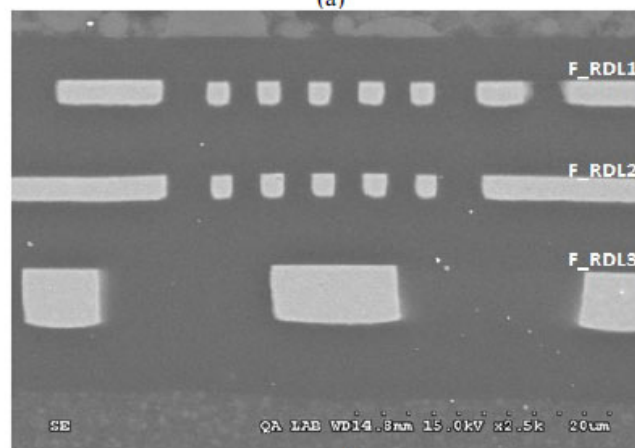


Fig.8The top view of the test chip

Chen-2016, at Fig. 9.



(a)



(b)

Fig.9 (a) Horizontal cut cross-section of a fan-out carrier die, and
(b) the vertical cut cross-section of a fan-out carrier die.

Chen-2016, at Fig. 10.

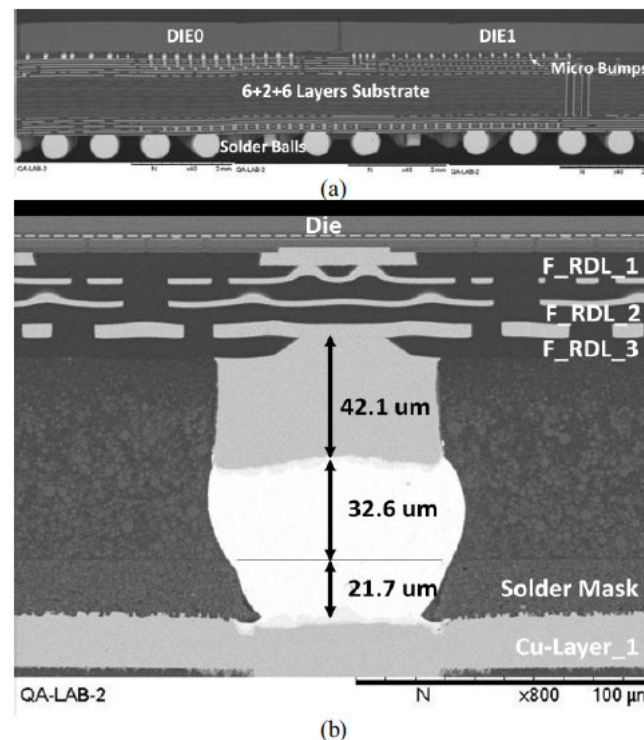


Fig.10 (a) Cross-section view of FOWLP, and (b) structure of micro bump connecting fan-out carrier and the BGA substrate.

Chen-2016, at Fig. 11.

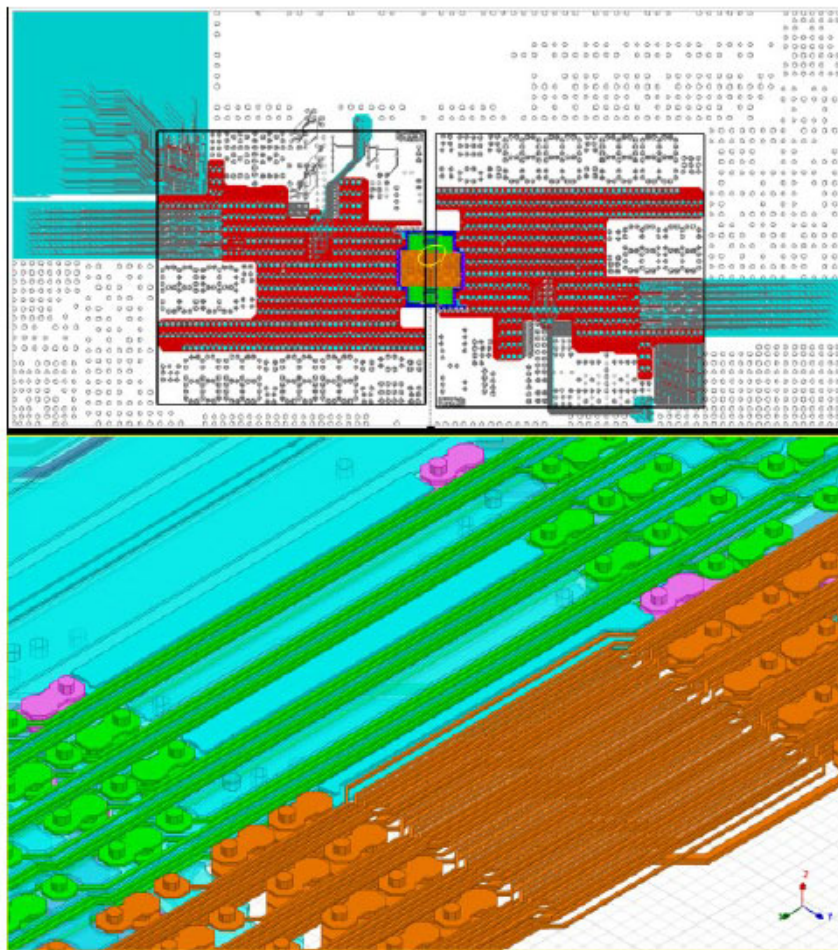


Fig.11 3D model of DIE to DIE interconnection

Chen-2016, at Abstract (“A novel multi-chips Fan-out Wafer Level Package (FOWLP) with fine pitch Cu pillar bumps was presented to accommodate volumes of I/O requirements on a BGA organic substrate. Due to the manufacturing limitation, cost and reliability consideration of fine-pitch BGA substrate, chips with extremely fine pitch RDL routing

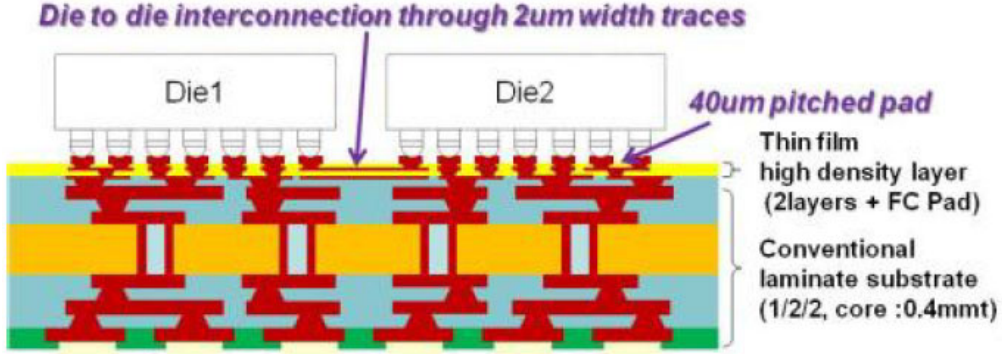
U.S. Patent No. 11,841,803	Fan-out Mediatek Innovation Package (“Mediatek FOMIP”)
	(width/space=2/2um) and a 2.nD wafer level carrier was integrated as system-in-package (SiP) to resolve I/O capacity issue. Besides, multi-chips with small die size connected through wafer-level RDL was also developed to resolve low yield issue of large chip in advanced wafer process. In this paper, a comprehensive design methodology was introduced to cover fan-out WLP interconnection and BGA substrate design. A test vehicle built in TSMC N16 process was presented to verify the design methodology. With fully consideration on signal integrity (SI) and power integrity (PI), hundreds of interconnection I/Os with 4Gbps speed were achieved. Besides, high speed SERDES with 25Gbit/sec data rate was also verified to enable high speed and high bandwidth data center application.”) Chen-2016, at 1495 (“I. INTRODUCTION High data rate and high bandwidth become more and more important in emerging data center market. In order to fulfill these requirements, silicon die size for data center application is usually enormous and suffers from high power consumption and low yield issues, especially in advanced wafer process. Thus, reducing single unit die size helps to increase gross die and improve the yield. Hence, a novel method to split a large die into two homogeneous smaller dice was introduced to resolve the yield loss and maintain the same bandwidth requirement. While the numbers of interconnection between multi-chips are more than 10 thousands, extremely fine pitch BGA substrate technology is necessary. However, it becomes a great challenge on BGA substrate manufacturing and causes loss on yield and cost. Moreover, the resulting reliability concerns increases during the extreme environment test under high/low temperature and humidity conditions. Therefore, a novel multi-chips Fan-out Wafer Level Package (FOWLP) was presented to accommodate volumes of I/O requirements.”) Chen-2016, at 1495 (“Fig. 1 demonstrates the concept of FOWLP by using fan-out technology with ultra fine pitch interconnections for multi-chips integration on a 2.nD wafer level carrier. The wafer level carrier is attached on an organic BGA substrate by fine pitch Cu pillar bumping process. Passive component, such as de-coupling capacitor, and heat spreader are also integrated on the BGA substrate as system in package (SiP). In this paper, a test vehicle built in TSMC N16 process was developed using FOWLP. Two homogeneous chips were bonded on a 2.nD wafer level carrier and connected with extremely fine pitch routings. The package design methodology is discussed in the Section II and III.

U.S. Patent No. 11,841,803	Fan-out Mediatek Innovation Package (“Mediatek FOMIP”)
	Chip-package co-design methodology and its correlation results including signal integrity (SI) and power integrity (PI) considerations are discussed in the Section IV. Finally, the conclusion is shown in Section V.”) Chen-2016, at 1495 (“II. FAN-OUT WLP INTERCONNECT The tremendous growth in the mobile handset, tablet and networking markets is fueled by consumer demand on increased mobility and functionality. This increased demand drives the development of 3D integration on IC devices, which requires more complex and sophisticated packaging techniques. A variety of advanced IC interconnection technologies, such as through-silicon via (TSV), chip-on-chip (CoC) and package-on-package (PoP), are proposed to fulfill the growing demands. In particular, emerging FOWLP technologies provide unique and innovative extensions into the 3D packaging arena. Compared with competing laminate-based technologies, FOWLP provides more aggressive design rules on advanced design [1].”) Chen-2016, at 1496 (“In addition, the unique characteristics of FOWLP enable the creation of innovative 3D structures that address the need for IC integration in emerging mobile and networking applications. Fig. 2 demonstrates the capabilities of different PKG carriers on advanced fan-out technologies so that we can know that only wafer level process can provide the ultra fine line width and space to meet the large bandwidth requirement. FOWLP is a 2D configuration by embedding die into a molded wafer as so-called ‘wafer reconstitution’. The molded wafer is processed through a standard WLP flow to create the final IC assembly structure. The active surface of the die is co-planar with the mold compound and allows fan-out of copper trace routing into the molded area using conventional redistribution layer (RDL) processing [2]. Therefore, FOWLP can benefit from higher I/O density, smaller form factor, good electrical and mechanical performance. It provides the scalability within a heterogeneous assembly platform and the opportunity for advanced 3D structures.”) Chen-2016, at 1496 (“Limited capability of line width and spacing for the RDL fan-out structure are the main restriction of traditional FOWLP. Typically, mask aligners image the photo resistance and dielectric materials for FOWLP packages. Because of the limited depth-of-field, this patterning technique is highly dependent on the surface topology to be developed. It is very difficult to control the planarization which spins on liquid dielectrics and photo-resists for a molded reconstituted wafer, especially when multiple layers of RDL are required.

U.S. Patent No. 11,841,803	Fan-out Mediatek Innovation Package (“Mediatek FOMIP”)
	The surface topology is also affected by the inherent warpage issue associated with molded wafers. Consequently, it is very difficult to achieve fine resolution features smaller than 10/10µm width/space on FOWLP [3]. An innovated 2.nD wafer level carrier is adopted to overcome those issues associated with conventional FOWLP technology. The added 2.nD wafer level carrier provides capabilities of extremely fine pitch RDL routing (width/space = 2/2 µm) and then bonded on a BGA organic substrate. Hence, it provides increased I/O and circuit density within a reduced footprint for single- or multi-chips applications.”) Chen-2016, at 1496 (“This innovative wafer-level packaging technique allows more aggressive design rules compared with competing FOWLP and laminate-based packaging techniques. In addition, the characteristics of the new process enable the creation of 2D and 3D structures that address the need for IC integration in emerging mobile and networking applications. Although the package appears to be a typical fine-pitch flip-chip construction, it incorporates some unique features against conventional IC packages, including: 1) polymer-based dielectrics; 2) multi-chips and large die size capability; 3) large package body size capability; 4) interconnect density which is critical for SoC partitioning applications [4].”) Chen-2016, at 1496 (“III. FAN-OUT WLP AND BGA SUBSTRATE DESIGN A. Fan-out WLP 2.nD design A test chip was developed to verify FOWLP and its electrical performance. Three layers RDL are used for high speed signal fan-out routings and interconnections. Fig. 3(a) shows the structure of FOWLP carrier and Fig. 3(b) shows how the Cu-pillar bump joint to BGA substrate. Two silicon chips covered by molding compound through re-construction process and connected to a FOWLP carrier with micro bumps W_RDL1. F_RDL1, 2, 3 are three layers RDL for interconnection and fan-out between micro bumps and Cu-pillar bumps. F_PM1, 2, 3, 4 are the polyamide based dielectric layers and are placed between each RDL layer. A UBM layer is placed above the 4th dielectric layer, F_PM4, for connection to BGA substrate by Cu-pillar (CuP) bumps. The CuP bumps are constructed by copper with lead-free (Sn-Ag) cap and the bump height is 70 nm with 130nm bump pitch. The floor plan of test chip

U.S. Patent No. 11,841,803	Fan-out Mediatek Innovation Package (“Mediatek FOMIP”)
	is shown in Fig. 4. There are 4 identical dice integrated in the test chip and every 2 dice are connected by FOWLP. Fig. 4(a) shows the placement and connection of 2 homogeneous dice on a FOWLP carrier. The fan-out area is mainly for high speed SERDES signals which are directly connected from micro bumps to CuP bumps with very short routing traces to avoid signal quality loss. The interconnection area on FOWLP carrier is shown in Fig. 4(b). The distance between 2 dice is 100um and 414 small bumps are designed on each die for signal interconnections. Fine-pitch trace with minimum trace width/space=2/2um is used on the 1st and 2nd RDL layer while 5/5um is used on 3rd RDL layer.”) Chen-2016, at 1497 (“Total 2 carrier dice (DIE0/1 and DIE2/3) are mounted on BGA substrate and connected to each other by substrate routings. The interconnection area between 2 dice consists of several slices of bump tile and each slice consists of 4x9 bump tile including 28 data bits, 2 strobes and 6 power/ground. Fig. 5(a) shows the 1 slice of bump tile for interconnection. Each column of bump tile has 7 bumps and is routed on F_RDL1 and F_RDL2 layers for signal routing. The bump pitch is 44 um and small size bumps with diameter equals 30um is adopted. Since the space between 2 vias is 22um, it allows maximum 5 traces routing between 2 vias based on minimum trace design rules (width/space=2/2um). Therefore, it allows total 10 traces routing on F_RDL1 & 2. 7 traces are used for signal and the rest 3 traces are used for GND shielding. The power and ground traces are routed on F_RDL3 for wider trace width and spacing requirement. Fig. 5(b) shows the layout topology of 4x9 bump tile. The ground trace is arranged between signals for signal integrity consideration. Fig. 5(c) demonstrates the real routing of interconnection traces on F_RDL1 & 2.”) Chen-2016, at 1498-99 (“C. Structures of fan-out carrier and BGA substrate A physical examination was performed to verify the test chip structure of the fan-out carrier and BGA substrate. Fig. 8 shows the top view of the test chip. There are 2 fan-out carrier dice on the substrate and each carrier die contains 2 identical dice (DIE0/1 and DIE2/3). The detail horizontal cut cross-section diagram of a fan-out carrier die (DIE0/1) is shown in Fig. 9(a). The average metal thickness of the interconnection traces is 2.04um and 4.23um for F_RDL1/2 and F_RDL3 respectively. Fig. 9(b) shows the vertical cut cross-section diagram between DIE0 and DIE1. The average measurement results of trace width and space are 2.09um/2.00um which is within the design tolerance. Fig. 10(a) shows the overall cross-

U.S. Patent No. 11,841,803	Fan-out Mediatek Innovation Package (“Mediatek FOMIP”)
	section structure of FOWLP SiP. DIE0 and DIE1 are connected by a 3-RDL carrier and then mount to a 6+2+6 layers BGA substrate by micro bumps. Fig. 10(b) shows the structure of micro bump for between connecting fan-out carrier and the BGA substrate. The bump height after reflow process is about 96.4 um and matches to the design as shown on Fig. 3(b). Based on the experiment verification on structure cross-section, the process of FOWLP are stable and assured for mass production.”) Chen-2016, at 1499 (“IV. CHIP-PKG CO-DESIGN AND CORRELATION In this test chip, bandwidth with 10Tbps throughput is necessary and thus 10,000 parallel buses with data transceiver working at 1Gbps are required. One way to reduce the interconnection between two homogenous dice is to raise the data rate of data transceiver. However, it may suffer severe signal and power integrity issue without appropriate design. Thus, a timing budget analysis of the complete data path between transceiver to receiver is shown in Table I. Here, we named the data path of proposed fan-out chip as MediaTek I/F Link (MIL). The time skew of the MIL must be smaller than 20.5ps and 83ps for 4Gbps and 3.2Gbps operation, respectively. Through systematic and comprehensive co-simulation methodology, a 4x9 bump tile with 28 data bits, 2 strobes and 6 power/ground was proposed to optimize the SI/PI performance. The shielding guard traces could be appropriately implemented with the proposed bump tile to mitigate the crosstalk noise induced from adjacent aggressors. Fig. 11 shows the 3D simulation platform of DIE-to-DIE interconnection. The signal integrity as well as the power integrity can be ensured through co-simulation for high data rate requirement. In order to maintain the target bandwidth of 10Tbps, total 90 bump tiles were duplicated.”) Chen-2016, at 1500 (“The 2.nD FOWLP successfully connected two homogeneous dice using ultra fine pitch process. This study discussed the design experience and concepts of wafer level carrier. The 2/2 um trace width and space realized how to routing 10 traces on 2 RDL layers with very fine pitch (44 um) bump placement. In additions, the 6+2+6 layers BGA substrate design is also important to ensure the signal integrity meets its requirement on a 55x55mm PKG. About the experiment on real samples, we know the structure of FOWLP and the cross-section diagrams help verify the process capability of the wafer level carrier can meet our

U.S. Patent No. 11,841,803	Fan-out Mediatek Innovation Package (“Mediatek FOMIP”)
	design spec and requirements. Based on the Chip-PKG co-design and correlation results, the 2.nD FOWLP can handle bandwidth with 10Tbps throughput and 25GHz high speed SERDES signals. Therefore, the FOWLP interconnection methodology was approved to connect two homogeneous dice which were separated from a huge die. It really helps us to increase the gross die rate and improve the wafer yield when the silicon die size is usually enormous and suffers low yield issues, especially in advanced wafer process.”) <u>Oi</u> Oi, at Fig.1.  Figure 1. Schematic cross sectional image of i-THOP Oi, at Fig.2.

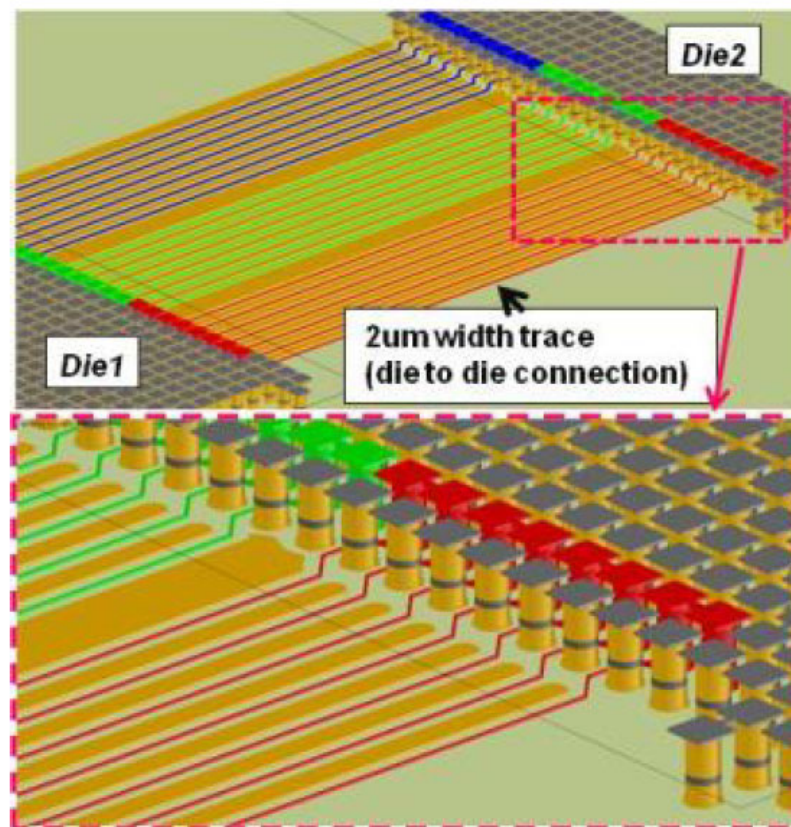


Figure 2. Die to die connection through 2 μ m width copper trace.

Oi, at Fig. 3.

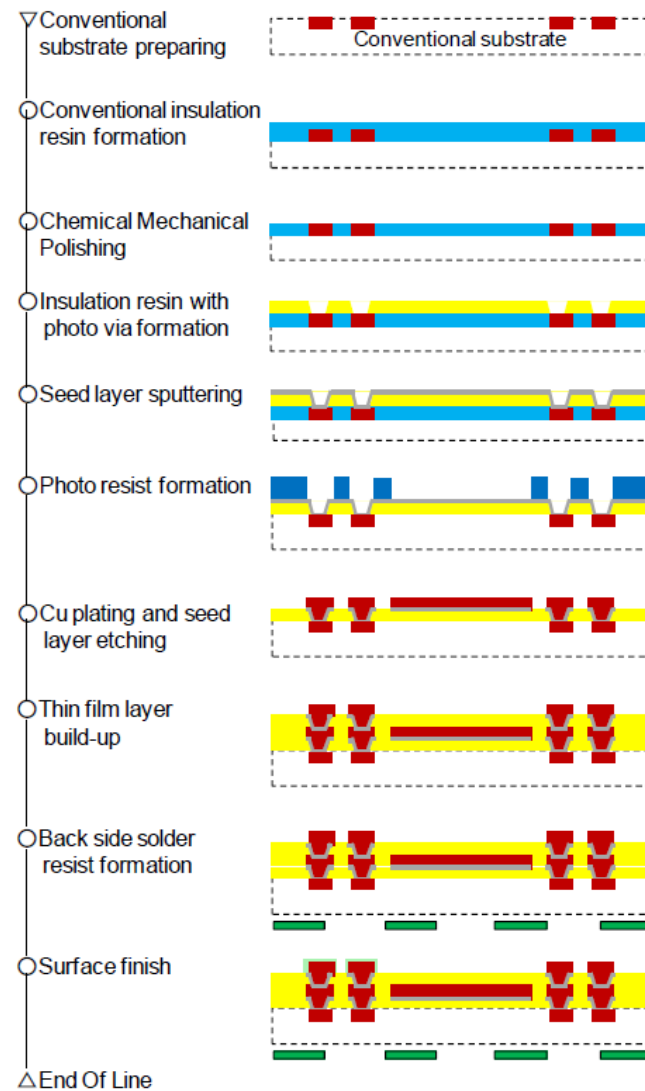


Figure 3. Schematic images of i-THOP process.

Oi, at Fig. 4.

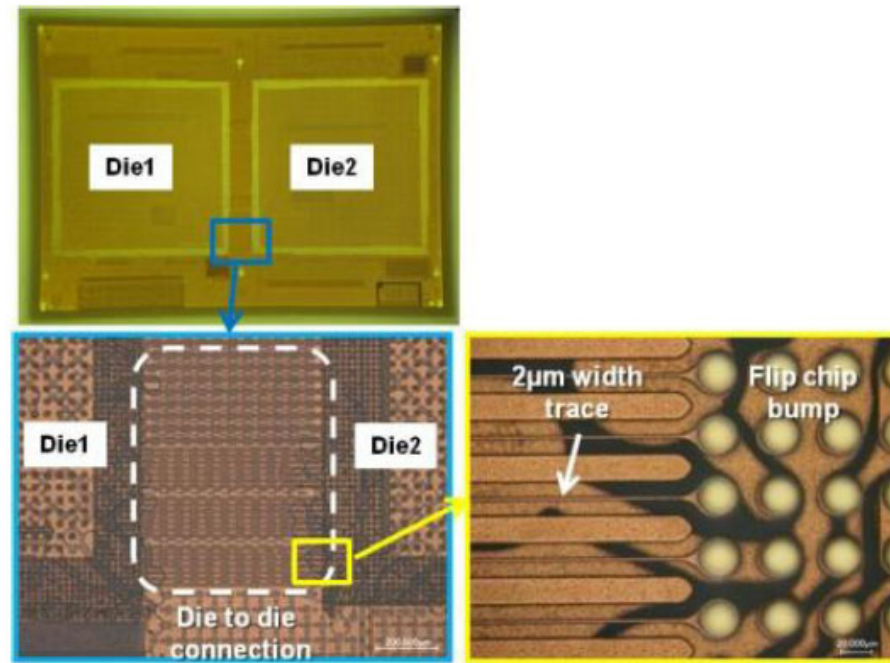


Figure 4. Die patterns on i-THOP

Oi, at Fig. 5.

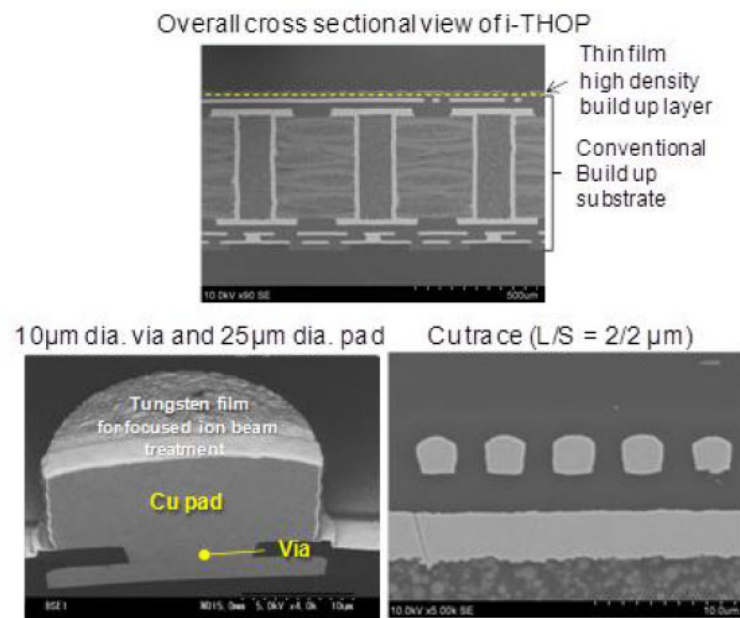
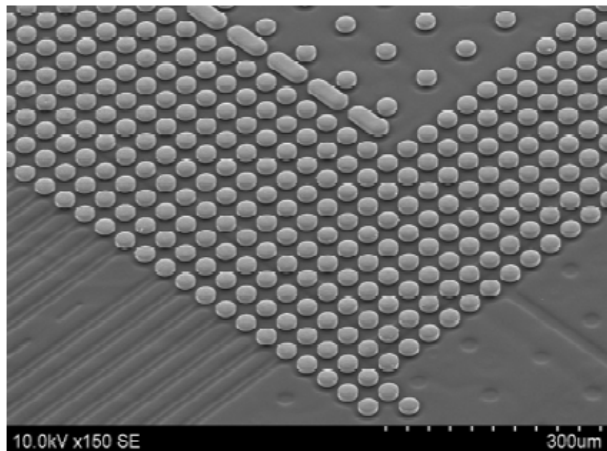


Figure 5. SEM image of i-THOP cross section

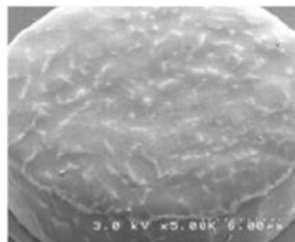
Oi, at Fig. 6.

A) 40µm pitch flip chip bonding pads



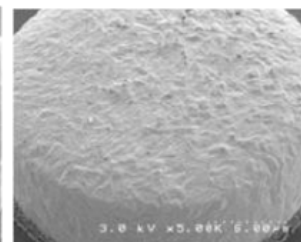
B) Enlarged image

Cu-OSP



Ra:0.30µm

Thin ENEPIG



Ra:0.31µm

Figure 6. SEM images of flip chip bonding pad.

Oi, at Fig. 8.

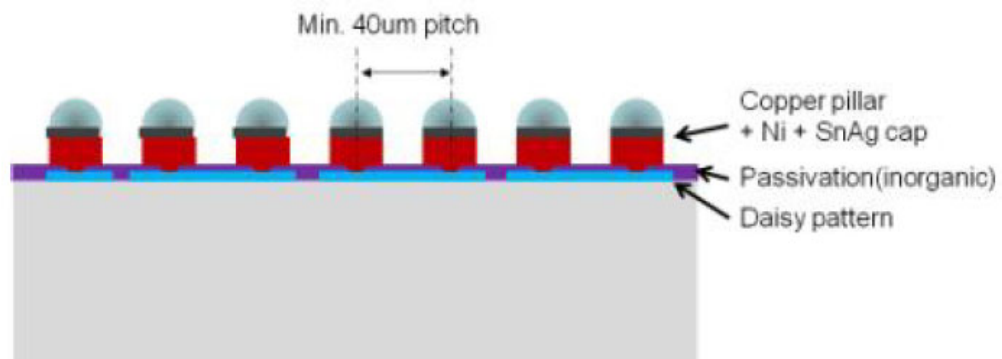


Figure 8. Schematic image of the test die.

Oi, at Fig.9.

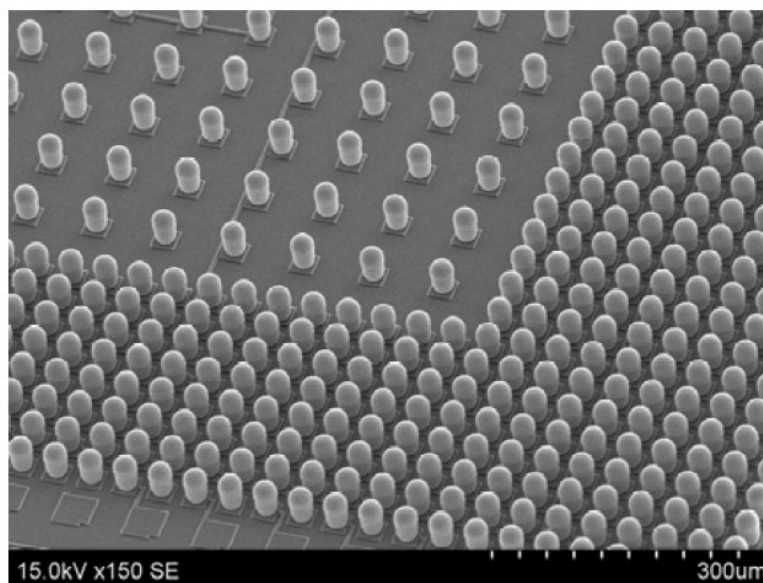
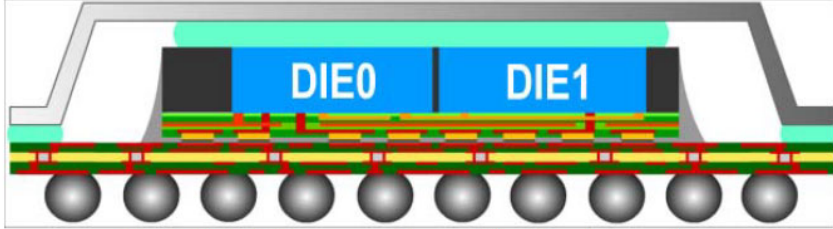
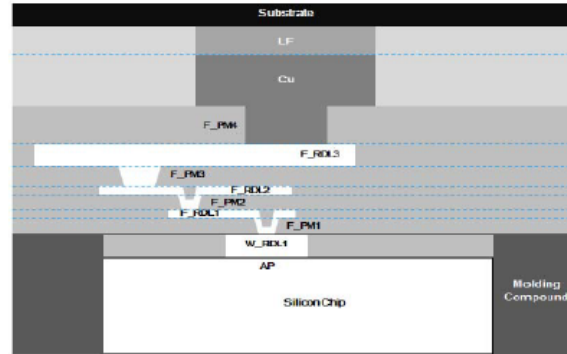


Figure 9. SEM image of test die

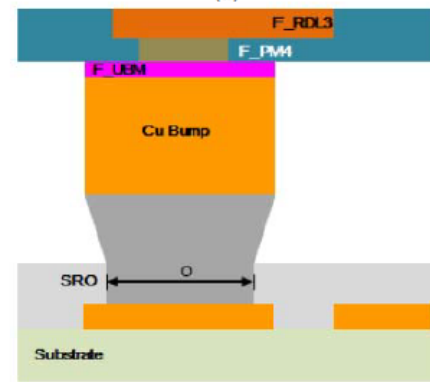
U.S. Patent No. 11,841,803	Fan-out Mediatek Innovation Package (“Mediatek FOMIP”)
	Oi, at Abstract (“2.5D packaging technology utilizing silicon interposers is being developed and used for high-performance applications as the demand for miniaturization and higher density continues to increase. Silicon interposers enable very high density interconnects using standard semiconductor fabrication process technology, but are challenged as size increases. An alternative solution, high density laminate interposer, can offer advantages in cost, form factor, and infrastructure over other interposer options available today. However, it is difficult to manufacture high density wiring and flip chip bump pads on laminate interposer because conventional build-up technology limits fine line and via diameter. Our solution is the combination of an integrated organic interposer substrate with high density interconnects and thermo-compression flip chip bonding. Our solution eliminates the backside integration process for silicon interposer and assembly of interposers onto substrates. Conventional assembly processes can be utilized for assembling dies onto the integrated organic interposer substrate. Feasibility of this 2.5D package has been demonstrated by assembling dual-die with 40um pitch copper pillar bumps onto this novel integrated organic interposer substrate with 2μm line and space.”) Oi, at 348 (“Introduction Recent advancement in mobile applications such as smart phones and tablets are progressing rapidly. The demand on miniaturization of form factor and high performance are becoming much more important factors. Die partitioning enables miniaturization of small form factor and high performance without the yield issues of other methods. The focus of this paper is die partitioning technology conducted by fine pitch flip chip interconnect and die to die connection through fine lines. The materials under consideration for such an interposer are silicon, glass, and organic dielectrics. Silicon is a leading candidate for 2.5D packaging [1]-[3], but is challenged as size increases. Glass offers the promise of performance at lower cost than silicon, but there is no manufacturing infrastructure. Organic materials used in substrates have a prevalent supply chain, but cannot meet 2.5D line and space requirements that are finer than conventional laminate.

U.S. Patent No. 11,841,803	Fan-out Mediatek Innovation Package (“Mediatek FOMIP”)
	By processing a fine-metal layer interposer directly on top of a conventional organic substrate, we have designed an integrated organic substrate and organic interposer. Our development progress has confirmed some basic properties [4], though we have not yet fully evaluated the die assembly. Key Points: • High accuracy alignment of die to interposer. • CTE mismatch and warpage control during soldering. • Elimination of underfill void. We have also measured thermal warpage of the organic interposer with die. We will report on the feasibility of our design for 2.5D packaging focusing on the assembly process.”) Oi, at 348 (“The Concept of Organic Interposer “ i-THOP” (Integrated Thin film High density Organic Package) Figure 1 shows a schematic cross-sectional image of i-THOP with die. Dual-die were assembled side by side and fine copper traces were used for die to die connection. A schematic image of die to die connections through fine line and fine pitched bump is shown in Figure 2. The pattern consisted of a thin film layer which is built on a conventional laminate substrate.”) This limitation also would have been obvious to a person of ordinary skill in the art based on (1) Mediatek FOMIP alone; (2) Mediatek FOMIP in combination with the general knowledge of a person of ordinary skill in the art; and/or (3) Mediatek FOMIP in combination with the teachings with respect to this claim element as detailed in NVIDIA’s other invalidity charts and cover pleading.
1[B]: wherein the GPU chiplet array includes: the first GPU chiplet communicably coupled to the CPU via a bus; and	Including at least under Complainant’s apparent infringement theory, Mediatek FOMIP discloses and/or renders obvious this limitation. <i>See, e.g.,</i> <i>See evidence cited for ’803 patent limitation 1[A].</i>

	This limitation also would have been obvious to a person of ordinary skill in the art based on (1) Mediatek FOMIP alone; (2) Mediatek FOMIP in combination with the general knowledge of a person of ordinary skill in the art; and/or (3) Mediatek FOMIP in combination with the teachings with respect to this claim element as detailed in NVIDIA's other invalidity charts and cover pleading.
1[C]: a second GPU chiplet communicably coupled to the first GPU chiplet via a passive crosslink,	Including at least under Complainant's apparent infringement theory, Mediatek FOMIP discloses and/or renders obvious this limitation. <i>See, e.g.,</i> <u>Chen-2016</u> Chen-2016, at Fig.1.  Fig.1 FOWLP process for multi-chips integration Chen-2016, at Fig. 3.



(a)



(b)

Fig.3 (a) FOWLP carrier structure, and (b) cu-pillar bump jointed to BGA substrate

Chen-2016, at Fig.4.

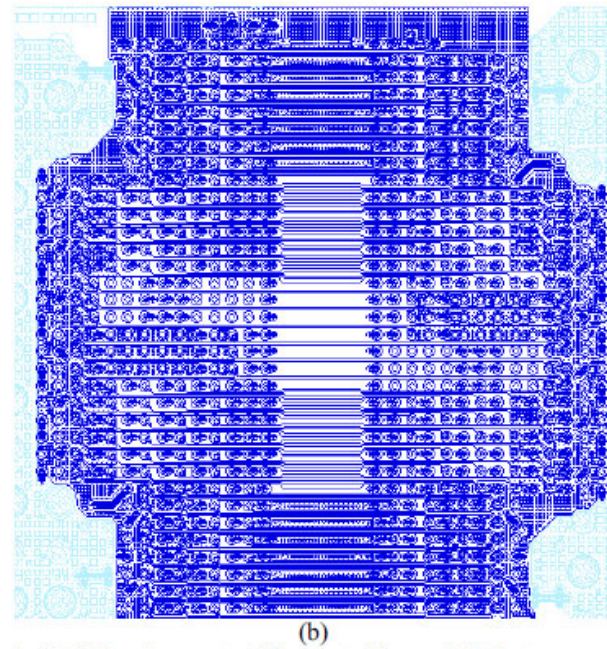
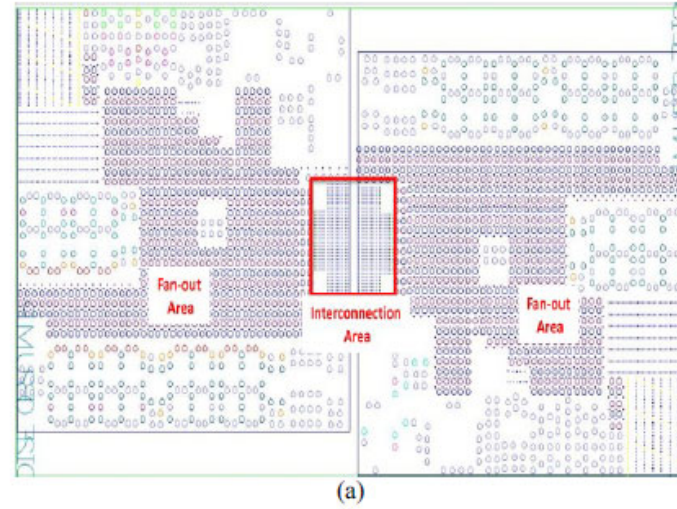


Fig.4 (a) Chip placement of the test chip, and (b) interconnection area on the FOWLP carrier.

Chen-2016, at Fig.5.

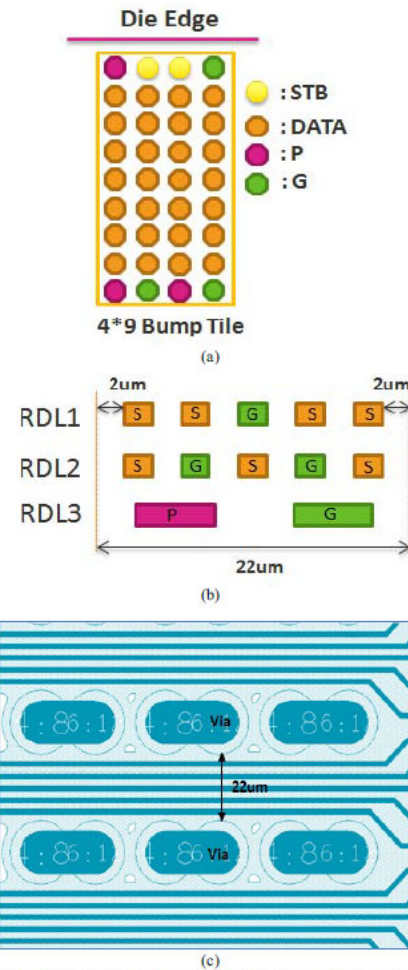


Fig.5 (a) Bump tile of interconnection bumps, (b) layout strategy of 4x9 bump matrix, and (c) an example of interconnection traces routing

Chen-2016, at Fig. 7.

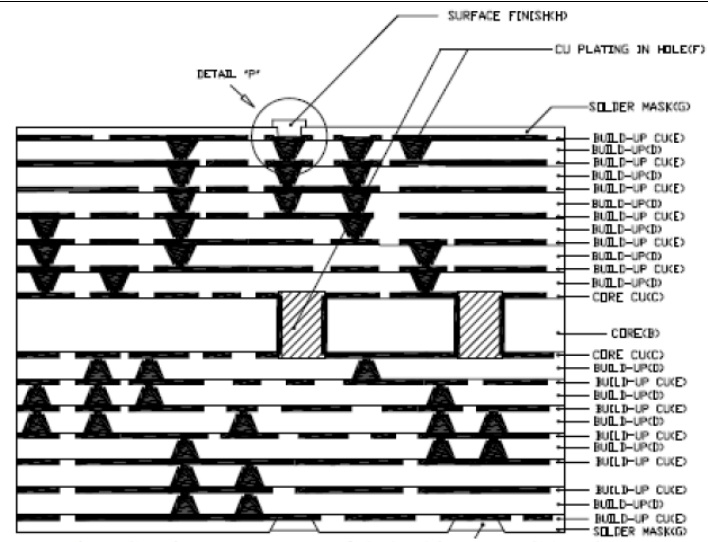
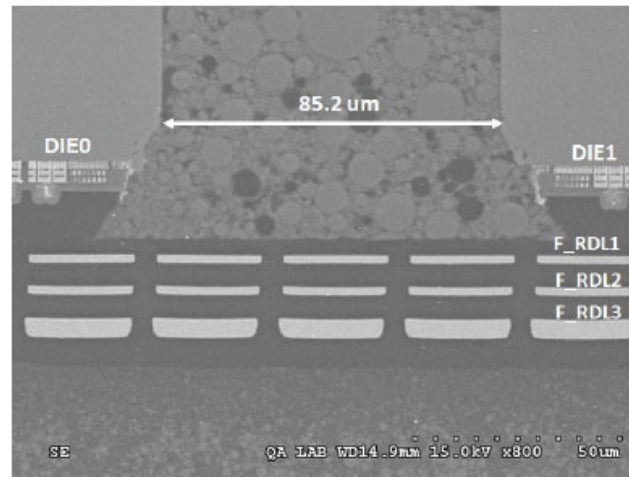
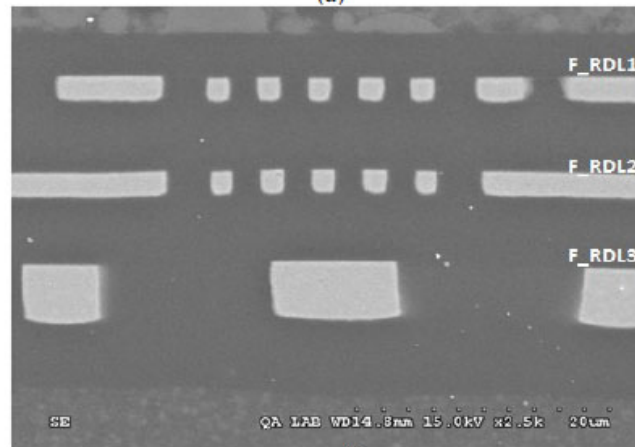


Fig.7 Stack-up structure of 6+2+6 layers BGA substrate

Chen-2016, at Fig. 9.



(a)



(b)

Fig.9 (a) Horizontal cut cross-section of a fan-out carrier die, and
(b) the vertical cut cross-section of a fan-out carrier die.

Chen-2016, at Fig. 10.

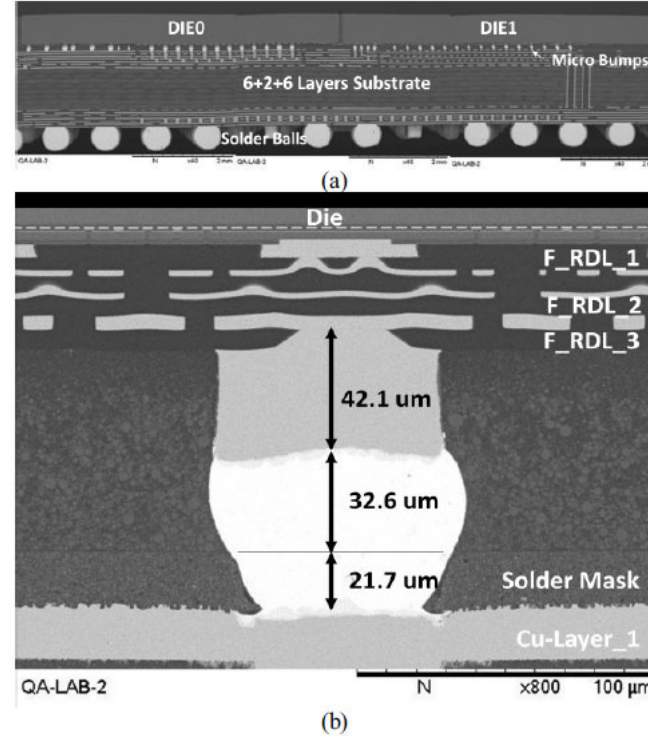


Fig.10 (a) Cross-section view of FOWLP, and (b) structure of micro bump connecting fan-out carrier and the BGA substrate.

Chen-2016, at Fig. 11.

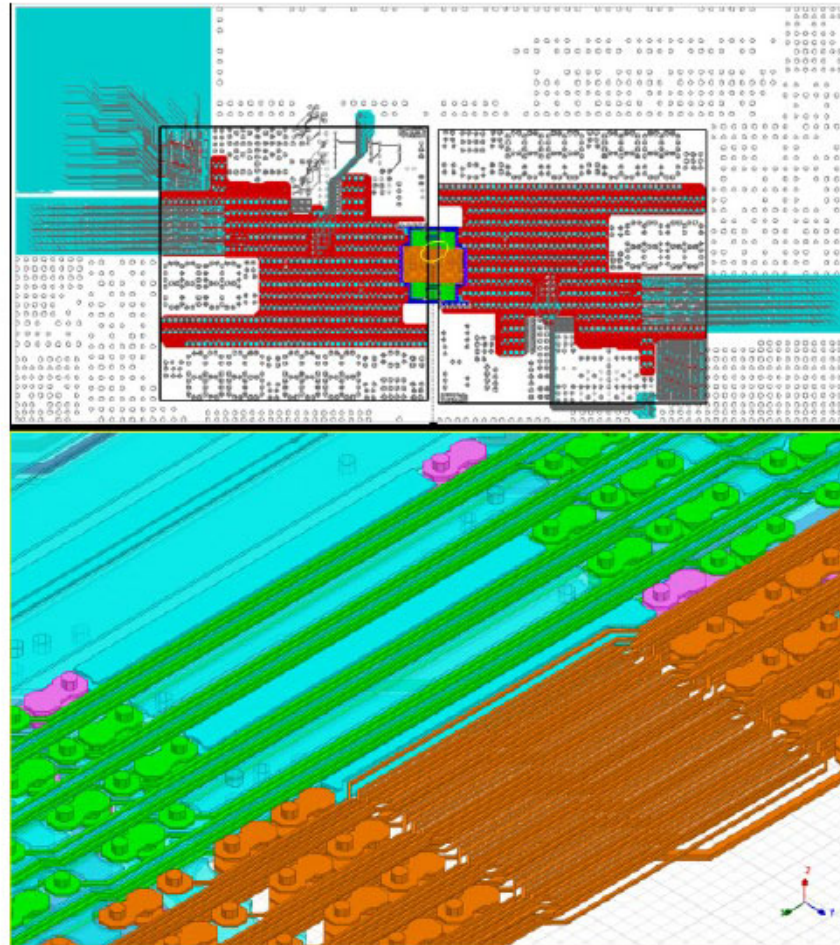


Fig.11 3D model of DIE to DIE interconnection

Chen-2016, at Abstract (“A novel multi-chips Fan-out Wafer Level Package (FOWLP) with fine pitch Cu pillar bumps was presented to accommodate volumes of I/O requirements on a BGA organic substrate. Due to the manufacturing limitation, cost and reliability consideration of fine-pitch BGA substrate, chips with extremely fine pitch RDL routing (width/space=2/2um) and a 2.nD wafer level carrier was integrated as system-in-package (SiP)

to resolve I/O capacity issue. Besides, multi-chips with small die size connected through wafer-level RDL was also developed to resolve low yield issue of large chip in advanced wafer process. In this paper, a comprehensive design methodology was introduced to cover fan-out WLP interconnection and BGA substrate design. A test vehicle built in TSMC N16 process was presented to verify the design methodology. With fully consideration on signal integrity (SI) and power integrity (PI), hundreds of interconnection I/Os with 4Gbps speed were achieved. Besides, high speed SERDES with 25Gbit/sec data rate was also verified to enable high speed and high bandwidth data center application.”)

Chen-2016, at 1495 (“I. INTRODUCTION

High data rate and high bandwidth become more and more important in emerging data center market. In order to fulfill these requirements, silicon die size for data center application is usually enormous and suffers from high power consumption and low yield issues, especially in advanced wafer process. Thus, reducing single unit die size helps to increase gross die and improve the yield. Hence, a novel method to split a large die into two homogeneous smaller dice was introduced to resolve the yield loss and maintain the same bandwidth requirement. While the numbers of interconnection between multi-chips are more than 10 thousands, extremely fine pitch BGA substrate technology is necessary. However, it becomes a great challenge on BGA substrate manufacturing and causes loss on yield and cost. Moreover, the resulting reliability concerns increases during the extreme environment test under high/low temperature and humidity conditions. Therefore, a novel multi-chips Fan-out Wafer Level Package (FOWLP) was presented to accommodate volumes of I/O requirements.”)

Chen-2016, at 1495 (“Fig. 1 demonstrates the concept of FOWLP by using fan-out technology with ultra fine pitch interconnections for multi-chips integration on a 2.nD wafer level carrier. The wafer level carrier is attached on an organic BGA substrate by fine pitch Cu pillar bumping process. Passive component, such as de-coupling capacitor, and heat spreader are also integrated on the BGA substrate as system in package (SiP).

In this paper, a test vehicle built in TSMC N16 process was developed using FOWLP. Two homogeneous chips were bonded on a 2.nD wafer level carrier and connected with extremely fine pitch routings. The package design methodology is discussed in the Section II and III. Chip-package co-design methodology and its correlation results including signal integrity (SI)

and power integrity (PI) considerations are discussed in the Section IV. Finally, the conclusion is shown in Section V.”)

Chen-2016, at 1495 (“II. FAN-OUT WLP INTERCONNECT

The tremendous growth in the mobile handset, tablet and networking markets is fueled by consumer demand on increased mobility and functionality. This increased demand drives the development of 3D integration on IC devices, which requires more complex and sophisticated packaging techniques. A variety of advanced IC interconnection technologies, such as through-silicon via (TSV), chip-on-chip (CoC) and package-on-package (PoP), are proposed to fulfill the growing demands. In particular, emerging FOWLP technologies provide unique and innovative extensions into the 3D packaging arena. Compared with competing laminate-based technologies, FOWLP provides more aggressive design rules on advanced design [1].”)

Chen-2016, at 1496 (“In addition, the unique characteristics of FOWLP enable the creation of innovative 3D structures that address the need for IC integration in emerging mobile and networking applications. Fig. 2 demonstrates the capabilities of different PKG carriers on advanced fan-out technologies so that we can know that only wafer level process can provide the ultra fine line width and space to meet the large bandwidth requirement. FOWLP is a 2D configuration by embedding die into a molded wafer as so-called ‘wafer reconstitution’. The molded wafer is processed through a standard WLP flow to create the final IC assembly structure. The active surface of the die is co-planar with the mold compound and allows fan-out of copper trace routing into the molded area using conventional redistribution layer (RDL) processing [2]. Therefore, FOWLP can benefit from higher I/O density, smaller form factor, good electrical and mechanical performance. It provides the scalability within a heterogeneous assembly platform and the opportunity for advanced 3D structures.”)

Chen-2016, at 1496 (“Limited capability of line width and spacing for the RDL fan-out structure are the main restriction of traditional FOWLP. Typically, mask aligners image the photo resistance and dielectric materials for FOWLP packages. Because of the limited depth-of-field, this patterning technique is highly dependent on the surface topology to be developed. It is very difficult to control the planarization which spins on liquid dielectrics and photo-resists for a molded reconstituted wafer, especially when multiple layers of RDL are required. The surface topology is also affected by the inherent warpage issue associated with molded wafers. Consequently, it is very difficult to achieve fine resolution features smaller than

10/10μm width/space on FOWLP [3]. An innovated 2.nD wafer level carrier is adopted to overcome those issues associated with conventional FOWLP technology. The added 2.nD wafer level carrier provides capabilities of extremely fine pitch RDL routing (width/space = 2/2 μm) and then bonded on a BGA organic substrate. Hence, it provides increased I/O and circuit density within a reduced footprint for single- or multi-chips applications.”)

Chen-2016, at 1496 (“This innovative wafer-level packaging technique allows more aggressive design rules compared with competing FOWLP and laminate-based packaging techniques. In addition, the characteristics of the new process enable the creation of 2D and 3D structures that address the need for IC integration in emerging mobile and networking applications. Although the package appears to be a typical fine-pitch flip-chip construction, it incorporates some unique features against conventional IC packages, including: 1) polymer-based dielectrics; 2) multi-chips and large die size capability; 3) large package body size capability; 4) interconnect density which is critical for SoC partitioning applications [4].”)

Chen-2016, at 1496 (“III. FAN-OUT WLP AND BGA SUBSTRATE DESIGN

A. Fan-out WLP 2.nD design

A test chip was developed to verify FOWLP and its electrical performance. Three layers RDL are used for high speed signal fan-out routings and interconnections. Fig. 3(a) shows the structure of FOWLP carrier and Fig. 3(b) shows how the Cu-pillar bump joint to BGA substrate. Two silicon chips covered by molding compound through re-construction process and connected to a FOWLP carrier with micro bumps W_RDL1. F_RDL1, 2, 3 are three layers RDL for interconnection and fan-out between micro bumps and Cu-pillar bumps. F_PM1, 2, 3, 4 are the polyamide based dielectric layers and are placed between each RDL layer.

A UBM layer is placed above the 4th dielectric layer, F_PM4, for connection to BGA substrate by Cu-pillar (CuP) bumps. The CuP bumps are constructed by copper with lead-free (Sn-Ag) cap and the bump height is 70 nm with 130nm bump pitch. The floor plan of test chip is shown in Fig. 4. There are 4 identical dice integrated in the test chip and every 2 dice are connected by FOWLP. Fig. 4(a) shows the placement and connection of 2 homogeneous dice on a FOWLP carrier. The fan-out area is mainly for high speed SERDES signals which are directly connected from micro bumps to CuP bumps with very short routing traces to avoid

signal quality loss. The interconnection area on FOWLP carrier is shown in Fig. 4(b). The distance between 2 dice is 100um and 414 small bumps are designed on each die for signal interconnections. Fine-pitch trace with minimum trace width/space=2/2um is used on the 1st and 2nd RDL layer while 5/5um is used on 3rd RDL layer.”)

Chen-2016, at 1497 (“Total 2 carrier dice (DIE0/1 and DIE2/3) are mounted on BGA substrate and connected to each other by substrate routings. The interconnection area between 2 dice consists of several slices of bump tile and each slice consists of 4x9 bump tile including 28 data bits, 2 strobes and 6 power/ground. Fig. 5(a) shows the 1 slice of bump tile for interconnection. Each column of bump tile has 7 bumps and is routed on F_RDL1 and F_RDL2 layers for signal routing. The bump pitch is 44 um and small size bumps with diameter equals 30um is adopted. Since the space between 2 vias is 22um, it allows maximum 5 traces routing between 2 vias based on minimum trace design rules (width/space=2/2um). Therefore, it allows total 10 traces routing on F_RDL1 & 2. 7 traces are used for signal and the rest 3 traces are used for GND shielding. The power and ground traces are routed on F_RDL3 for wider trace width and spacing requirement. Fig. 5(b) shows the layout topology of 4x9 bump tile. The ground trace is arranged between signals for signal integrity consideration. Fig. 5(c) demonstrates the real routing of interconnection traces on F_RDL1 & 2.”)

Chen-2016, at 1498-99 (“C. Structures of fan-out carrier and BGA substrate

A physical examination was performed to verify the test chip structure of the fan-out carrier and BGA substrate. Fig. 8 shows the top view of the test chip. There are 2 fan-out carrier dice on the substrate and each carrier die contains 2 identical dice (DIE0/1 and DIE2/3). The detail horizontal cut cross-section diagram of a fan-out carrier die (DIE0/1) is shown in Fig. 9(a). The average metal thickness of the interconnection traces is 2.04um and 4.23um for F_RDL1/2 and F_RDL3 respectively. Fig. 9(b) shows the vertical cut cross-section diagram between DIE0 and DIE1. The average measurement results of trace width and space are 2.09um/2.00um which is within the design tolerance. Fig. 10(a) shows the overall cross-section structure of FOWLP SiP. DIE0 and DIE1 are connected by a 3-RDL carrier and then mount to a 6+2+6 layers BGA substrate by micro bumps. Fig. 10(b) shows the structure of micro bump for between connecting fan-out carrier and the BGA substrate. The bump height after reflow process is about 96.4 um and matches to the design as shown on Fig. 3(b). Based

	on the experiment verification on structure cross-section, the process of FOWLP are stable and assured for mass production.”) Chen-2016, at 1499 (“IV. CHIP-PKG CO-DESIGN AND CORRELATION In this test chip, bandwidth with 10Tbps throughput is necessary and thus 10,000 parallel buses with data transceiver working at 1Gbps are required. One way to reduce the interconnection between two homogenous dice is to raise the data rate of data transceiver. However, it may suffer severe signal and power integrity issue without appropriate design. Thus, a timing budget analysis of the complete data path between transceiver to receiver is shown in Table I. Here, we named the data path of proposed fan-out chip as MediaTek I/F Link (MIL). The time skew of the MIL must be smaller than 20.5ps and 83ps for 4Gbps and 3.2Gbps operation, respectively. Through systematic and comprehensive co-simulation methodology, a 4x9 bump tile with 28 data bits, 2 strobes and 6 power/ground was proposed to optimize the SI/PI performance. The shielding guard traces could be appropriately implemented with the proposed bump tile to mitigate the crosstalk noise induced from adjacent aggressors. Fig. 11 shows the 3D simulation platform of DIE-to-DIE interconnection. The signal integrity as well as the power integrity can be ensured through co-simulation for high data rate requirement. In order to maintain the target bandwidth of 10Tbps, total 90 bump tiles were duplicated.”) Chen-2016, at 1500 (“The 2.nD FOWLP successfully connected two homogeneous dice using ultra fine pitch process. This study discussed the design experience and concepts of wafer level carrier. The 2/2 um trace width and space realized how to routing 10 traces on 2 RDL layers with very fine pitch (44 um) bump placement. In additions, the 6+2+6 layers BGA substrate design is also important to ensure the signal integrity meets its requirement on a 55x55mm PKG. About the experiment on real samples, we know the structure of FOWLP and the cross-section diagrams help verify the process capability of the wafer level carrier can meet our design spec and requirements. Based on the Chip-PKG co-design and correlation results, the 2.nD FOWLP can handle bandwidth with 10Tbps throughput and 25GHz high speed SERDES signals.
--	---

Therefore, the FOWLP interconnection methodology was approved to connect two homogeneous dice which were separated from a huge die. It really helps us to increase the gross die rate and improve the wafer yield when the silicon die size is usually enormous and suffers low yield issues, especially in advanced wafer process.”)

Oi

Oi, at Fig.1.

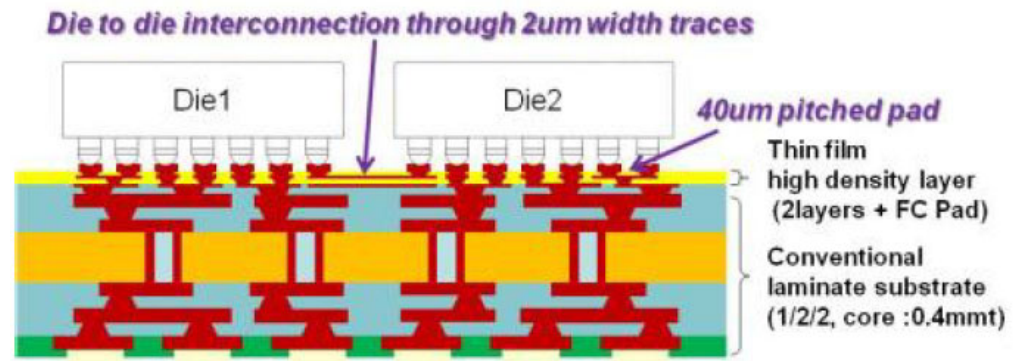


Figure 1. Schematic cross sectional image of i-THOP

Oi, at Fig.2.

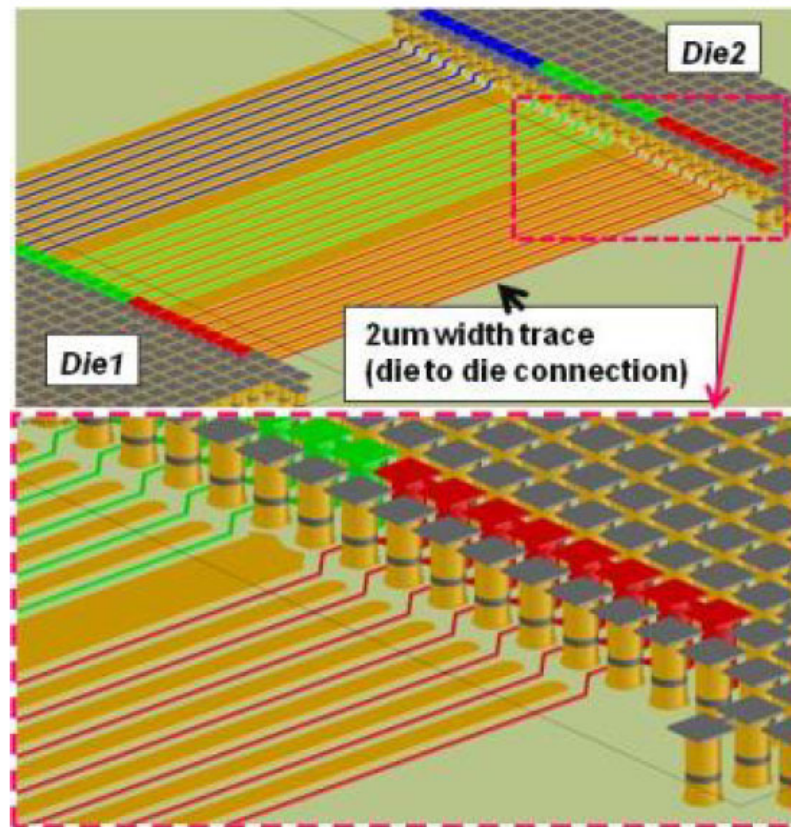


Figure 2. Die to die connection through 2 μ m width copper trace.

Oi, at Fig. 3.

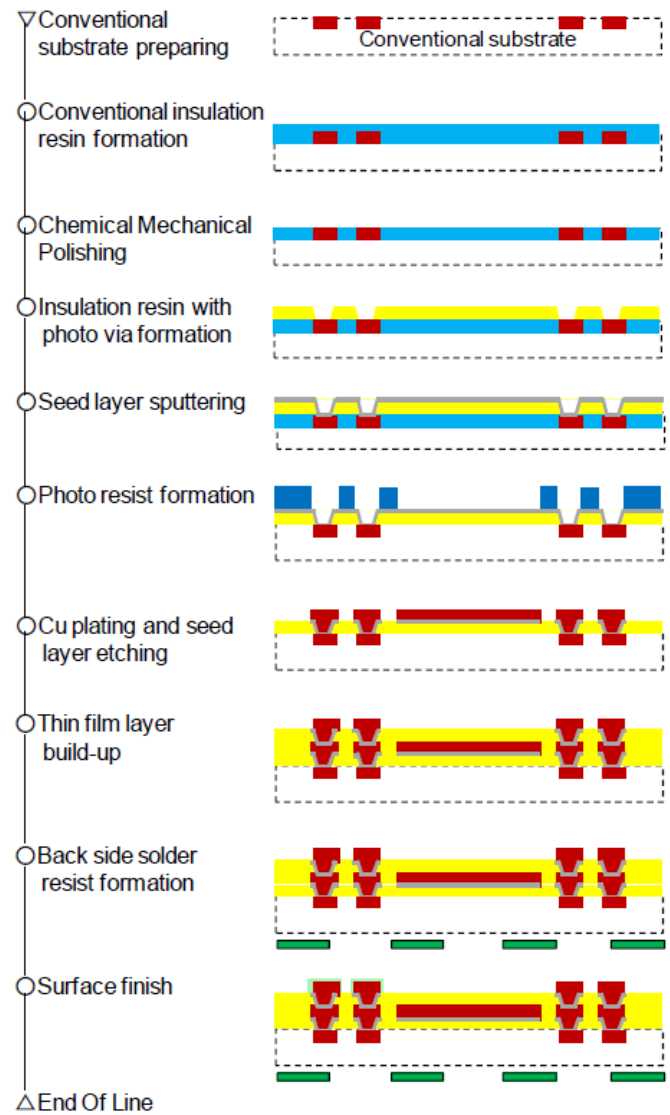


Figure 3. Schematic images of i-THOP process.

Oi, at Fig. 4.

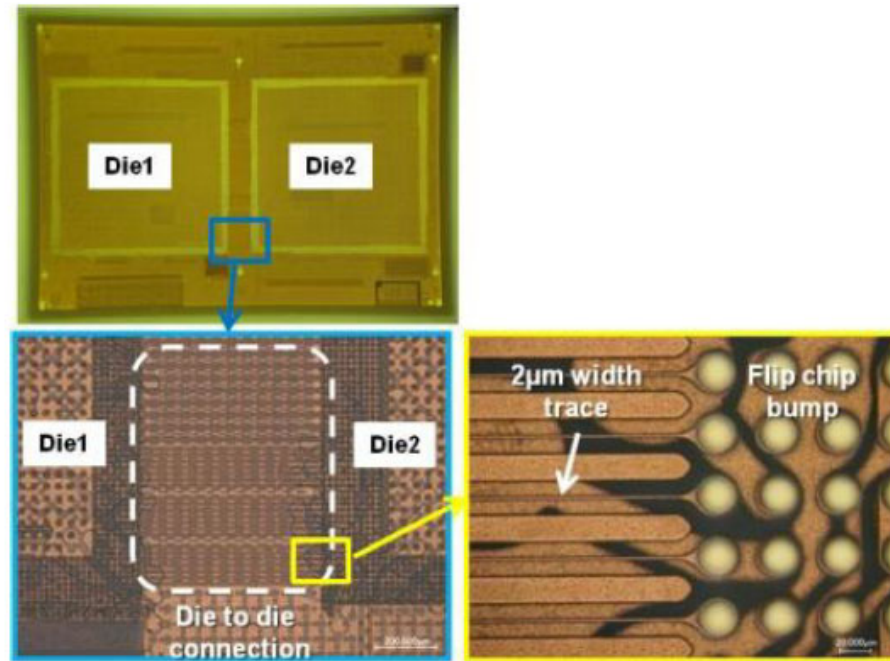


Figure 4. Die patterns on i-THOP

Oi, at Fig. 5.

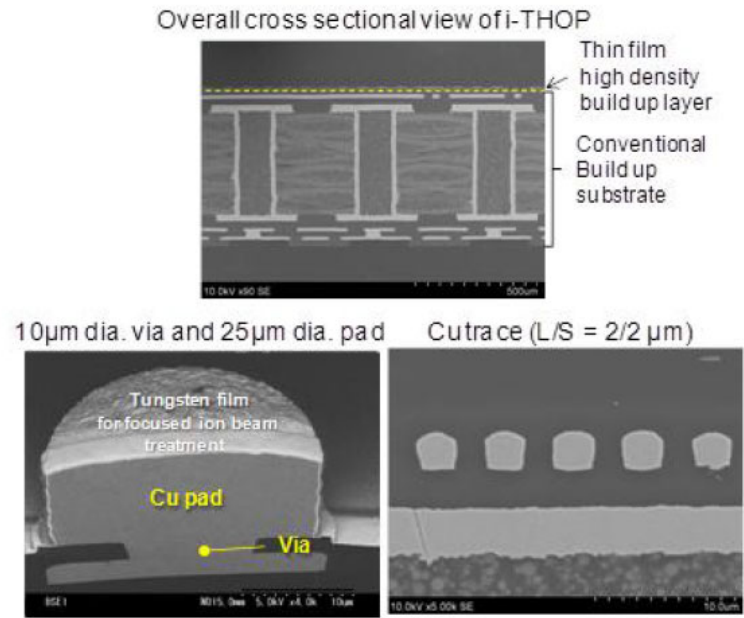
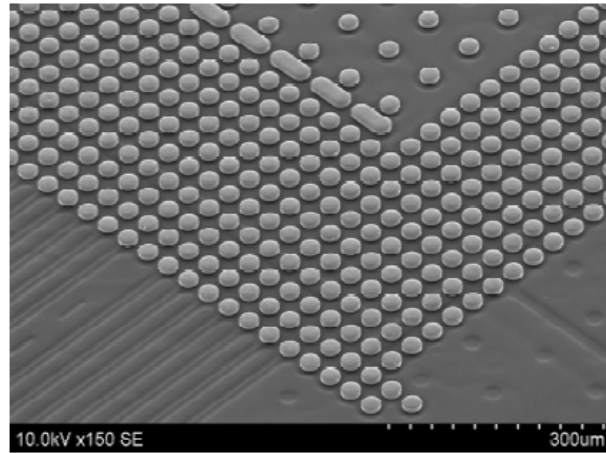


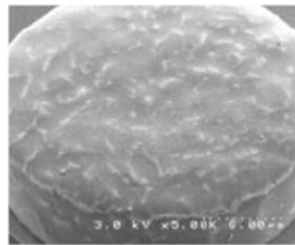
Figure 5. SEM image of i-THOP cross section

Oi, at Fig. 6.

A) 40 μ m pitch flip chip bonding pads

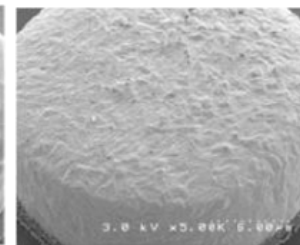


B) Enlarged image
Cu-OSP



Ra:0.30 μ m

Thin ENEPIG



Ra:0.31 μ m

Figure 6. SEM images of flip chip bonding pad.

Oi, at Fig.9.

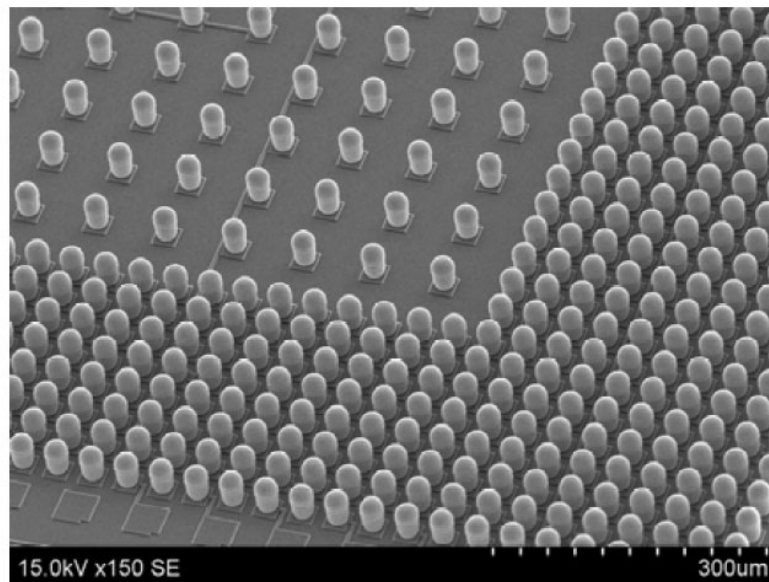


Figure 9. SEM image of test die

Oi, at Fig. 10

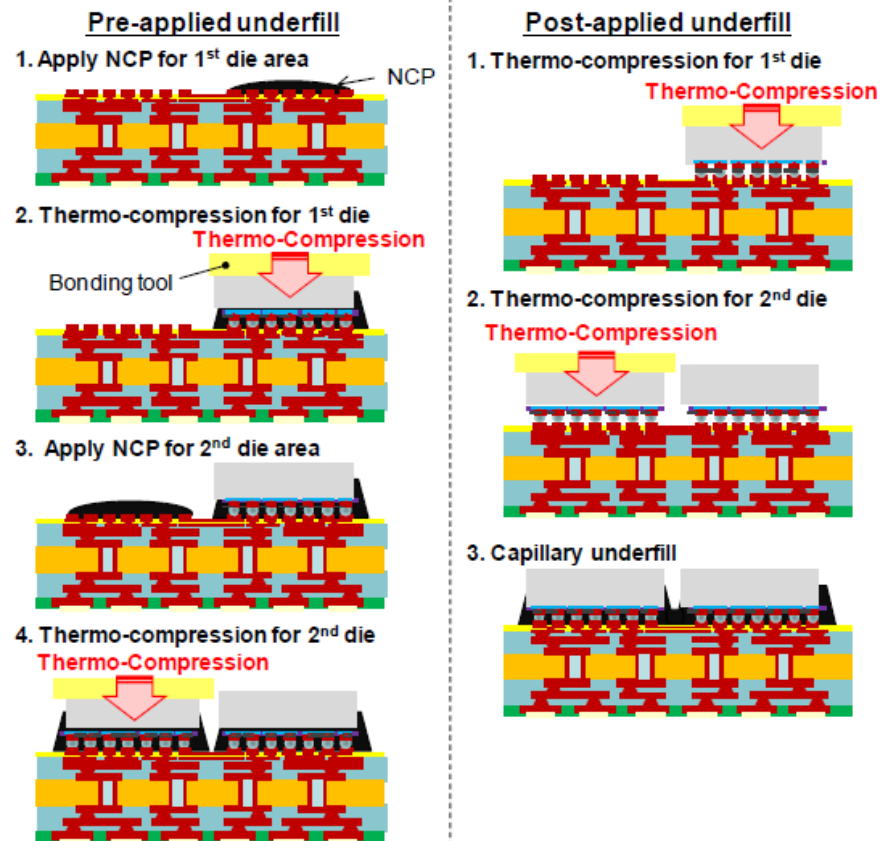


Figure 10. Evaluation process

Oi, at Abstract (“2.5D packaging technology utilizing silicon interposers is being developed and used for high-performance applications as the demand for miniaturization and higher density continues to increase. Silicon interposers enable very high density interconnects using standard semiconductor fabrication process technology, but are challenged as size increases. An alternative solution, high density laminate interposer, can offer advantages in cost, form factor, and infrastructure over other interposer options available today. However, it is difficult to manufacture high density wiring and flip chip bump pads on laminate interposer because

	conventional build-up technology limits fine line and via diameter. Our solution is the combination of an integrated organic interposer substrate with high density interconnects and thermo-compression flip chip bonding. Our solution eliminates the backside integration process for silicon interposer and assembly of interposers onto substrates. Conventional assembly processes can be utilized for assembling dies onto the integrated organic interposer substrate. Feasibility of this 2.5D package has been demonstrated by assembling dual-die with 40um pitch copper pillar bumps onto this novel integrated organic interposer substrate with 2µm line and space.”) Oi, at 348 (“Introduction Recent advancement in mobile applications such as smart phones and tablets are progressing rapidly. The demand on miniaturization of form factor and high performance are becoming much more important factors. Die partitioning enables miniaturization of small form factor and high performance without the yield issues of other methods. The focus of this paper is die partitioning technology conducted by fine pitch flip chip interconnect and die to die connection through fine lines. The materials under consideration for such an interposer are silicon, glass, and organic dielectrics. Silicon is a leading candidate for 2.5D packaging [1]-[3], but is challenged as size increases. Glass offers the promise of performance at lower cost than silicon, but there is no manufacturing infrastructure. Organic materials used in substrates have a prevalent supply chain, but cannot meet 2.5D line and space requirements that are finer than conventional laminate. By processing a fine-metal layer interposer directly on top of a conventional organic substrate, we have designed an integrated organic substrate and organic interposer. Our development progress has confirmed some basic properties [4], though we have not yet fully evaluated the die assembly. Key Points: • High accuracy alignment of die to interposer. • CTE mismatch and warpage control during soldering.
--	--

	• Elimination of underfill void. We have also measured thermal warpage of the organic interposer with die. We will report on the feasibility of our design for 2.5D packaging focusing on the assembly process.”) Oi, at 348 (“The Concept of Organic Interposer “ i-THOP” (Integrated Thin film High density Organic Package) Figure 1 shows a schematic cross-sectional image of i-THOP with die. Dual-die were assembled side by side and fine copper traces were used for die to die connection. A schematic image of die to die connections through fine line and fine pitched bump is shown in Figure 2. The pattern consisted of a thin film layer which is built on a conventional laminate substrate.”) This limitation also would have been obvious to a person of ordinary skill in the art based on (1) Mediatek FOMIP alone; (2) Mediatek FOMIP in combination with the general knowledge of a person of ordinary skill in the art; and/or (3) Mediatek FOMIP in combination with the teachings with respect to this claim element as detailed in NVIDIA’s other invalidity charts and cover pleading.
1[D]: wherein the passive crosslink is dedicated for inter-chiplet communications.	Including at least under Complainant’s apparent infringement theory, Mediatek FOMIP discloses and/or renders obvious this limitation. <i>See, e.g.,</i> <i>See</i> evidence cited for ’803 patent limitations 1[A]-1[C]. This limitation also would have been obvious to a person of ordinary skill in the art based on (1) Mediatek FOMIP alone; (2) Mediatek FOMIP in combination with the general knowledge of a person of ordinary skill in the art; and/or (3) Mediatek FOMIP in combination with the teachings with respect to this claim element as detailed in NVIDIA’s other invalidity charts and cover pleading.
2. The system of claim 1, wherein the passive crosslink comprises a passive interposer die.	Including at least under Complainant’s apparent infringement theory, Mediatek FOMIP discloses and/or renders obvious this limitation.

See, e.g.,

Lau

Lau, at Fig. 1.28

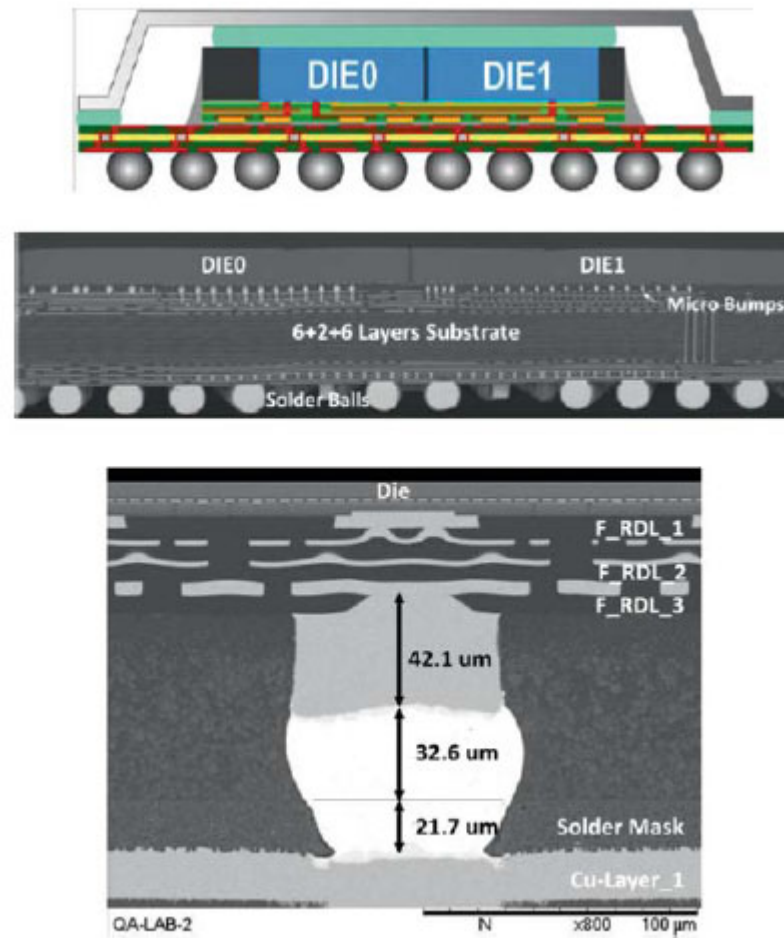


Fig. 1.28 MediaTek's fan-out RDL substrate

Lau, at 27 (“1.10.3 MediaTek’s RDLs by FOWLP

In 2016, MediaTek [133] proposed similar TSV-less interposer RDLs fabricated with FOWLP technology as shown in Fig. 1.28. Instead of the C4 bump, they used a microbump (Cu pillar + solder cap) to connect the bottom RDL to the 6-2-6 package substrate.”)

Chen-2016

Chen-2016, at 1495 (“Fig. 1 demonstrates the concept of FOWLP by using fan-out technology with ultra fine pitch interconnections for multi-chips integration on a 2.nD wafer level carrier. The wafer level carrier is attached on an organic BGA substrate by fine pitch Cu pillar bumping process. Passive component, such as de-coupling capacitor, and heat spreader are also integrated on the BGA substrate as system in package (SiP).

In this paper, a test vehicle built in TSMC N16 process was developed using FOWLP. Two homogeneous chips were bonded on a 2.nD wafer level carrier and connected with extremely fine pitch routings. The package design methodology is discussed in the Section II and III. Chip-package co-design methodology and its correlation results including signal integrity (SI) and power integrity (PI) considerations are discussed in the Section IV. Finally, the conclusion is shown in Section V.”)

Oi

Oi, at Abstract (“2.5D packaging technology utilizing silicon interposers is being developed and used for high-performance applications as the demand for miniaturization and higher density continues to increase. Silicon interposers enable very high density interconnects using standard semiconductor fabrication process technology, but are challenged as size increases. An alternative solution, high density laminate interposer, can offer advantages in cost, form factor, and infrastructure over other interposer options available today. However, it is difficult to manufacture high density wiring and flip chip bump pads on laminate interposer because conventional build-up technology limits fine line and via diameter. Our solution is the combination of an integrated organic interposer substrate with high density interconnects and thermo-compression flip chip bonding. Our solution eliminates the backside integration process for silicon interposer and assembly of interposers onto substrates. Conventional

	assembly processes can be utilized for assembling dies onto the integrated organic interposer substrate. Feasibility of this 2.5D package has been demonstrated by assembling dual-die with 40um pitch copper pillar bumps onto this novel integrated organic interposer substrate with 2μm line and space.”) Oi, at 348 (“Introduction Recent advancement in mobile applications such as smart phones and tablets are progressing rapidly. The demand on miniaturization of form factor and high performance are becoming much more important factors. Die partitioning enables miniaturization of small form factor and high performance without the yield issues of other methods. The focus of this paper is die partitioning technology conducted by fine pitch flip chip interconnect and die to die connection through fine lines. The materials under consideration for such an interposer are silicon, glass, and organic dielectrics. Silicon is a leading candidate for 2.5D packaging [1]-[3], but is challenged as size increases. Glass offers the promise of performance at lower cost than silicon, but there is no manufacturing infrastructure. Organic materials used in substrates have a prevalent supply chain, but cannot meet 2.5D line and space requirements that are finer than conventional laminate. By processing a fine-metal layer interposer directly on top of a conventional organic substrate, we have designed an integrated organic substrate and organic interposer. Our development progress has confirmed some basic properties [4], though we have not yet fully evaluated the die assembly. Key Points: • High accuracy alignment of die to interposer. • CTE mismatch and warpage control during soldering. • Elimination of underfill void.
--	--

	We have also measured thermal warpage of the organic interposer with die. We will report on the feasibility of our design for 2.5D packaging focusing on the assembly process.”) Oi, at 348 (“The Concept of Organic Interposer “ i-THOP” (Integrated Thin film High density Organic Package) Figure 1 shows a schematic cross-sectional image of i-THOP with die. Dual-die were assembled side by side and fine copper traces were used for die to die connection. A schematic image of die to die connections through fine line and fine pitched bump is shown in Figure 2. The pattern consisted of a thin film layer which is built on a conventional laminate substrate.”) This limitation also would have been obvious to a person of ordinary skill in the art based on (1) Mediatek FOMIP alone; (2) Mediatek FOMIP in combination with the general knowledge of a person of ordinary skill in the art; and/or (3) Mediatek FOMIP in combination with the teachings with respect to this claim element as detailed in NVIDIA’s other invalidity charts and cover pleading.
3[A]: The system of claim 1, wherein the first GPU chiplet comprises a first PHY region including conductor structures for chiplet-to-chiplet communications,	Including at least under Complainant’s apparent infringement theory, Mediatek FOMIP discloses and/or renders obvious this limitation. <i>See, e.g.,</i> <u>Lau</u> Lau, at Fig. 1.28

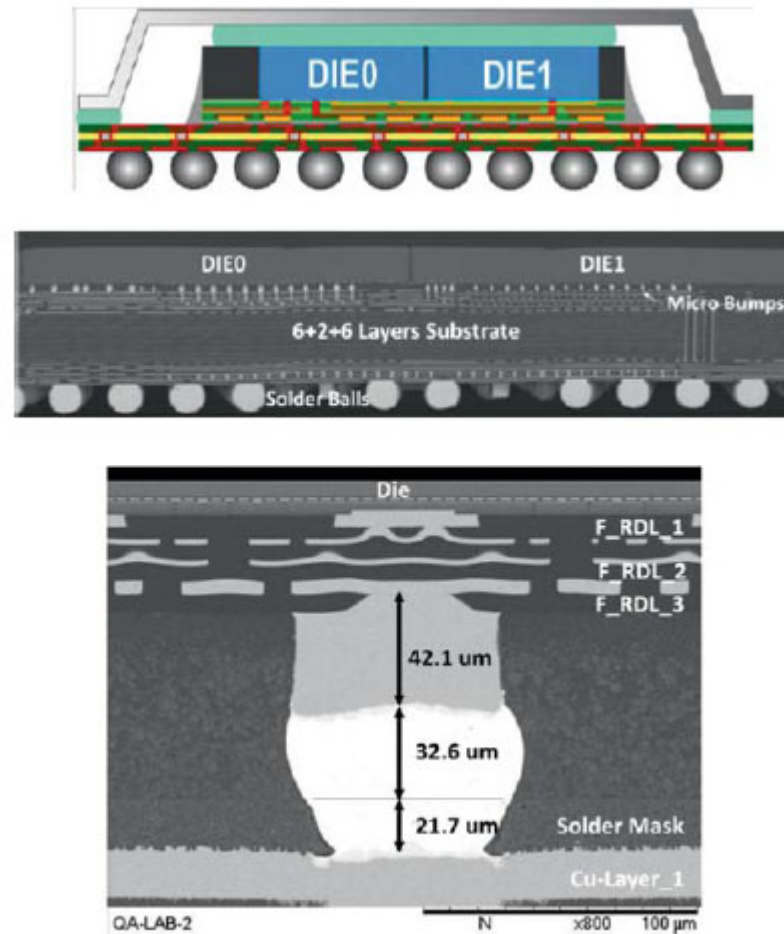


Fig. 1.28 MediaTek's fan-out RDL substrate

Lau, at 27 (“1.10.3 MediaTek’s RDLs by FOWLP

In 2016, MediaTek [133] proposed similar TSV-less interposer RDLs fabricated with FOWLP technology as shown in Fig. 1.28. Instead of the C4 bump, they used a

microbump (Cu pillar + solder cap) to connect the bottom RDL to the 6-2-6 package substrate.”)

Chen-2016

Chen-2016, at Fig.1.

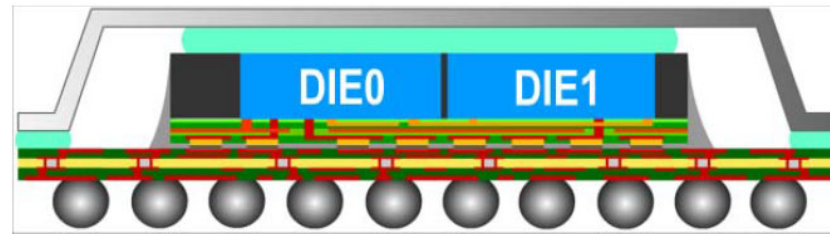


Fig.1 FOWLP process for multi-chips integration

Chen-2016, at Fig. 3.

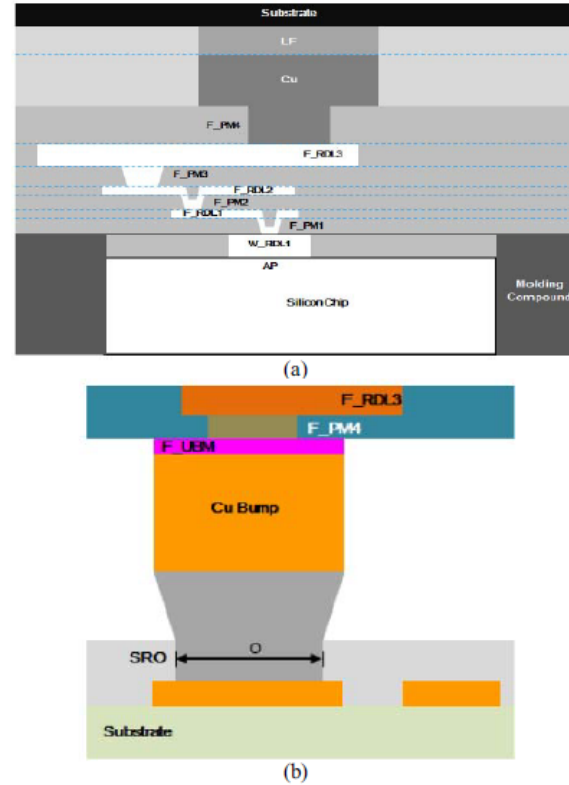


Fig.3 (a) FOWLP carrier structure, and (b) cu-pillar bump joined to BGA substrate

Chen-2016, at Fig.5.

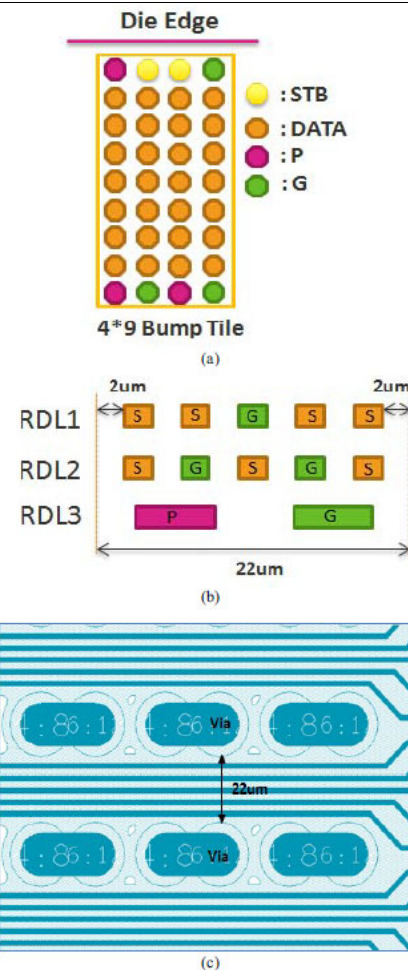


Fig.5 (a) Bump tile of interconnection bumps, (b) layout strategy of 4x9 bump matrix, and (c) an example of interconnection traces routing

Chen-2016, at Fig. 7.

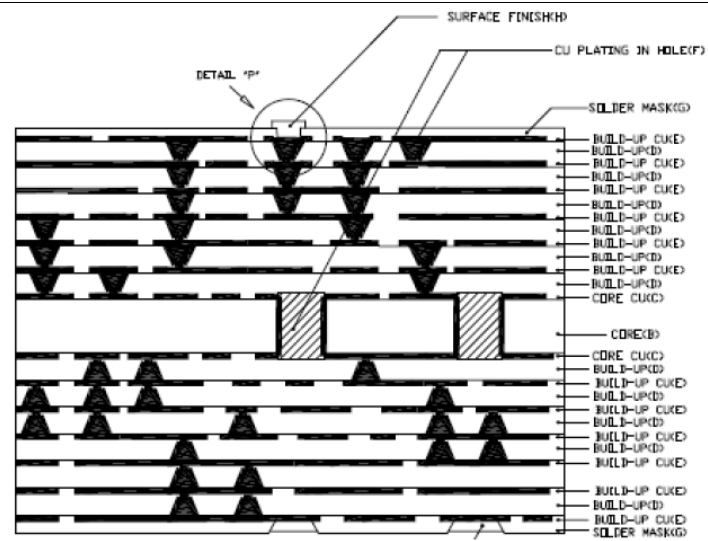
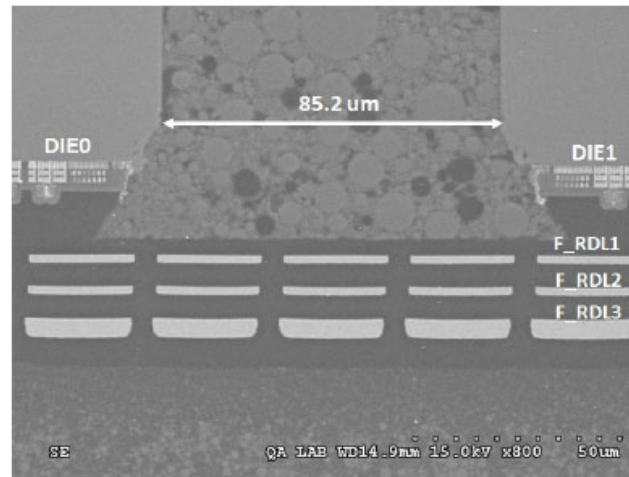
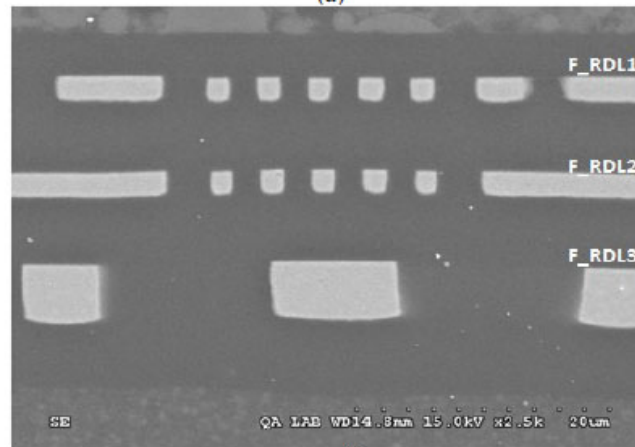


Fig.7 Stack-up structure of 6+2+6 layers BGA substrate

Chen-2016, at Fig. 9.



(a)



(b)

Fig.9 (a) Horizontal cut cross-section of a fan-out carrier die, and (b) the vertical cut cross-section of a fan-out carrier die.

Chen-2016, at Fig. 10.

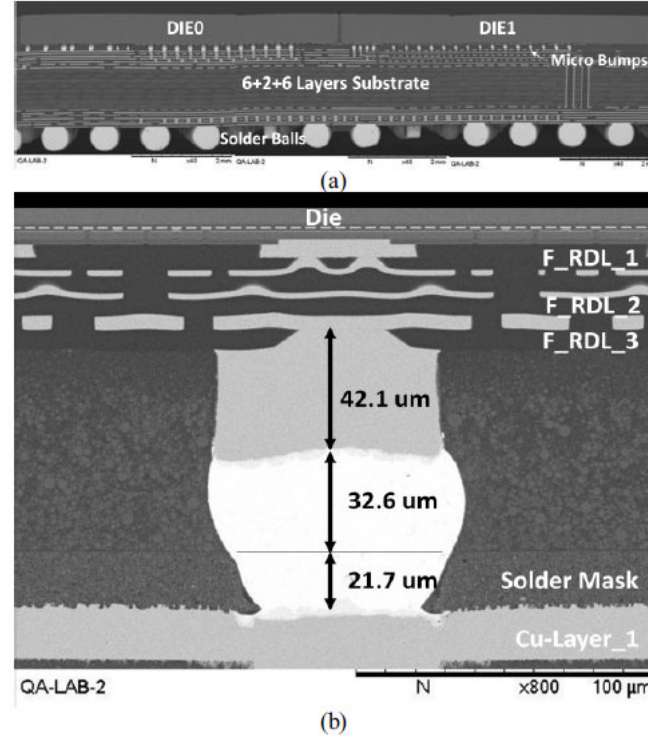


Fig.10 (a) Cross-section view of FOWLP, and (b) structure of micro bump connecting fan-out carrier and the BGA substrate.

Chen-2016, at Fig. 11.

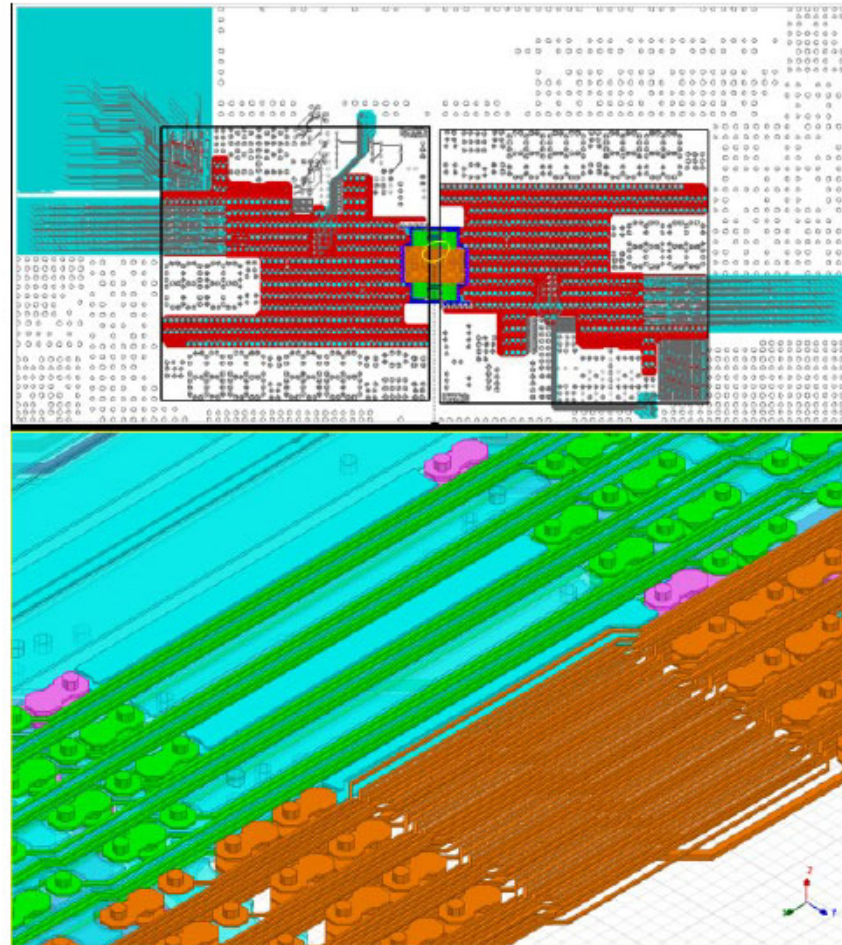


Fig.11 3D model of DIE to DIE interconnection

Chen-2016, at Abstract (“A novel multi-chips Fan-out Wafer Level Package (FOWLP) with fine pitch Cu pillar bumps was presented to accommodate volumes of I/O requirements on a BGA organic substrate. Due to the manufacturing limitation, cost and reliability consideration of fine-pitch BGA substrate, chips with extremely fine pitch RDL routing (width/space=2/2um) and a 2.nD wafer level carrier was integrated as system-in-package (SiP)

to resolve I/O capacity issue. Besides, multi-chips with small die size connected through wafer-level RDL was also developed to resolve low yield issue of large chip in advanced wafer process. In this paper, a comprehensive design methodology was introduced to cover fan-out WLP interconnection and BGA substrate design. A test vehicle built in TSMC N16 process was presented to verify the design methodology. With fully consideration on signal integrity (SI) and power integrity (PI), hundreds of interconnection I/Os with 4Gbps speed were achieved. Besides, high speed SERDES with 25Gbit/sec data rate was also verified to enable high speed and high bandwidth data center application.”)

Chen-2016, at 1495 (“Fig. 1 demonstrates the concept of FOWLP by using fan-out technology with ultra fine pitch interconnections for multi-chips integration on a 2.nD wafer level carrier. The wafer level carrier is attached on an organic BGA substrate by fine pitch Cu pillar bumping process. Passive component, such as de-coupling capacitor, and heat spreader are also integrated on the BGA substrate as system in package (SiP).

In this paper, a test vehicle built in TSMC N16 process was developed using FOWLP. Two homogeneous chips were bonded on a 2.nD wafer level carrier and connected with extremely fine pitch routings. The package design methodology is discussed in the Section II and III. Chip-package co-design methodology and its correlation results including signal integrity (SI) and power integrity (PI) considerations are discussed in the Section IV. Finally, the conclusion is shown in Section V.”)

Chen-2016, at 1496 (“III. FAN-OUT WLP AND BGA SUBSTRATE DESIGN

A. Fan-out WLP 2.nD design

A test chip was developed to verify FOWLP and its electrical performance. Three layers RDL are used for high speed signal fan-out routings and interconnections. Fig. 3(a) shows the structure of FOWLP carrier and Fig. 3(b) shows how the Cu-pillar bump joint to BGA substrate. Two silicon chips covered by molding compound through re-construction process and connected to a FOWLP carrier with micro bumps W_RDL1. F_RDL1, 2, 3 are three layers RDL for interconnection and fan-out between micro bumps and Cu-pillar bumps. F_PM1, 2, 3, 4 are the polyamide based dielectric layers and are placed between each RDL layer.

A UBM layer is placed above the 4th dielectric layer, F_PM4, for connection to BGA substrate by Cu-pillar (CuP) bumps. The CuP bumps are constructed by copper with lead-free (Sn-Ag) cap and the bump height is 70 nm with 130nm bump pitch. The floor plan of test chip is shown in Fig. 4. There are 4 identical dice integrated in the test chip and every 2 dice are connected by FOWLP. Fig. 4(a) shows the placement and connection of 2 homogeneous dice on a FOWLP carrier. The fan-out area is mainly for high speed SERDES signals which are directly connected from micro bumps to CuP bumps with very short routing traces to avoid signal quality loss. The interconnection area on FOWLP carrier is shown in Fig. 4(b). The distance between 2 dice is 100um and 414 small bumps are designed on each die for signal interconnections. Fine-pitch trace with minimum trace width/space=2/2um is used on the 1st and 2nd RDL layer while 5/5um is used on 3rd RDL layer.”)

Chen-2016, at 1497 (“Total 2 carrier dice (DIE0/1 and DIE2/3) are mounted on BGA substrate and connected to each other by substrate routings. The interconnection area between 2 dice consists of several slices of bump tile and each slice consists of 4x9 bump tile including 28 data bits, 2 strobes and 6 power/ground. Fig. 5(a) shows the 1 slice of bump tile for interconnection. Each column of bump tile has 7 bumps and is routed on F_RDL1 and F_RDL2 layers for signal routing. The bump pitch is 44 um and small size bumps with diameter equals 30um is adopted. Since the space between 2 vias is 22um, it allows maximum 5 traces routing between 2 vias based on minimum trace design rules (width/space=2/2um). Therefore, it allows total 10 traces routing on F_RDL1 & 2. 7 traces are used for signal and the rest 3 traces are used for GND shielding. The power and ground traces are routed on F_RDL3 for wider trace width and spacing requirement. Fig. 5(b) shows the layout topology of 4x9 bump tile. The ground trace is arranged between signals for signal integrity consideration. Fig. 5(c) demonstrates the real routing of interconnection traces on F_RDL1 & 2.”)

Chen-2016, at 1498-99 (“C. Structures of fan-out carrier and BGA substrate

A physical examination was performed to verify the test chip structure of the fan-out carrier and BGA substrate. Fig. 8 shows the top view of the test chip. There are 2 fan-out carrier dice on the substrate and each carrier die contains 2 identical dice (DIE0/1 and DIE2/3). The detail horizontal cut cross-section diagram of a fan-out carrier die (DIE0/1) is shown in Fig. 9(a). The average metal thickness of the interconnection traces is 2.04um and 4.23um for

F_RDL1/2 and F_RDL3 respectively. Fig. 9(b) shows the vertical cut cross-section diagram between DIE0 and DIE1. The average measurement results of trace width and space are 2.09um/2.00um which is within the design tolerance. Fig. 10(a) shows the overall cross-section structure of FOWLP SiP. DIE0 and DIE1 are connected by a 3-RDL carrier and then mount to a 6+2+6 layers BGA substrate by micro bumps. Fig. 10(b) shows the structure of micro bump for between connecting fan-out carrier and the BGA substrate. The bump height after reflow process is about 96.4 um and matches to the design as shown on Fig. 3(b). Based on the experiment verification on structure cross-section, the process of FOWLP are stable and assured for mass production.”)

Chen-2016, at 1499 (“IV. CHIP-PKG CO-DESIGN AND CORRELATION

In this test chip, bandwidth with 10Tbps throughput is necessary and thus 10,000 parallel buses with data transceiver working at 1Gbps are required. One way to reduce the interconnection between two homogenous dice is to raise the data rate of data transceiver. However, it may suffer severe signal and power integrity issue without appropriate design. Thus, a timing budget analysis of the complete data path between transceiver to receiver is shown in Table I. Here, we named the data path of proposed fan-out chip as MediaTek I/F Link (MIL). The time skew of the MIL must be smaller than 20.5ps and 83ps for 4Gbps and 3.2Gbps operation, respectively. Through systematic and comprehensive co-simulation methodology, a 4x9 bump tile with 28 data bits, 2 strobes and 6 power/ground was proposed to optimize the SI/PI performance. The shielding guard traces could be appropriately implemented with the proposed bump tile to mitigate the crosstalk noise induced from adjacent aggressors. Fig. 11 shows the 3D simulation platform of DIE-to-DIE interconnection. The signal integrity as well as the power integrity can be ensured through co-simulation for high data rate requirement. In order to maintain the target bandwidth of 10Tbps, total 90 bump tiles were duplicated.”)

Chen-2016, at 1500 (“The 2.nD FOWLP successfully connected two homogeneous dice using ultra fine pitch process. This study discussed the design experience and concepts of wafer level carrier. The 2/2 um trace width and space realized how to routing 10 traces on 2 RDL layers with very fine pitch (44 um) bump placement. In additions, the 6+2+6 layers BGA substrate design is also important to ensure the signal integrity meets its requirement on a 55x55mm

PKG. About the experiment on real samples, we know the structure of FOWLP and the cross-section diagrams help verify the process capability of the wafer level carrier can meet our design spec and requirements. Based on the Chip-PKG co-design and correlation results, the 2.nD FOWLP can handle bandwidth with 10Tbps throughput and 25GHz high speed SERDES signals.

Therefore, the FOWLP interconnection methodology was approved to connect two homogeneous dice which were separated from a huge die. It really helps us to increase the gross die rate and improve the wafer yield when the silicon die size is usually enormous and suffers low yield issues, especially in advanced wafer process.”)

Oi

Oi, at Fig.1.

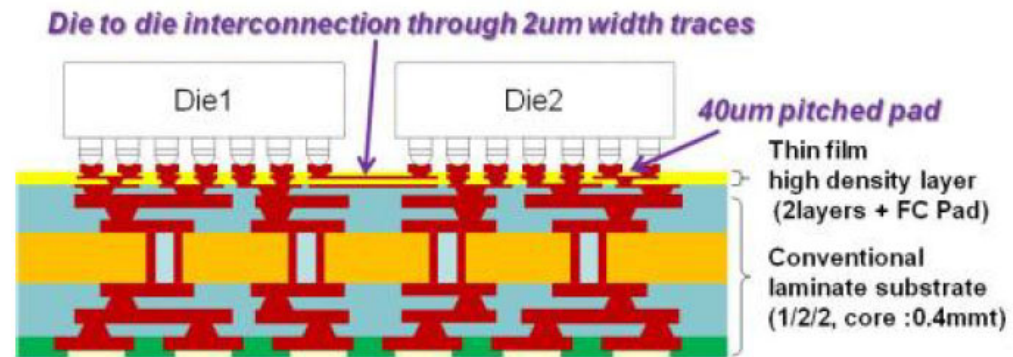


Figure 1. Schematic cross sectional image of i-THOP

Oi, at Fig.2.

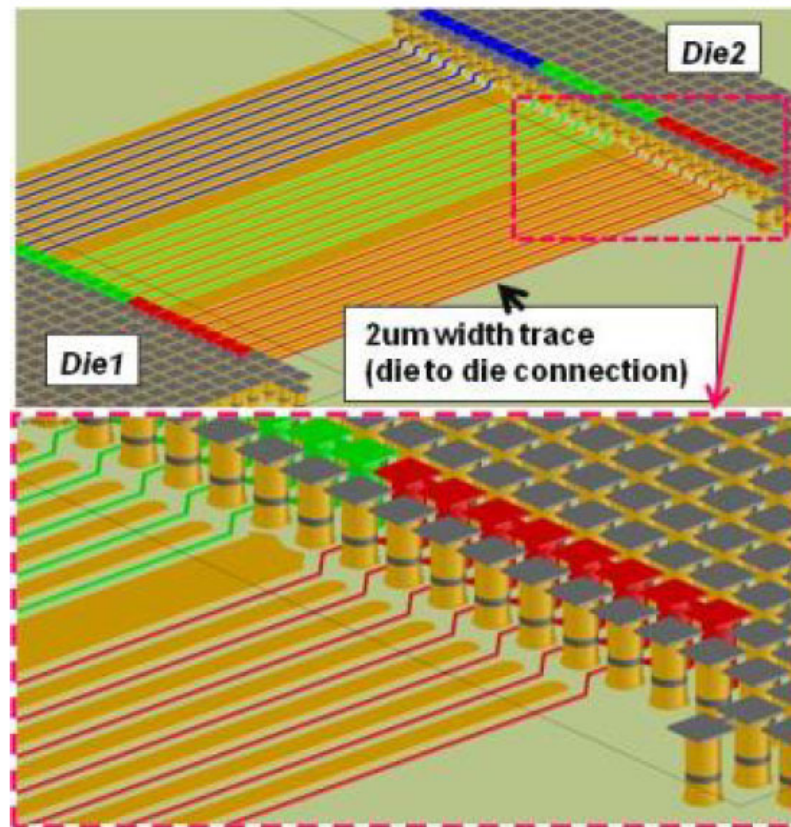


Figure 2. Die to die connection through 2 μ m width copper trace.

Oi, at Fig. 4.

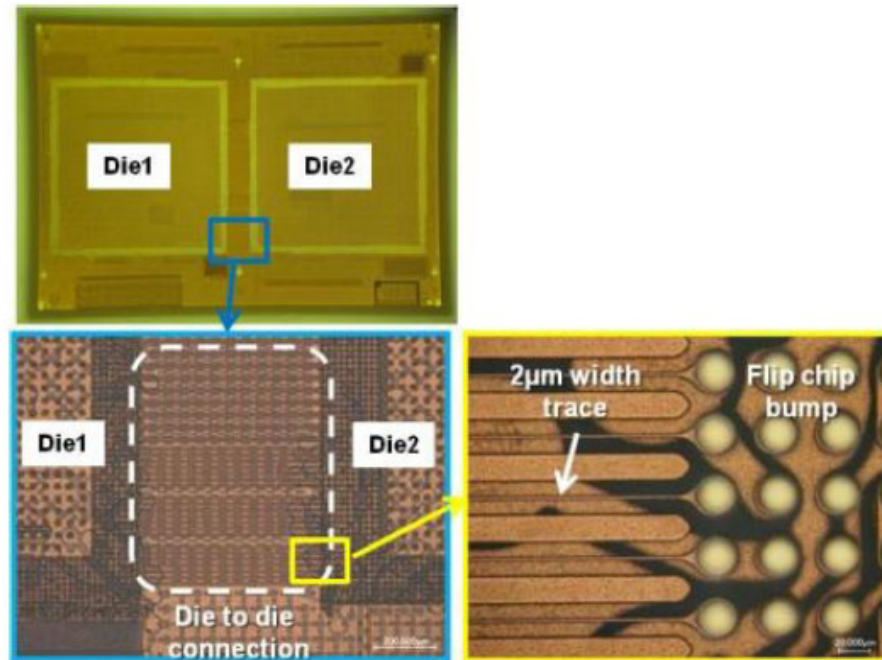


Figure 4. Die patterns on i-THOP

Oi, at Fig. 5.

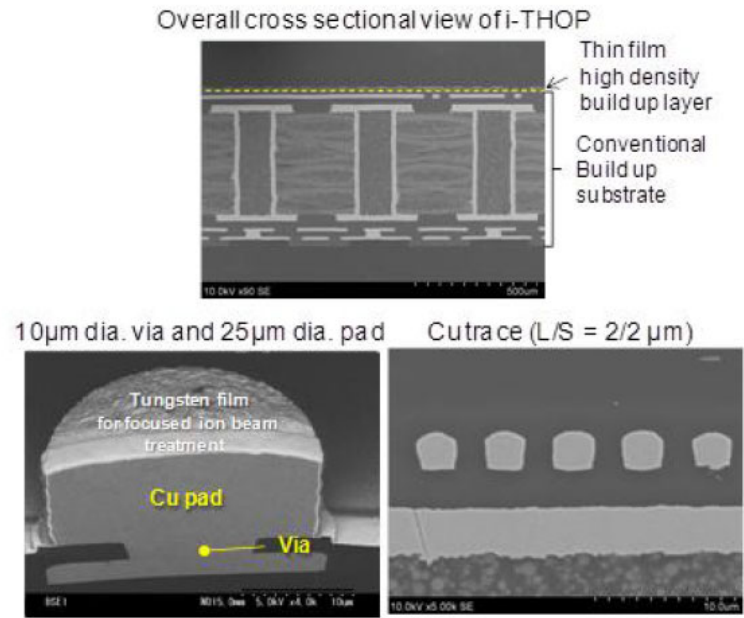
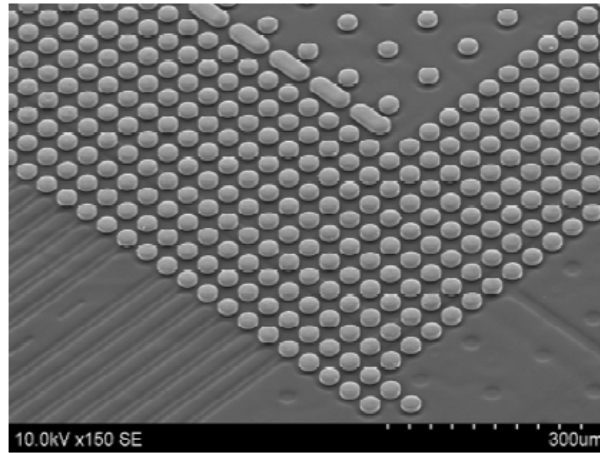


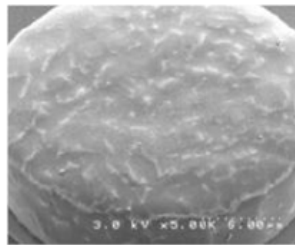
Figure 5. SEM image of i-THOP cross section

Oi, at Fig. 6.

A) 40 μ m pitch flip chip bonding pads

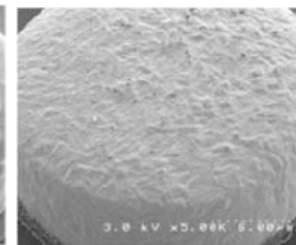


B) Enlarged image
Cu-OSP



Ra:0.30 μ m

Thin ENEPIG



Ra:0.31 μ m

Figure 6. SEM images of flip chip bonding pad.

Oi, at Fig. 8.

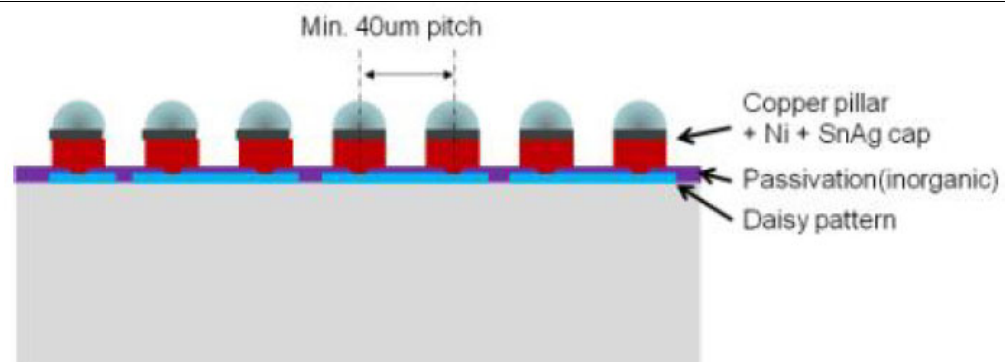


Figure 8. Schematic image of the test die.

Oi, at Fig.9.

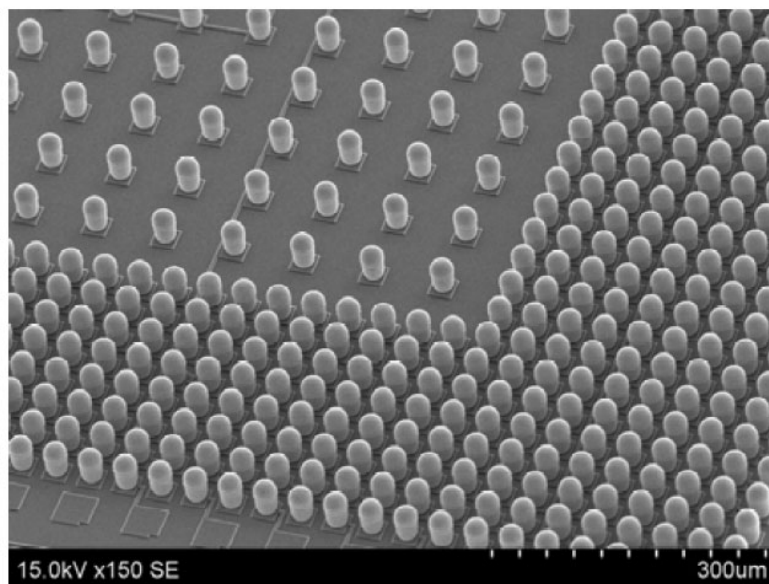


Figure 9. SEM image of test die

	Oi, at Abstract (“2.5D packaging technology utilizing silicon interposers is being developed and used for high-performance applications as the demand for miniaturization and higher density continues to increase. Silicon interposers enable very high density interconnects using standard semiconductor fabrication process technology, but are challenged as size increases. An alternative solution, high density laminate interposer, can offer advantages in cost, form factor, and infrastructure over other interposer options available today. However, it is difficult to manufacture high density wiring and flip chip bump pads on laminate interposer because conventional build-up technology limits fine line and via diameter. Our solution is the combination of an integrated organic interposer substrate with high density interconnects and thermo-compression flip chip bonding. Our solution eliminates the backside integration process for silicon interposer and assembly of interposers onto substrates. Conventional assembly processes can be utilized for assembling dies onto the integrated organic interposer substrate. Feasibility of this 2.5D package has been demonstrated by assembling dual-die with 40um pitch copper pillar bumps onto this novel integrated organic interposer substrate with 2μm line and space.”) Oi, at 348 (“The Concept of Organic Interposer “ i-THOP” (Integrated Thin film High density Organic Package) Figure 1 shows a schematic cross-sectional image of i-THOP with die. Dual-die were assembled side by side and fine copper traces were used for die to die connection. A schematic image of die to die connections through fine line and fine pitched bump is shown in Figure 2. The pattern consisted of a thin film layer which is built on a conventional laminate substrate.”) This limitation also would have been obvious to a person of ordinary skill in the art based on (1) Mediatek FOMIP alone; (2) Mediatek FOMIP in combination with the general knowledge of a person of ordinary skill in the art; and/or (3) Mediatek FOMIP in combination with the teachings with respect to this claim element as detailed in NVIDIA’s other invalidity charts and cover pleading.
3[B]: and further wherein the second GPU chiplet comprises a second PHY region including	Including at least under Complainant’s apparent infringement theory, Mediatek FOMIP discloses and/or renders obvious this limitation. <i>See, e.g.,</i>

conductor structures for chiplet-to-chiplet communications.	<i>See</i> evidence cited for '803 patent limitation 3[A]. This limitation also would have been obvious to a person of ordinary skill in the art based on (1) Mediatek FOMIP alone; (2) Mediatek FOMIP in combination with the general knowledge of a person of ordinary skill in the art; and/or (3) Mediatek FOMIP in combination with the teachings with respect to this claim element as detailed in NVIDIA's other invalidity charts and cover pleading.
7. The system of claim 1, wherein the passive crosslink communicably couples all GPU chiplets in the GPU chiplet array.	Including at least under Complainant's apparent infringement theory, Mediatek FOMIP discloses and/or renders obvious this limitation. <i>See, e.g.,</i> <i>See</i> evidence cited for '803 patent limitations 1[D]. This limitation also would have been obvious to a person of ordinary skill in the art based on (1) Mediatek FOMIP alone; (2) Mediatek FOMIP in combination with the general knowledge of a person of ordinary skill in the art; and/or (3) Mediatek FOMIP in combination with the teachings with respect to this claim element as detailed in NVIDIA's other invalidity charts and cover pleading.
8[A]: The system of claim 1, further comprising: a first cache memory hierarchy at the first GPU chiplet,	Including at least under Complainant's apparent infringement theory, Mediatek FOMIP discloses and/or renders obvious this limitation. <i>See, e.g.,</i> <i>See</i> evidence cited for '803 patent claims 1-7. This limitation also would have been obvious to a person of ordinary skill in the art based on (1) Mediatek FOMIP alone; (2) Mediatek FOMIP in combination with the general knowledge of a person of ordinary skill in the art; and/or (3) Mediatek FOMIP in combination with the teachings with respect to this claim element as detailed in NVIDIA's other invalidity charts and cover pleading.

8[B]: wherein a first level of the first cache memory hierarchy is coherent within the first GPU chiplet; and	Including at least under Complainant's apparent infringement theory, Mediatek FOMIP discloses and/or renders obvious this limitation. <i>See, e.g.,</i> <i>See</i> evidence cited for '803 patent limitation 8[A]. This limitation also would have been obvious to a person of ordinary skill in the art based on (1) Mediatek FOMIP alone; (2) Mediatek FOMIP in combination with the general knowledge of a person of ordinary skill in the art; and/or (3) Mediatek FOMIP in combination with the teachings with respect to this claim element as detailed in NVIDIA's other invalidity charts and cover pleading.
8[C]: a second cache memory hierarchy at the second GPU chiplet,	Including at least under Complainant's apparent infringement theory, Mediatek FOMIP discloses and/or renders obvious this limitation. <i>See, e.g.,</i> <i>See</i> evidence cited for '803 patent limitation 8[A]. This limitation also would have been obvious to a person of ordinary skill in the art based on (1) Mediatek FOMIP alone; (2) Mediatek FOMIP in combination with the general knowledge of a person of ordinary skill in the art; and/or (3) Mediatek FOMIP in combination with the teachings with respect to this claim element as detailed in NVIDIA's other invalidity charts and cover pleading.
8[D]: wherein a first level of the second cache memory hierarchy is coherent within the second GPU chiplet.	Including at least under Complainant's apparent infringement theory, Mediatek FOMIP discloses and/or renders obvious this limitation. <i>See, e.g.,</i> <i>See</i> evidence cited for '803 patent limitation 8[B]. This limitation also would have been obvious to a person of ordinary skill in the art based on (1) Mediatek FOMIP alone; (2) Mediatek FOMIP in combination with the general knowledge of a person of ordinary skill in the art; and/or (3) Mediatek FOMIP in combination with the

	teachings with respect to this claim element as detailed in NVIDIA's other invalidity charts and cover pleading.
9[A]: The system of claim 8, further comprising: a unified cache memory including both a last level of the first cache memory hierarchy and a last level of the second cache memory hierarchy,	Including at least under Complainant's apparent infringement theory, Mediatek FOMIP discloses and/or renders obvious this limitation. <i>See, e.g.,</i> <i>See also</i> evidence cited for '803 patent claim 8. This limitation also would have been obvious to a person of ordinary skill in the art based on (1) Mediatek FOMIP alone; (2) Mediatek FOMIP in combination with the general knowledge of a person of ordinary skill in the art; and/or (3) Mediatek FOMIP in combination with the teachings with respect to this claim element as detailed in NVIDIA's other invalidity charts and cover pleading.
9[B]: wherein the unified cache memory is coherent across all GPU chiplets of the GPU chiplet array.	Including at least under Complainant's apparent infringement theory, Mediatek FOMIP discloses and/or renders obvious this limitation. <i>See, e.g.,</i> <i>See</i> evidence cited for '803 patent claim 8. This limitation also would have been obvious to a person of ordinary skill in the art based on (1) Mediatek FOMIP alone; (2) Mediatek FOMIP in combination with the general knowledge of a person of ordinary skill in the art; and/or (3) Mediatek FOMIP in combination with the teachings with respect to this claim element as detailed in NVIDIA's other invalidity charts and cover pleading.
10.The system of claim 1, further comprising: a plurality of conductive pillars coupling a circuit board substrate to a first non-PHY region of the first GPU	Including at least under Complainant's apparent infringement theory, Mediatek FOMIP discloses and/or renders obvious this limitation. <i>See, e.g.,</i> <u>Lau</u>

chiplet and a second non-PHY region of the second GPU chiplet.

Lau, at Fig. 1.28

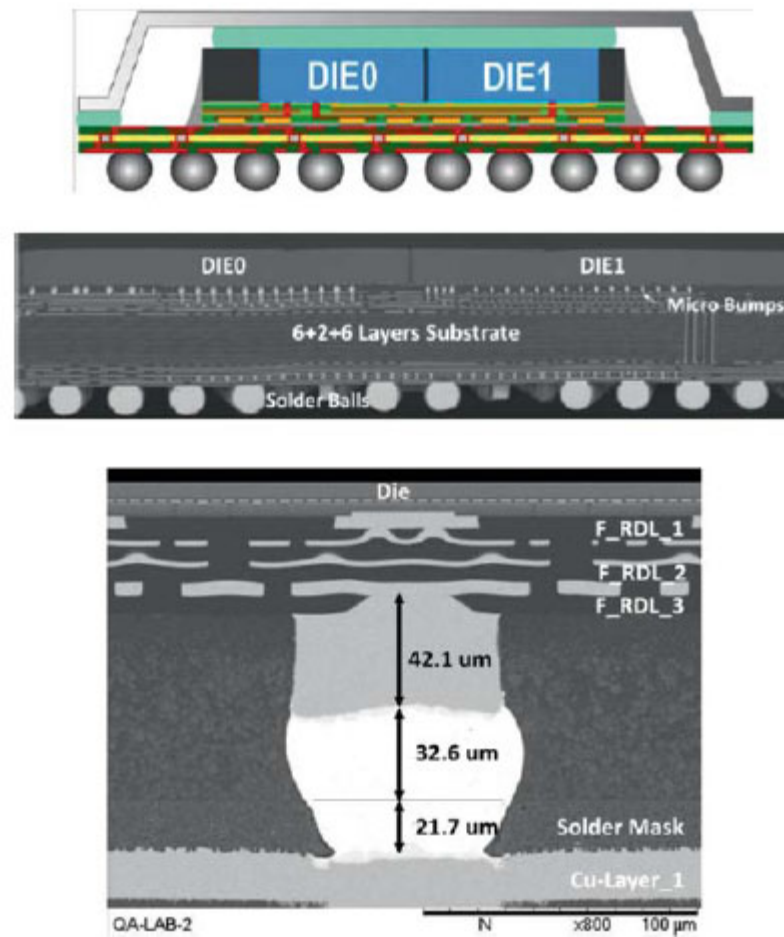


Fig. 1.28 MediaTek's fan-out RDL substrate

Lau, at 27 ("1.10.3 MediaTek's RDLs by FOWLP

In 2016, MediaTek [133] proposed similar TSV-less interposer RDLs fabricated with FOWLP technology as shown in Fig. 1.28. Instead of the C4 bump, they used a microbump (Cu pillar + solder cap) to connect the bottom RDL to the 6-2-6 package substrate.”)

Chen-2016

Chen-2016, at Fig.1.

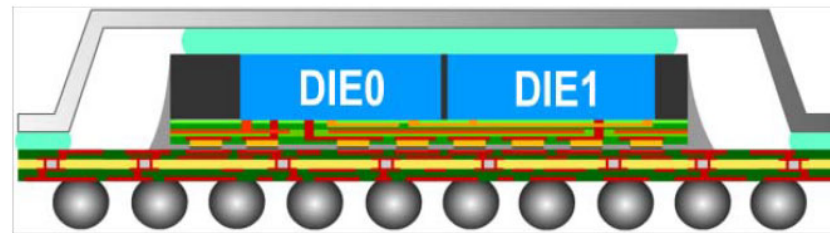
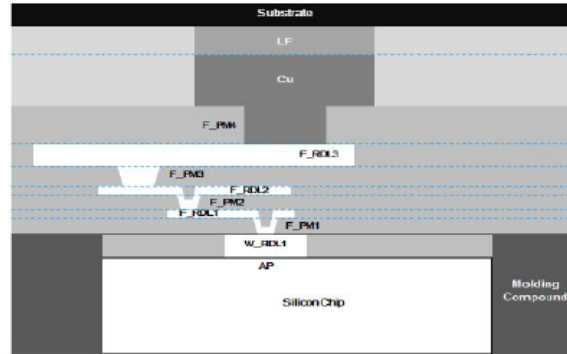
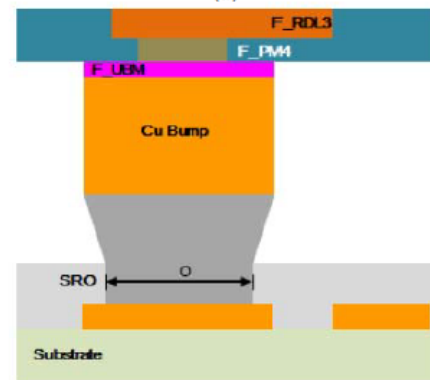


Fig.1 FOWLP process for multi-chips integration

Chen-2016, at Fig. 3.



(a)



(b)

Fig.3 (a) FOWLP carrier structure, and (b) cu-pillar bump jointed to BGA substrate

Chen-2016, at Fig. 7.

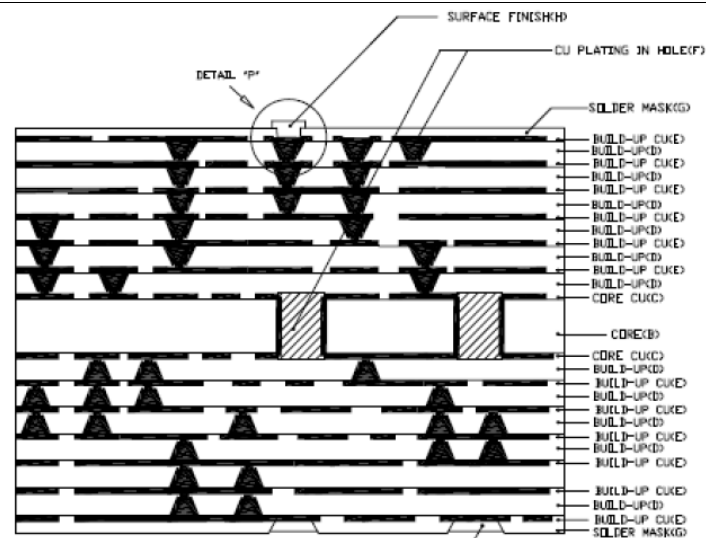


Fig.7 Stack-up structure of 6+2+6 layers BGA substrate

Chen-2016, at Abstract (“A novel multi-chips Fan-out Wafer Level Package (FOWLP) with fine pitch Cu pillar bumps was presented to accommodate volumes of I/O requirements on a BGA organic substrate. Due to the manufacturing limitation, cost and reliability consideration of fine-pitch BGA substrate, chips with extremely fine pitch RDL routing (width/space=2/2um) and a 2.nD wafer level carrier was integrated as system-in-package (SiP) to resolve I/O capacity issue. Besides, multi-chips with small die size connected through wafer-level RDL was also developed to resolve low yield issue of large chip in advanced wafer process. In this paper, a comprehensive design methodology was introduced to cover fan-out WLP interconnection and BGA substrate design. A test vehicle built in TSMC N16 process was presented to verify the design methodology. With fully consideration on signal integrity (SI) and power integrity (PI), hundreds of interconnection I/Os with 4Gbps speed were achieved. Besides, high speed SERDES with 25Gbit/sec data rate was also verified to enable high speed and high bandwidth data center application.”)

Chen-2016, at 1495 (“Fig. 1 demonstrates the concept of FOWLP by using fan-out technology with ultra fine pitch interconnections for multi-chips integration on a 2.nD wafer level carrier. The wafer level carrier is attached on an organic BGA substrate by fine pitch Cu pillar

bumping process. Passive component, such as de-coupling capacitor, and heat spreader are also integrated on the BGA substrate as system in package (SiP).

Chen-2016, at 1496 (“III. FAN-OUT WLP AND BGA SUBSTRATE DESIGN

A. Fan-out WLP 2.nD design

A test chip was developed to verify FOWLP and its electrical performance. Three layers RDL are used for high speed signal fan-out routings and interconnections. Fig. 3(a) shows the structure of FOWLP carrier and Fig. 3(b) shows how the Cu-pillar bump joint to BGA substrate. Two silicon chips covered by molding compound through re-construction process and connected to a FOWLP carrier with micro bumps W_RDL1. F_RDL1, 2, 3 are three layers RDL for interconnection and fan-out between micro bumps and Cu-pillar bumps. F_PM1, 2, 3, 4 are the polyamide based dielectric layers and are placed between each RDL layer.

A UBM layer is placed above the 4th dielectric layer, F_PM4, for connection to BGA substrate by Cu-pillar (CuP) bumps. The CuP bumps are constructed by copper with lead-free (Sn-Ag) cap and the bump height is 70 nm with 130nm bump pitch. The floor plan of test chip is shown in Fig. 4. There are 4 identical dice integrated in the test chip and every 2 dice are connected by FOWLP. Fig. 4(a) shows the placement and connection of 2 homogeneous dice on a FOWLP carrier. The fan-out area is mainly for high speed SERDES signals which are directly connected from micro bumps to CuP bumps with very short routing traces to avoid signal quality loss. The interconnection area on FOWLP carrier is shown in Fig. 4(b). The distance between 2 dice is 100um and 414 small bumps are designed on each die for signal interconnections. Fine-pitch trace with minimum trace width/space=2/2um is used on the 1st and 2nd RDL layer while 5/5um is used on 3rd RDL layer.”)

Oi

, at Fig.1.

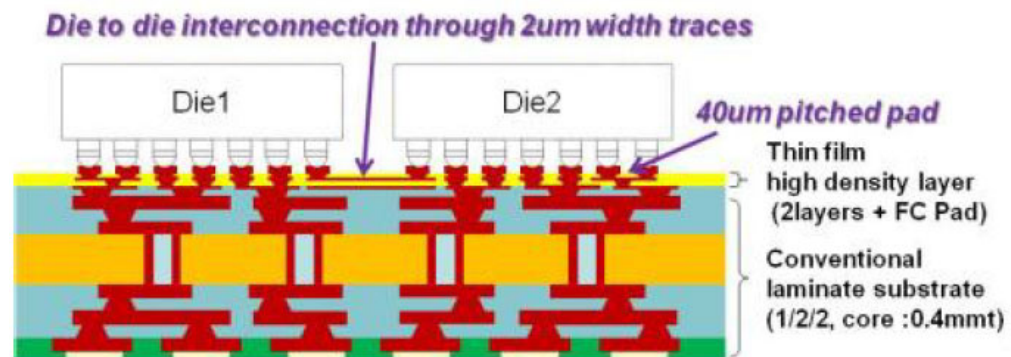


Figure 1. Schematic cross sectional image of i-THOP

Oi, at Fig.2.

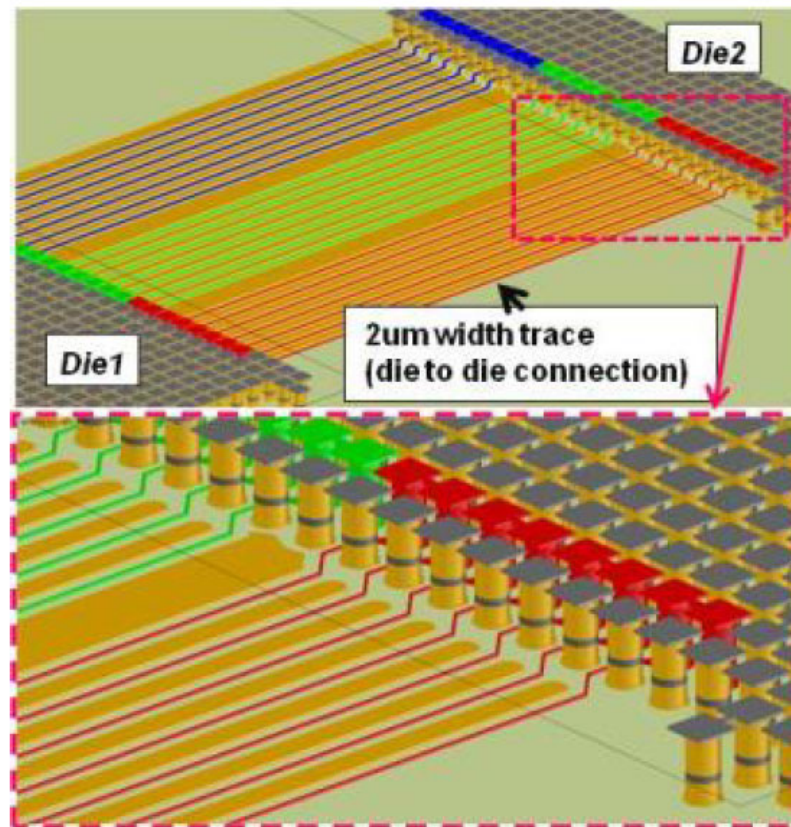


Figure 2. Die to die connection through 2 μ m width copper trace.

Oi, at Fig. 4.

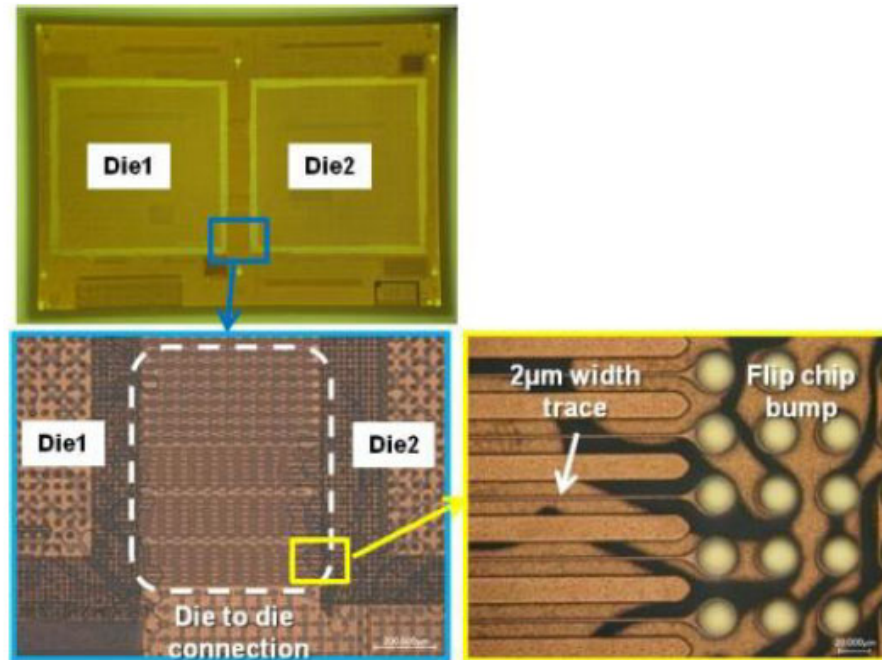


Figure 4. Die patterns on i-THOP

Oi, at Fig. 5.

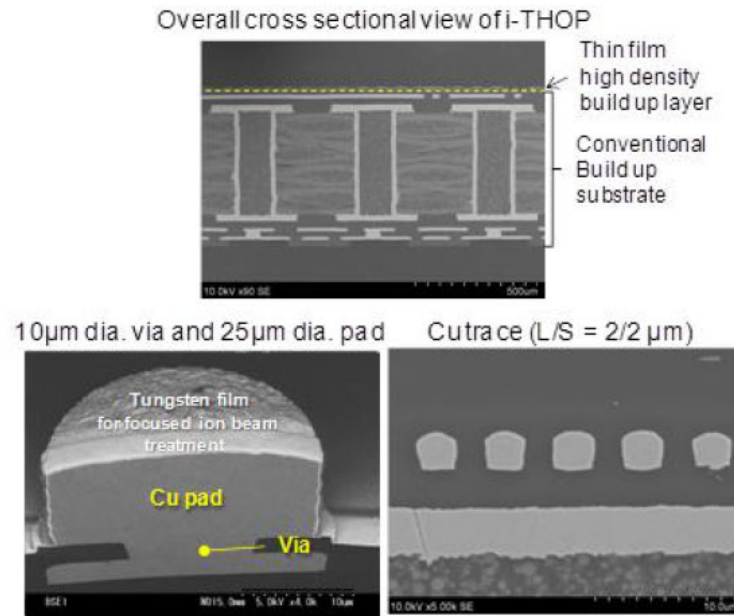


Figure 5. SEM image of i-THOP cross section

Oi, at Abstract (“Abstract

2.5D packaging technology utilizing silicon interposers is being developed and used for high-performance applications as the demand for miniaturization and higher density continues to increase. Silicon interposers enable very high density interconnects using standard semiconductor fabrication process technology, but are challenged as size increases. An alternative solution, high density laminate interposer, can offer advantages in cost, form factor, and infrastructure over other interposer options available today. However, it is difficult to manufacture high density wiring and flip chip bump pads on laminate interposer because conventional build-up technology limits fine line and via diameter. Our solution is the combination of an integrated organic interposer substrate with high density interconnects and thermo-compression flip chip bonding. Our solution eliminates the backside integration process for silicon interposer and assembly of interposers onto substrates. Conventional assembly processes can be utilized for assembling dies onto the integrated organic interposer

	substrate. Feasibility of this 2.5D package has been demonstrated by assembling dual-die with 40um pitch copper pillar bumps onto this novel integrated organic interposer substrate with 2um line and space.”) Oi, at 350 (“1) Pre-applied underfill process: Figure 11 shows cross sectional images of solder joints formed with different bonding forces. When bonding force was low at 0.55gf/bump, the solder at the top of copper pillar did not make sufficient connection to the pad. When the bonding force was high at 1.09gf/bump, undulation of the copper pad and copper trace under a pad caused via cracks. Taking into account the results, we considered that it was difficult to have enough process margin for the connection. In addition, an optimization of the NCP material would be required.”) Oi, at 351 (“3) Process optimization for flip chip bonding: Solder thickness between copper pillar and copper trace was found to be one of the most important factors to form a good solder joint. If the solder thickness was out of control, insufficient solder wetting or solder bridging could occur. Figure 13 shows the relationship between the post-bond solder thickness and bonding force. The solder thickness was found to decrease with increasing bonding force. When bonding force was higher than 0.4gf/bump, the solder thickness saturated around 3.0um. There was no difference in the solder thickness between the two surface finishes tested.”) This limitation also would have been obvious to a person of ordinary skill in the art based on (1) Mediatek FOMIP alone; (2) Mediatek FOMIP in combination with the general knowledge of a person of ordinary skill in the art; and/or (3) Mediatek FOMIP in combination with the teachings with respect to this claim element as detailed in NVIDIA’s other invalidity charts and cover pleading.
--	---