半導体パッケージの 10 年後に向けての課題

西尾 俊彦*

Semiconductor Package Challenges toward 10 Years Later

Toshihiko NISHIO*

* 株式会社 SBR テクノロジー (〒 124-0014 東京都葛飾区東四つ木 3-2-12)

* SBR Technology Co., Ltd. (3-2-12, Higashi-yotsugi, Katsushika, Tokyo 124-0014)

はじめに

3 年前の JIEP の特集号¹⁾にて, “2.1, 2.5 および 3D パッケージの大きなマーケットは本当に存在するのか?” という問いかけについて “本格量産になるまで時間を要したオーガニック基板上の Flip Chip (FC) の歴史と重ねてみるとそこにヒントが存在する” と説明した。今回は Fan Out - Wafer Level Package (FO-WLP) は本格的な普及に至るのか? という問いかけも追加してみたい。

また, 5G モバイル, 自動運転, 1.6TbE ネットワークさらにクラウド AI などこれからビジネスを牽引するインフラやアプリケーションについて要求される半導体性能を満たすために相応しいパッケージテクノロジーのロードマップが正しく描かれているのか検証し, 今から 10 年後に向けてのパッケージの課題を共有したい。

1992 年に世に登場したオーガニック基板上の Flip Chip - Ball Grid Array (FC-BGA) は²⁾, 1998 年以後に PC に採用され, ゲーム用などマーケットを徐々に拡大した。そして 2007 年に登場し現代の世の中に大きな変貌をもたらしたスマホのアプリケーションプロセッサ (AP) が Flip Chip - Chip Scale Package (FC-CSP) を使用するという展開により本格的なフリップチップ時代となった³⁾。その普及には 15 年が費やされたことになる。FO-WLP は 5 mm 前後のパッケージサイズで 2009 年ごろより採用が開始されていた⁴⁾。その後 2016 年に主要なスマホの Application Processor (AP) に採用され注目度が一気に上がった⁵⁾。しかしウエハレベルでの製造を Wafer Level Package (WLP) と同じ Re-Distribution-Layer (RDL) により回路形成するために FC-CSP 基板より単位面積当たりの製造コストが高かったことで FC-CSP パッケージの倍以上のコストとなり, その他の AP へは採用されていない。低コスト化のために FO-WLP を基板や液晶製造並みの大きさで作成しコストダウンを狙う Fan Out Panel Level Package (FO-PLP) の開発は進んでいる。製造フォーマットが大きくなることで RDL のデザインルールの微細化は FO-WLP に比べて難しいなど, 課題を多くあるが, 成功

すれば適用マーケットの爆発的な拡大になる。数社が 2018 年後半から少量量産を開始し, 2021 年以後には現在開発中のテクノロジーのほとんどの量産適用が期待される。2009 年を FO-WLP 元年として, フリップチップの例を当てはめれば 15 年後は 2024 年前後になり, その頃が本格的量産時期になる。

Through Silicon Via (TSV) の展開を振り返ってみる。CMOS Image Sensor (CIS) が比較的大きなマーケットで初めて適用されたのが 2008 年になる⁶⁾。High Band Memory (HBM) は 2015 年よりデスクトップパソコンへのグラフィックスカード (Add In Board : AIB) 適用開始された⁷⁾。2016 年には次の世代の HBM2 の発表があり⁸⁾, 2017 年に出荷が始まった。さらに 2020 年前後より次世代ゲームコンソールにリードされ次世代ゲームソフトは現在の 2 倍以上のメモリバンド幅を使用し, 2023 年ごろには本格化するだろう。また 2023 年前後の 5G のミリ波導入⁹⁾ による Virtual Reality (VR) や Augmented Reality (AR) のマーケット拡大が期待できる。これらは HBM の大幅な需要拡大につながる。Double-Data-Rate (DDR) 4 メモリーは, 2014 年よりメモリ最大手はフリップチップ化し, サーバー用の積層構造には TSV 構造とした¹⁰⁾。2019 年後半から 2020 年に登場する予定の DDR5 は全てのメモリメーカーがフリップチップと TSV の適用を開始する見込みである。DDR の世代は 4 年から 5 年の製品開発周期としているため, 2023 年あたりは TSV 化もかなり進むと推定する。2019 年に適用が予定されているモバイル用メモリーの次世代の Low Power DDR (LPDDR) はワイヤーボンドを踏襲することが明らかになっている。しかし同世代の DDR5 が全面フリップチップ化することから早ければ次世代, すくなくとも次々世代 (モバイル用メモリーは 2 年ごとにデータレート上げる製品サイクル) の 2023 年までには FC 化が予測される。LPDDR は 32 bits を 2 段積層して 64 bits に拡張して使うので, フリップされると半数のウエハに TSV が必要となる。これらを考えると遅くとも 2023 年ごろは TSV の本格的な立上りと推定され, CIS 適用後 15 年で本格的な適用がみられることになる。

図1. 5G時代を担うアプリケーション

つまりFO-WLPやTSVの普及は過去の歴史から見て、けっして遅くはなく、逆に少し早いぐらいでの展開のようにみえる。そう理解することで半導体パッケージのロードマップやトレンドについての見方が変わる。

図1に5G時代時代の示す対象アプリケーションに必要なパッケージのロードマップをみる。2年後の2020年から始まる5Gモバイル(5G)⁸⁾運用に伴いモバイルトラフィックの急増に対応するミリ波通信、自動運転向けなどの低遅延を実現するサブ6GHz通信、さらにIoTビッグデータ用大量センサー接続を実現するサブ1GHzなど、さまざまなアプリケーションへの5G適用の期待は大きく、さまざまな形態の基地局およびそれにつながるクライアント機器の通信デバイスが重要となる。

サーバーはクラウド化が進み、5Gで吸い上げられたビッグデータを活用するAI、仮想現実(VR)や拡張現実(AR)、4K以上の動画配信など要求されるデータ処理の急増に対応する必要がある。大量データトラフィックに応えるためにEthernetも2018年より100GbEから400GbEへアップグレードする¹¹⁾。さらに2021年前後に800GbE、2024年以後に1.6TbEとデータレートを上げる必要がある¹¹⁾。

スマホはミリ波のアンテナおよび通信デバイスを複数個装備し、大容量データとそれを活用するAIを含めたアプリケーションのためにさらにAPの性能向上が求められる。

ビッグデータ収集用やインフラモニター用センサーは、アンプ、データ処理用マイコンさらに、5G用通信デバイスを小型一体化する必要がある¹²⁾。

自動および自律運転においては、ADASの延長としてミリ波による物体認識レーダー、レーザーによる3D物体認識(Lidar)、さらにカメラとしてCMOSイメージセンサー(CIS)を中心に30を超えるセンサーおよびそのデータ処理デバイスの性能要求と搭載サイズの実現、センサーからの

状況を認識して運転判断するAI用制御デバイスにはPCを超えた性能を車載での実現という大きな課題、さらにそれぞれのセンサーとAIを遅延なく接続する車規格Ethernet接続が求められている¹³⁾。

5Gにより従来の1GHz以下の帯域により基地局にIoTの膨大な数のセンサーを直接接続することも可能になり、従来のWiFiベースのさまざまなLow Power Wide Area (LPWA)と役割分担しながら真のIoT時代を実現する。

以下の章においては、これらの主要なアプリケーションに対して要求される性能(検出、処理、さらに通信性能など)を実現するための半導体デバイスのパッケージのロードマップを確認し、それらの実現に対する課題を示す。

さらに、その中でもすべてのアプリケーションにとって今後期待値の高いFO-WLPおよび2.1/2.5Dパッケージに注目して、現在の量産テクノロジーとそのロードマップ、さらにはそれをコストパフォーマンスで上回る開発中のテクノロジーの可能性について着目して、テクノロジーの比較することで、今後のそれぞれのテクノロジーの役割と存在価値、さらにはデファクト化をけん引する可能性について評価する。

サーバー

クラウドデータセンター(クラウド)へのデータトラフィックはCAGR 28%(2016年から2021年)で増加し2016年から2021年までに約3.5倍の性能向上が必要になる見込み。クラウドのデータストレージはCAGR28%(同)で増加し、その中でもビッグデータについてはCAGR51%(同)で急速に成長する。このビッグデータの急増は人口知能(AI)アプリケーションの増加による¹⁴⁾。

人工知能(AI)アプリケーションにおいては画像認識、音声認識あるいは文章認識に対してDeep Learning (Deep

Learning: DL) が活用されそのアルゴリズムも日々進化している。AI アプリケーションに対するハードウェアおよびソフトウェアの環境整備がクラウドにおいて重要となっている。大手クラウドは DL 用に最適化されたサーバーの導入を進めている。

High Performance Computing (HPC) に対して General Purpose Graphics Processing Unit (GPGPU) を適用することで期待される効果が大きい。計算高速化を加速する Acceleration Processing Unit (APU) として 2012 年前後から導入が始まり、2018 年にはゲーム用の Graphics Processing Unit (GPU) を APU として適用したパソコンがトップ性能を 2 年連続で示した¹⁵⁾。その後 GPU を GPGPU として拡張して使用されるようになった。その後、DL に対して GPGPU 適用の効果が示されて以後 GPGPU は DL 実施のための重要デバイスとなった¹⁶⁾。

DL 実施の牽引をしている GPGPU はグラフィック描画用に開発された GPU をグラフィックス以外のアプリケーションに適用することからその名前の由来があり、GPU と GPGPU は世代ごとの GPU コアをそれぞれ共有するロードマップとなっている。DL のアルゴリズムを Application Specific Integrated Circuit (ASIC) あるいは Field Programmable Gate Array (FPGA) で実行すれば GPGPU より高速な演算が可能となる。DL による学習の成果の出ている DL アルゴリズムを ASIC 化し商用適用する製品は AI スピーカーなど増えている。クラウドにおいても特定アルゴリズムに対して ASIC 化の動きがみられる。さらに DL のアルゴリズム変更への対応も可能な FPGA は書込む時間がかかるものの GPGPU と比較して高速実施ができるメリットは大きい。FPGA 各社が DL アルゴリズムの書込みに対してのソフトウェア環境の整備が進みさらに活用が増えている。それぞれの方法に対して大量のビッグデータをアルゴリズムに基づく計算を行う上で使用するメモリバンド幅の拡張要求は共通であるが、ソフトウェアの依存度が高い GPGPU はより大きなバンド幅が必要となる。

GPU に対してはメモリバンド幅を拡張するために最大で 512 bits (32 bits/個) のバス幅を持つ Graphics Double Data Rate (GDDR) メモリが適用されてきた。GPGPU として GPU が HPC に使用されるようになるとメモリバンド幅がさらに必要とされ、マイクロンが中心となり 2013 年に Hybrid Band Memory Cube (HMC) 1.0¹⁷⁾ が 2014 年に JEDEC により High Band Memory (HBM) の仕様が制定された¹⁸⁾。

HMC は高速シリアルインターフェース SERializer/DESerializer (SerDdes) に最大 16 接続 (links) とした。SerDes なので信号線は双方向で 2 本で最大 16 リンクとなり 32 本の信号線で成立つために従来の基板で配線が可能となった。最大 8 モジュール (Cube) 構成として 320 GB/s のデータバンド幅を達成した。HMC 2.0 でバンド幅を 2 倍として SerDes のデータレートは 30 Gbps になった¹⁹⁾。さらに次世

図2. GPGPU と HBM 搭載 2.5D パッケージ

代としてバンド幅を2倍に上げるためにはSerDesを60 Gbpsとする必要があるが、マザーボード上のほぼ限界のSerDes性能になる(次章のネットワークの高速化で背景を説明する)。業界ではすでにHBMが散在することもあり、今までHMCを使用してきたGPGPUもHBMへの切替えが進む見込みである。

図2にGPGPUとHBMの典型的なシリコンインターポザーによる2.5D構造を示す。HBMのシリコンインターポザーの接続端子デザインを示す。HBMの接続端子は25 μmパッドサイズで、55 μmピッチ、48/96 μmスタガードで24列の引出しを2層の2/2 μmのL/Sルールとした。

図3にHBMの断面構造を示す³²⁾。HBMのメモリダイは2段(HBM2以後)、4段、8段のオプションがある。HBMは40 μmピッチのTSV接続がなされる。今後メモリの微細化によるダイ面積の縮小によりTSV部の面積も合わせて縮小するので、HBM2の後半ではTSVピッチは30 μm世代、さらにHBM3の後半では20 μmピッチまで採用されるロードマップになる。

GPGPUがHBMを採用する前に、デスクトップPC用の高性能GPU用のさらなるバンド幅を求めてHBM適用製品が2015年に商品化された。GPUモジュールとGDDR5x16の実装面積90 x 110 mm。GPGPUとHBMx4ではそれを55 x 55 mmのSiP上に実現可能で、バンド幅も204.8 GB/sから512 GBpsへ拡張できた⁷⁾。翌年にはHBM2としてそのバンド幅を2倍とした製品の仕様が示され、2017年より量産化された⁸⁾。HBMが出てからHBM2としてサーバー用としては異例の速度で世代交代した背景には、HBM X4モジュールのバンド幅がGDDR5をさらに最大で1.75倍ほどバンド幅を挙げたGDDR5Xx16によりバンド幅の差はほぼ同等となりコストパフォーマンスとして見劣りするよ

図3. HBM 構造³³⁾

うになったことがある。

HBM適用のHPCも今後次第に増える見込みである。2018年6月のスパコンのトップ500ランキングでは、アメリカ製スパコンがGPGPUとHBMのAPUを採用し性能のトップに立ち²⁰⁾、2年間続いた中国製スパコンの性能をやっと上回った。日本でもスパコン大手がコアをARMとしたMPUを開発中でGPGPUとの組み合わせで東京オリンピックの開催される2020年にスパコントップの座を狙っている²¹⁾。HBMを扱うパッケージのロードマップについて後の章で解説する。

ネットワーク

データセンターのイーサネットにおいては100Gbits Ethernet (100GbE) から400GbEへの移行が始まろうとしている¹¹⁾。400GbEは2017年末に仕様がリリースされた。400GbEのルーター(Router)やスイッチ(SW)のASICはすでにリーダー企業からは16nmテクノロジーノードで製品発表はしているが、本格的な導入は7nmテクノロジーノード適用製品から始まり、実際にマーケットに導入されるのは2019年からになる見込みである。先に述べたAIなどのアプリケーションの急速な伸びによりクラウドへのデータトラフィックは急増しているため、100GbEでは性能不足となり200GbEさらに400GbEへの移行が進み、2022年にはクラウド内においては50%以上がアップグレードする。

Ethernetは基地局やWiFiからのデータトラフィックをクラウドへつなげる際のボトルネックにならないようにする必要がある。ネットワーク性能はサーバー演算能力を中心に、データトラフィックとのバランスで3-4年で2倍、6から7年で4倍にアップグレードしてきた。現在の100GbEの仕様が制定されたのは2010年になる¹¹⁾。これは約4年でDDRメモリのバンド幅が2倍になりメモリチャネル(Ch)の平均増加もこれに追加される。EthernetからMPUへのデータ転送を行うPCI express (PCIe)のバンド幅が3年ごと

に倍になるので、それぞれ6-7年で性能が4倍になることに対応する結果である。

400GbEが動きしたので、3年後に向けての800GbE実現に向けて動きだしている。その800GbEは7年後の2024年の1.6TbEへの橋渡しとして重要な役割を担うのは当然である。

100GbEでは25 Gbps/レーンで4レーンを1チャネル(ポート)としてデータセンター向けのSWは32ポート仕様が一般的となっている。400GbEへのアップグレードに際しては、4レーン/チャネル実現のためにはレーンあたりの100 Gbpsのデータレートを実現する必要がある、マザーボード配線において各層をつなぐPTHによる反射ノイズが大きな理由から100 Gbpsの実現ができなかった。さらに結果56GbEのSerDdesは12.8 GHzのPulse-Amplitude Modulation 4 (PAM4)が採用され、50 Gbps×8レーンで400GbEとなった²²⁾。

図4に示すように、SWは28GbEを4レーンとしていたASICから同じポート数実現するためには、倍のSerdes接続が必要となり、FC-BGAのサイズは70 mmまで大きくなる。それに伴い、Serdesの配線長は長くなり、さらなる低Dfのビルドアップ材料が必要となり、100GbE世代では0.007-0.009だったDfが400GbEでは0.004台のDfの材料となった。マザーボードでは周波数を上げずPAM4採用となったが、従来の30 cm前後の配線長では信号ノイズが大きく、光信号を電気信号に変換するフォトニクスをSWの近くに搭載し、ボードの光リンクからさらにボード内で光接続を延長させるOn Board Optics (OBO)の採用が始まる。マザーボードの材料、PTHサイズ変更などコストアップの方法か、OBOの採用かはSW装置開発各社に委ねられる。

図5に示すように800GbE実現に対しては、400GbEの8レーン/ポートを16レーン/ポートにする方法、100GbE実現のために14 Gbpsを8倍にして112 Gbpsを実現するPAM8を用いる方法が、検討されている。さらにその倍に対して

図 4. Ethernet パッケージトレンド

図 5. SW ASIC 用パッケージロードマップ

はフォトニクスを SW に直接搭載し、信号はマザーボードを経由しない方法が検討される。フォトニクスから SW の間の信号についてはパッケージ上で絶縁層を薄くすることで Serdes の信号ピッチを下げ密度を上げるなどで、PAM4 で 56 Gbps x 32 lane あるいは PAM8 により 100 Gbps x 16 などのオプションが検討される。800GbE 世代に対してはその基本技術となる PAM8 などの確立ニーズがある。さらにクラウドからはオンパッケージのフォトニクスを 800GbE からの実現ニーズも高い。1.6TbE におけるオンボードフォトニクスのイメージは図 5 のようになる。ベースとなるサブストレート材料はオーガニック、ガラスあるいはシリコンをベースにした構造の開発が行われているが、今後構造以外に材料、フォトニクスへの光接続の方法など課題は多い。また 1 つのフォトニクスとしては 400 Gbps が最高性能で²³⁾でこれを 1.6Tbps に引上げる必要があり、当面は 800 Gbps の出現が待たれる。

モバイル

スマホの AP の進化は AI アプリの増加につれ性能要求を増す。必然としてコア数が増え、LPDDR のデータバンド幅

が増えるスマホのハイエンド大手は AP に FO-WLP を適用し従来のフリップチップ構造にくらべて電源ノイズで 50% 以上、熱抵抗で 20% 以上の改善を示し、今後の高性能化に大きく貢献するテクノロジーとなった。この詳細は後章にて説明する。図 6 に AP のロードマップを示す。ハイエンドパッケージは FO-WLP を前提としている。

2016 年にメモリのバンド幅は 25.6 GB/s、2017 年から 2018 年モデルではその 1.3 倍の 34.1 GB/s となり、いよいよ 2019 年より 51.2 GB/s を達成する LPDDR5 の実現が期待される。LPDDR5 の次世代については、データレートをさらに従来の 2 倍にすることは簡単ではない。バンド幅を倍増させるために 2CH から 4CH として 100 GB/s を超える事は検討されるであろう。そしてさらなるバンド幅の増加については、やはりバス幅を増加する方法となり、過去 2 回ほど議論された WideIO は将来の候補として残るだろう。それに伴う、AP の IO 接続デザインルールは 64 bits で 10/10 μm 平均だった L/S が 128 bits では 5/5 m, 512 bits では 2/2 μm が必要となる。さらに 5n あるいは 3n へ移行するタイミングで AP のダイを分割し 2 ダイ構成とするダイパーティショニングが行われる。その際にはダイ間の接続に 2/2 μm

All numbers below: Estimated by SBR Technology Co., Ltd

	Production year							
	2016	2017	2018	2019	2020	2021	2023	2025
SoC technology node	16n+	10n	7n	7n	7n+ EUV	7n+ EUV -> (5n) -> Die partitioning		
Memory	LPDDR4	LPDDR4 X(*1)		LPDDR5 (LPDDR5 Wide)		LPDDR5 Wide (Wide IOx)		Wide IOx
Band width	25.6GB/s	51.2GB/s 34.1GB/s		(102.4GB/s) 51.2GB/s		(204.8GB/s) 102.4GB/s		(409.6GB/s) 204.8GB/s
RDL bottom layer counts	3	3		3		3		4
RDL bottom L/S (um)	10/10	5/5	4/4	3/3	2.5/2.5	2/2		1.5/1.5
RDL top layer counts	0	0		1		2		2
RDL top L/S (um)	NA	NA		5/5		3/3		2/2

(*1): LPDDR4X: 4266Mbps, VDDQ/VDD = 0.6/1.1V
IO power reduction with 0.6V VDDQ -> IO power was 45% of LP4
34.1GB/s (64 bits) @ 4266Mbps (16 bit) x 4 ch

図 6. スマホ AP 性能およびパッケージロードマップ

L/S が要求される。

5G 時代のスマホの RF Front End (RFFE) について、4G, サブ 6 GHz, ミリ波に対するブロックダイアグラムを図 7 に示す。4G までの RFFE は SAW フィルターあるいは BAW フィルターを SW で切替え、微弱な受信波を AMP で増幅してトランシーバーからベースバンドにつなげる。サブ 6 GHz においては 4G の RFFE のどれかにその周波数を追加し、SW や AMP を場合によっては 2 個搭載するなどより複雑な SiP として対応する。ミリ波についてはトランシーバーは RFFE 側に持つことになるので 28 GHz の RFFE とトランシーバーの SoC 化が想定される。ユーザの手の干渉などを考慮するとスマホの両端に 2 個、さらに 60 GHz を使用する次世代の WiFi の WiGig を扱うアンテナ SiP が 1 個の合計 3 個のミリ波アンテナパッケージが要求される。アンテナと SoC とを一体にする Antenna in Package (AiP) とアンテナを低損失基板や FPC とし SoC パッケージを別に搭載するセパレート構造が想定される。28 GHz についてはパッケージサイズが比較的大きく AiP に対しては 10-12 mm が想定されるので、セパレート構造も強みとなるが、アンテナからのミリ波信号を SoC へパッケージ外部接続を通して行うためにその信号ノイズが懸念される。60 GHz については 7 mm 程度の AiP となるので一体型になると思われる。

その他スマホ内の Power Management IC (PMIC), Audio Coder and Decoder (CODEC) などは FO-WLP へ移行し、認証やクレジット機能やセンサー類もさらに種類が増えると共に場所制約から SiP への流れがあり、それを FO-WLP の実現に期待がかかる。

さまざまなアプリケーションで期待される FO-WLP のロードマップについては後の章で述べる。

図 7. 4G, 5G RF FE ブロックダイアグラムおよびパッケージ要求

基地局

次にバックホールにおけるパッケージテクノロジーを見てみる。LTE-A から 5G へ向けて、高周波帯域の電波を使う事で、電波到達距離が短くなりフェムトセルやピコセルと呼ばれるスモールセル（基地局）の設置が進む。図 8 に示すように、これら基地局で重要となるのが Massive Multiple Input Multiple Output アンテナ（数十から数百に及

図 8. 基地局からのミリ波 MIMO ビームフォーミング

ぶアンテナ数による MIMO) 技術で、ビームフォーミングにより対象とするモバイル端末に焦点を定めて通信を行う⁹⁾。同じ電波周波数で従来の 10 倍以上の距離での通信を可能にすることで、基地局の設置負担を最小化する。実現するパッケージは図 9 に示すように表層に多くのパッチアンテナを配置し、ビームフォーマと呼ばれるデバイスをアンテナインパッケージ (AiP) 構造で実現する²⁴⁾。AiP の材料はプリント配線板であればテフロンや LCP などの低損失材料でかつデバイスの高耐熱でかつ温度および湿度による性能変化が少ない材料が要求される。基板のガラス転移温度と同じかそれ以上の発熱になる高出力の Gallium Nitride (GaN) には高温に耐えられるセラミック基板が想定される。

ミリ波の特性から基地局から可視的に見える範囲にしか電波は届かない。基地局から影になる場所や室内に対しては、その受信した電波を WiFi で受けて、クライアントデバイスとつながる。WiFi についてはすでに 7 Gbps 性能を 60 GHz ミリ波で実現する WiGig が導入済みで²⁵⁾、ミリ波受信を WiGig あるいは従来 WiFi に接続するフェムトセルの数が現在の 4G 世代から大幅に増える。WiGig 対応フェムトセルは 5G の 28 GHz と WiFi の 60 GHz の両方の RFFE と Base band を持つ。アンテナはスマホと同様に SiP あるいは SoC を含む AiP であるが、スマホほど大きさの制約が無いので、AiP のアンテナ部分の送受信感度を上げるための十分な大きさ、受信の指向性から複数の配置などスマホよりは自由度がある。しかし設置要場所に合わせた外形や寸法など依然、AiP はコンパクトであることが求められる。AiP 実現に対しては、FC-CSP、FO-WLP あるいは部品内蔵パッケージによる AiP となり、今後 5G や WiGig のデータレートなどの性能がアップグレードされ、それらのパッケージ

図 9. 基地局用ビームフォーミングアンテナモジュールパッケージ

の主に要求材料のロードマップが重要になる。ミリ波はビルなど建物により遮られるために、スモールセルから WiGig など WiFi につなぐフェムトセル設置が急増する。WiGig の 60 GHz も建物内の壁を通せない事を考えると従来の 5 GHz WiFi との組み合わせとなる。フェムトセルにおいてもビームフォーミングを適用するので、ミリ波スモールセルの低出力版に WiFi を加えた構造になる。それらのパッケージについてもアンテナを搭載したビームフォーマーとそれを制御し状況に応じて AP, ASIC や FPGA により多様化に対応する。パッケージの多くは FC-BGA となる。

IoT センサー

図 10 に示すように、ビッグデータ収集用のさまざまなセンサーは 5G 世代ではセンサーから直接基地局に長距離データ転送し、かつ消費電力を抑えた Low Power Wide Area (LPWA)²⁶⁾ が中心となる。そのパッケージ実現はセンサー、アンプ、マイコンに加えて RFFE の構成となり、その電源

2016	2021	2026
FO-WLP 12インチWafer製造	500mmサイズ以上のパネル製造	パネル製造にて3D構造
PMIC, Audio CODEC, 77GHzレーダーなど	LPDDR、NANDメモリーなどFCCSP化が必要な周波数、バンド幅に対しての適用	3D構造を前提とするIoTセンサーモジュール、RFアンテナデバイスなど
5mmサイズ前後でFBGAよりコストアップ、10mmサイズ前後でFCCSPよりコストアップ	8mmサイズ前後までのFBGA、15mm前後までのFCBGAのアプリケーションをカバーする	3D構造を前提としたSiPに対して部品内蔵パッケージ以上にコスト競争力を持つ

図 10. IoT センサー統合のためのパッケージロードマップ

を太陽光に依存する場合にはさらなるデバイスの追加が必要のために複雑な SiP になる。

SiP 実現として理想的には、FO-WLP と部品内蔵パッケージが候補に挙がるがそれぞれ適用にはコストの壁が存在する。低コスト化にはできるだけ大きなフォームファクターでの製造が必要であり、すでに基板プロセスをベースにした部品内蔵のさらなる量産化が期待される。

部品内蔵基板パッケージは基板製造に部品搭載を組込むプロセスとして開発された。基板のデザインルールや部品搭載の位置精度からベアチップを搭載することは想定しておらず、SoC などアクティブ IC は WLP 化して搭載としていた。また、製造が基板メーカーなのでパッケージ後のテストインフラが不十分でもあり、パッケージとして対応するには不十分だった。近年基板メーカーから OSATS への展開が図られつつありその拡大に期待がかかる。FO-WLP については後の章で解説する FO-PLP に期待が集まる。

自動車

自動車の今後は夢の自律走行で、長距離ドライブにおける社内でのエンターテインメントな過ごし方、車がオフィスになるなど期待が高まる。ドライバーが責任持つものの自動車の AI が運転制御を行う自律運転の実用化（レベル 3 自動運転）²⁷⁾ の実用化が 2019 年より実用化される国もあり 2025 年までには先進主要国では限定的ながら運用が始まる見込みである。

表 1 に自律運転を支える AI システム構成とその中心になる半導体のパッケージテクノロジートレンドを示す。周囲の状況把握のセンサー類とその状況から目的に沿った運転指示を与える AI コントローラ、それを受け実際の走行制御を行う膨大なマイクロコンピュータ (MC) を統括する制御ユニットにより運転が実施される。

現在自律運転用 AI (AI) はクラウドで膨大な運転に関連するビッグデータで予備学習され、さらにその AI に実走

表 1. 自動運転（自律運転）用デバイス

デバイス	機能	要求半導体	要求パッケージ
AI コントローラ	自動（自律）運転	MPU-GPGPU MPU-ASIC MPU-FPGA	FC-BGA
全体指示 コントローラ	全体制御	MPU - ASIC	FC-BGA
LRR Rader	200 - 100 m 物体認識	SiGe	FO-WLP
MRR/SRR Rader	100 m 以内 物体認識	RF-CMOS	FO-WLP/ FC-CSP
Lidar	3D 地形、物体認識	Photonics laser	SiP
CIS/Camera	物体、標識、道路認識	CIS	imBGA など ³⁵⁾

行を経験させて追加学習を行う。この実走行追加学習用のシステムは AI デバイス開発各社ともサーバークラスの複数 MPU に複数 GPGPU や ASIC/FPGA などクラウドシステムを全体で約 500 W と車載ギリギリまでダウンサイズしたシステム構成²⁸⁾ となっている。

商用実現時には AI 用に 30 W クラス²⁸⁾ までの SoC が中心となり、この例では 8CPU コア、512 の GPU コアをワンチップに実現している。ソフトウェアに加えて ASIC による AI アルゴリズム実施を行う事も想定される。これらはすべて FC-BGA であり、信頼性に対しての要求は民生用より高いが、マルチ構成されている CPU や GPU については、冗長構成を前提としてのシステム構成でありフェイルセーフに対応させている。メモリについては LPDDR を最大 8ch まで接続する。

走行に必要な人間の目に匹敵する情報はミリ波レーダー (Radar)、赤外線レーザーレーダー (Lidar)、イメージセンサー (CIS) さらに超音波センサーなどのセンサーにより入力される。以下センサーと要求される仕様を示す。

Rader については、100 - 200 m 先までを見るレーダーを長距離レーダー (LRR) と呼び 77 GHz の周波数で帯域は 1 GHz 以下、100 m までの距離に対してはその出力により 50 - 100 m の中距離 (MRR) あるいはそれ以下の短距離レー

ダー (SRR) に区別される。LRR については、アンテナからトランシーバへのミリ波入出力損失の小さな FO-WLP が適用され、自動車搭載の信頼性基準を満たす。MRR や SRR 用には 24 GHz が想定されていたが、79 GHz の規格制定後は距離検出力に優れているという理由から今後の主流となる見込み。

Lidar については、赤外線レーザーの到達距離の限界から 100 m 以下を対象にしている。3D 物体認識について、昼夜間問わず可能なことからセンサーの中心となる。レーザーの受発信システムを回転スキャンするタイプはカバレッジが高いが装置が効果で車の屋根に突起物として搭載する必要があり、タクシーロボットなど 2020 年前後での実用化を狙っているアプリケーションでは採用されるが、一般車に対しては、レーザーをフェーズドアレイ方式でビームフォーミングを行う方法の導入が始まろうとしている。さらに MEMS の活用などでサイズ、コストダウンへの開発競争となる。

CIS は標識や信号などの認識に必須である。反射光、トンネル内外などの明暗変化さらに夕暮れなどの暗視野への対応など Radar や Lidar など同時に搭載することができる技術だけが、カバレッジを広げることが求められる。

さらに発展途上国においては道路インフラなど変化が大きい上に工事も多いので、正確な 3D 地図情報の頻繁なアップデートとそれを 5G により車のデータベースのアップデー

トを行う必要がある。この場合の 5G の接続にはサブ 6 GHz を主体として、ミリ波も使用可能な場所で使用される。さらにミリ波は動く監視カメラの位置づけからトラック、バス、緊急車両あるいは対象車両の走行モニターなどサブ 6 GHz との併用することでのメリットは高い。車用には屋根の上にコンパクトに取付ける必要から AiP 構造で搭載マザーボード含めてコンパクトなシステム実装が必要とされる。

センサーから AI へ、その制御指示を実現する制御デバイスについての相互接続は車載規格のイーサネット (Ethernet) となろうと最終な規格検討がされていて、2018 年内には規格制定される。多くのセンサーについては光接続が必要な 1 Gbps を超えないレベルで、CIS からの出力をそのままモニター用に 5G 接続するなどの用途については光接続が必要になるかもしれない。Ethernet のパッケージはポート数も少ないので QFN が中心となる。

FO-WLP パッケージ

FO-WLP は Re Distribution Layer (RDL) より先にダイを搭載するプロセスをダイファースト、RDL が先でダイをフリップチップ接続するプロセスをダイラストと呼ぶ。前者の FO-WLP は 2009 年ごろより本格量産が開始され、2015 年には 200 M 個以上の出荷数 (US\$20 M 以上の売) を OSATS が中心となり実現してきた³⁴⁾。図 11 に示すように、

図 11. APU 用従来パッケージと FO-WLP の比較⁵⁾

図 12. FO-WLP/PLP と従来タイプパッケージのコストトレンド

2016 年の後半ハイエンドスマホのリーディングカンパニーが AP 用に FO-WLP を採用した結果、熱そりが大幅に少なくなり FC-CSP より約 30% の低背化が実現でき、さらに電源ノイズを半分以下になった。さらに熱抵抗も約 20% 下がるなどスマホ用の AP が求めていた要求を満足させる高性能パッケージとしての実力を示した²⁹⁾。

しかし、2017 年同じスマホの次期モデルへ継続適用されているが、その他の AP には適用されていない。その理由は、コストにある。図 12 に構造に対するコスト比較を示す。FO-WLP とフリップチップサブストレートを比較した場合、絶縁膜の厚みがそれぞれ $7\mu\text{m}$ 前後対 $30\mu\text{m}$ 前後で膜厚精度の差、めっきのシード層に対してそれぞれスパッタリングと無電解の差、めっき厚みが $5\mu\text{m}$ と $15\mu\text{m}$ の差、さらに大きな違いにつながる回路形成が 300 mm の Wafer と $500 \times 500\text{ mm}$ 前後のパネルの違いがあるため、プロセスが高価になる上に製造サイズに差がある。図 12 に示すように 12 インチ製造の FO-WLP では約 10 mm サイズ以上の FC-CSP、約 5 mm 以上の FBGA のコストに及ばない。価格を下げる方法として製造サイズを大きくすることが有効な方法になる。仮に $500 \times 500\text{ mm}$ の製造プロセスを確立したとすれば、約 15 mm サイズの FCCSP、約 7 mm サイズの FBGA までと競争できるコスト計算結果となっている。

FO-WLP を業界でリードする OSAT ではコストを下げるためのアプローチとして、製造サイズを大きくし、単位面積あたりの RDL コストを下げるロードマップを持つ³⁰⁾。この時点においては配線ルールは $5/5\mu\text{m}$ の L/S であったが、開発のタイミングが遅れ先の AP のロードマップで示したように FO-WLP が Wide IO クラスのメモリまで対応するとなるとファンアウトには部分的に $2/2\mu\text{m}$ レベ

ルの配線ルールも想定する必要がある、それをパネルレベルで実現できるかどうかの検討段階で、開発投資が停滞してしまった。しかし最近になってパネルレベル開発が進みつつある。5G 時代に向けての IoT センサー SiP や RFIC の AiP や SiP、自動運転用のセンサーなど期待値が高くなった背景がある。これら開発のロードマップとして $2/2\mu\text{m}$ の L/S が加わるができるかどうかにか今後の期待値はさらに膨らむ。

一方で FO-WLP は $2/2\mu\text{m}$ L/S 以後のロードマップを見据えて、ロードマップについて変化が起きている。

2.1/2.5D パッケージ

図 2 に示した GPGPU と HBM においてシリコンインターポーザのコストと電気特性の 2 点に課題がある。最近 GPGPU のコア数増大、あるいは HBM を 6 サイト搭載しバンド幅を拡大する動きなどが増えつつある。その場合にシリコンインターポーザのリソグラフィを 1 ショットで行う限界サイズを超えて、2 ショットさらには 4 ショットとする必要があるためシリコンインターポーザのコストがさらに高くなる。HBM2 のバス速度は 2 Gbps でシリコンインターポーザで Signal Integrity (SI) についての限界はまだ無い。しかし HBM4 (仮名) 世代としては 8 Gbps のバス速度となる。これは GDDR5 と同じバス速度となり、GDDR5 が 32 ビットであっても SI に対して厳しい状況などから、1,024 ビットの HBM にとってシリコンインターポーザでの限界に近いと思われる。いずれにせよ、シリコンインターポーザからオーガニック素材をベースにした FO-WLP あるいは 2.1D に切替える必要が出てくる Embedded Multi-die Interconnect Bridge (EMIB) と同じ状況で、シリコンブリッジの限

図 13. 2.1/2.5D 構造ロードマップと課題

界は見える。

図 13 にシリコンインターポーザに代わるパッケージ候補のロードマップとその課題をまとめた。

FO-WLP は $2/2\ \mu\text{m}$ の L/S 以下のファインピッチに対しては Non TSV Interconnection (NTI) としてシリコンキャリアに形成した SiO_2 による RDL に対してダイラストの FO-WLP プロセス後にシリコンキャリアを除去することで、BGA 搭載側に PI による RDL を追加するというテクノロジーでサブ $1\ \mu\text{m}$ の配線実現のロードマップを実現できるとしていた³¹⁾。しかし最近 FO-WLP を牽引するリーダー企業からの論文³²⁾では $0.8/0.8\ \mu\text{m}$ の L/S まで PI にダマシンプロセスで可能にできるとした。これによりサブ $2\ \mu\text{m}$ に対して FO-WLP のロードマップが見えてきた。

一方で 2.1D については依然製品化の動きが見られない。2014 年前後には $2/2\ \mu\text{m}$ の L/S 実現についての論文がいくつか発表されている³³⁾。その後それら各社から量産化についての発表はない。背景として、2.1D の量産対象とするアプリケーションが“逃げ水”のように遠ざかっていることにある。もちろん HBM を適用するアプリケーションは対象ではあるが、その数量は限定されている。テクノロジーノードの進化は永遠ではなく、10 年前には $7\ \text{nm}$ が最後の世代かもしれないと、ムーア則の終焉という言葉がささやかれた。そして進化できないテクノロジーノードに対して、必要な機能をそれに相応しいテクノロジーノードや CMOS 以外の化学化合物半導体の適用を行い、パッケージで 3D あるいは 2.5D 接続するヘテロジニアス SiP あるいはダイパーティショニング SiP がその役割を担うというシナリオがあった。これは前工程のテクノロジーノードの進化に対してコストが高くなり量産可能なテクノロジーノードのダイに必要な機能実現させてそれをパッケージで SiP とする方がコストパフォーマンスが高いという仮説に基づく。実際にこのシナリオとなるテクノロジーノードの限界

は $7\ \text{nm}$ なのかあるいは $5\ \text{nm}$ を超え、 $3\ \text{nm}$ まで SoC 時代は続くのか？。前工程の $7\ \text{nm}$ 世代の Extreme UltraViolet (EUV) など巨大な投資は今始まったばかりで、この議論の結論は差送り状態でそれを“逃げ水”と表現した。

まとめ

フリップチップの例から先端パッケージの本格的立上りは、少量量産から 15 年前後かかると仮定し、TSV、さらに FO-WLP について推定した。TSV の本格的な立上りは、DDR5, HBM4 さらに LPDDR5 の次世代の導入ラッシュの 2023 年ごろで、2008 年の CIS 適用から 15 年となる見込みである。この推論で FO-WLP の本格的な立上りは 2009 年の初期量産適用から本格的な量産まで 15 年の 2024 年になるとすると、まだ FO-PLP による適用マーケットの拡大には間に合うことになるので、現在の開発の成果に期待したい。

クラウドにおいての今後急速に AI 関連アプリケーションが拡大する。それにともない、GPGPU, ASIC, FPGA が HBM を使用する展開になる。現在そのためのパッケージとしてシリコンインターポーザによる 2.5D 構造であるが、コア数やロジック数の増加でシリコンインターポーザの面積が拡大し、リソグラフィのコストが上昇しており、パッケージのコストダウン要求が強い。さらに HBM のデータレートが $8\ \text{Gbps}$ を超える世代では、SI の限界になる。その意味からオーガニックインターポーザあるいはサブストレートに対する要望は高まっている。

ネットワークは $400\ \text{GbE}$ が開始され、すでに次世代 $800\ \text{GbE}$ さらには $1.6\ \text{TbE}$ のパッケージ仕様についての議論が盛上ってきた。 $400\ \text{GbE}$ の延長とするか、 $1.6\ \text{TbE}$ コンセプトの導入を $800\ \text{GbE}$ から行うのか、これらの議論より方向性は 2019 年前半には判明する見込みである。

5G 基地局のサブ $6\ \text{GHz}$ 対応のスマートセルについての

技術的な課題は4G世代の延長で考えられる。ミリ波適用になると、多くの課題を克服する必要がある。ビームフォーミング実現のためのSoCあるいはSiCの最適化、アンテナ構造とそれを実現するパッケージタイプの材料特性、さらに発熱に耐えられ基板材料などを2020年までに克服する必要がある。

5G世代のスマホにおける課題は、クラウドのAIをエッジで活用するための高性能APを消費電力を抑えて実現しなければならない。その意味でFO-WLPが最も省電力化を実現できるので、その普及が望まれるがコストアップの壁を超える必要がある。その実現にはFO-PLPの適用が必須で、その成功が鍵となる。

IoTのデバイスは5Gサブ1GHzのインフラ普及により従来から言われているトリリオンセンサー時代に突入する。そのパッケージとしては、小型化と高性能化と省電力化を普及可能なコストでの実現が必要なる。そのためのパッケージとしては部品内蔵パッケージあるいはFO-PLPが重要な役割を果たす。部品内蔵はOSATSの本格的な参入がみられ、基板の大型化により低コストの実現がされようとしている。この流れを他のOSATSの参入により加速したい。FO-PLPについては、先にすべてのアプリケーションが必要としており、フリップチップより低電源ノイズ、最少のSiPサイズ実現など多くの先進性の実現が待たれる。

以上、現状のパッケージテクノロジーの考察を通して、これからのパッケージへのロードマップの期待値の実現に向けての課題を挙げ、この業界に携わる企業の方々の指針にして頂けることを願う。

(2018.8.7-受理)

文 献

- 西尾俊彦, “IoT時代における2.1/2.5/3D実装マーケットの展望,” JIEP 2016年5月号
- Y. Tsukada, Y. Mashimoto, T. Nishio, and N. Mii: “Reliability and Stress Analysis of Encapsulated Flip Chip Joint of Epoxy Base Printed Circuit Board,” Proceedings of ASME/JSME Joint Conference on Electronic Packaging, Vol. 2, pp. 827–835 (1992)
- iPhone 2G 620 MHz ARM CPU: <https://www.engadget.com/2007/07/01/iphone-processor-found-620mhz-arm/>
- S. W. Yoon, Y. Lin, and P. C. Marimuthu: “Thin SiP and 3D eWLB (embedded Wafer Level BGA) Technology for Advanced Packaging,” ICEP 2012, Apr. 2012
- C. F. Tseng, C.-S. Liu, C.-H. Wu, and D. Yu: “InFO (Wafer Level Integrated Fan-Out) Technology,” 2016 IEEE 66th Electronic Components and Technology Conference (ECTC)
- 東芝プレスリリース: http://www.toshiba.co.jp/about/press/2007_10/pr_j0102.htm
- AMD press release: Jun/16/2015, <https://www.amd.com/en/press-releases/new-era-pc-gaming-2015jun16>
- Samsung press release: Jul/18/2017: <https://news.samsung.com/global/samsung-increases-production-of-industrys-fastest-dram-8gb-hbm2-to-address-rapidly-growing-market-demand>
- DOCOMO 5G モバイルホワイトペーパー: https://www.nttdocomo.co.jp/binary/pdf/corporate/technology/whitepaper_5g/DOCOMO_5G_White_PaperJP_20141006.pdf
- 半導体プロセスまるわかり 微細化よりも高密度化に進むDRAMの未来, ASCII デジタル, Jun/2014, <http://ascii.jp/elem/000/000/906/906781/img.html>
- OFC Ethernet Road Map: <http://www.ethernetalliance.org/roadmap/>
- T. Nishio: 43rd Internecon Technical Conference, Jan'14, 2014
- Ixia Automotive Ethernet White Paper: https://support.ixiacom.com/sites/default/files/resources/whitepaper/ixia-automotive-ethernet-primer-whitepaper_1.pdf
- Cisco Global Cloud Index: Forecast and Methodology, 2016–2021 (Cisco white paper): <https://www.cisco.com/c/en/us/solutions/collateral/service-provider/global-cloud-index-gci/white-paper-c11-738085.html>
- Roadrunner ウィキペディア: <https://ja.wikipedia.org/wiki/Roadrunner>
- GPGPU ウィキペディア: <https://ja.wikipedia.org/wiki/GPGPU>
- Hybrid Memory Cube Specification 1.0: 2013
- JEDEC Spec: High Bandwidth Memory (HBM) DRAM, JESD235, 2013
- Hybrid Memory Cube Specification 2.0: 2014
- Top500 statistics: <https://www.top500.org/statistics/>
- 富士通: “ポスト「京」開発状況について,” 2018年6月21日, <http://pr.fujitsu.com/jp/news/2018/06/21.html>
- S. Bhoja and F. Chang: “PAM Modulation for 400G SMF,” IEEE P802.3bs 400 Gb/s Ethernet Task Force, May 2014
- 日経エレクトロニクス: “PETRAと富士通, MPUに載る400Gbpsの光送受信器,” 2017年9月19日 <https://tech.nikkeibp.co.jp/dm/atcl/news/16/091409175/>
- IBM: “Enabling 5G: mmWave Silicon Integration and Packaging,” IEEE 5G Summit, Nov. 2015
- WiGig ウィキペディア: <https://ja.wikipedia.org/wiki/WiGig>
- LPWA ウィキペディア: [https://ja.wikipedia.org/wiki/LPWA_\(%E7%84%A1%E7%B7%9A\)](https://ja.wikipedia.org/wiki/LPWA_(%E7%84%A1%E7%B7%9A))
- 内閣官房IT総合戦略室: “自動運転レベルの定義を巡る動きと今後の対応,” 2016年12月7日
- nVidia: “自動運転向けAI車載コンピューター,” <https://www.nvidia.com/ja-jp/self-driving-cars/drive-platform/>
- InFO (Wafer Level Integrated Fan-Out Technology): 2016 IEEE 66th Electronic Components and Technology Conference
- 西尾俊彦: “FO-WLPは次世代モバイル機器実装の主役になり得るか?,” 2017 JPCA 技術セミナー

- 31) S. Ramalingam: "HBM Package Integration: Technology Trends, Challenges and Applications," Hot Chips Aug 21, 2016
- 32) H.-P. Pu, H. J. Kuo, C. S. Liu, and D. C. H. Yu: A Novel Submicron Polymer Re-Distribution Layer Technology for Advanced InFO Packaging45, 2018 ECTC
- 33) S. Watanabe: "Development of Organic Multi Chip Package for High Performance Application," i-NEMI Substrate & Packaging Technology Workshop, April 22, 2014
- 34) B. Prior: "High Density Interconnects Migrating from Concept to Mainstream," Prismark, 3D ASIP 2017
- 35) J. L. Huang and KingPak: "Image Sensor Advanced Package

Solution," Semicon Taiwan 2013, <https://pdfs.semanticscholar.org/presentation/639f/1a52fc90e89847288966119b137412e61f27.pdf>, Prepared by : RD division

著者紹介

西尾俊彦（にしお としひこ）
1988年5月：日本IBM入社
野洲研究所にてオーガニック基板へのフリップ
チップ実装技術開発
1993年：先端実装技術アプリケーション開発
2003年：IBM Distinguished Engineer（技術理事）
2010年12月：STATSChipPAC Ltd. 入社
2013年1月：同日本法人代表
2014年11月：SBR Technologyとしてコンサル
ティングサービス