

〈研究・調査報告〉

グローバル AI 半導体企業の動向

吉 岡 美 愛

【要旨】

近年生成AI「ChatGPT」の出現によって、世界中で爆発的にAI半導体の需要が高まっている。その中でも米国の半導体大手NVIDIAが、AI半導体の80%を占めている。

NVIDIAのGPUが、生成AIの開発や利用によって世界中に広がった背景には、2006年から提供していたCUDAというソフトウェア基盤の存在がある。当初は画像処理を目的としたが汎用的な計算を可能にしたため、その後世界中のAI開発者のスタンダードバイブルとなっている。そしてNVIDIAの製品の特徴として、自社が設計、韓国のSKハイニックス及びサムスンが高性能メモリーHBMを供給、世界最大のファウンドリー企業である台湾のTSMCが製造し、最終的に日本企業の後工程管理によってパッケージングされる流れで作られている。

また世界各地で建設されるAI向けデータセンターは、高額なGPUと莫大な電力消費が最大の弱点であるため、米国のビッグテックは、AI処理に特化した省電力の独自半導体開発に動き出したが、モデルの性能は「ChatGPT」のようなレベルには達しておらず、テキストの要約や画像検索など日常生活の用途にとどまっている。このような米国のビッグテックの新たな動きは、今後独り勝ちのNVIDIAに、潜在的な脅威を与えるに違いない。

本論文では、グローバルAI半導体関連企業であるNVIDIA、TSMC、サムスンやSKハイニックスなどの戦略と動向について考察している。

キーワード: AI半導体、GPU、HBM、NVIDIA、TSMC

1. はじめに

まず2016年2月にマイナス金利政策が導入されて以来8年ぶりに、2024年3月日本銀行は、金融政策決定会合でマイナス金利政策を含む大規模緩和の解除を決定した。今回の大規模緩和解除により、マイナス0.1%だった政策金利を0~0.25%へ引き上げた。

そして同年3月27日外国為替市場で34年ぶりに円相場が、一時151円97銭と円安が進んだが、これはマイナス金利政策が解除されたことや円キャリートレードの加速、個人の海外投資が拡大したことが要因とされ、更に円安が加速すると見られる。

また今年11月の米国大統領選挙を控えて「もしもトランプから確実トランプへ」の選挙戦

が熱い。最初は「高齢者二人」の戦いからスタートしたものの、途中バイデン大統領からハリス現副大統領へと民主党の大統領候補が交代された。これは「反トランプ」の共和党支持層や白人以外のエスニックグループ、無党派層の票を集める狙いがあると言われている。

最近「米国第一主義」トランプ氏の諸発言を受けて、世界の株式市場に警戒が広がり、米国のダウ工業株30種平均や米国ラッセル2000などの指数が上昇した一方、半導体関税引き上げの懸念から台湾と日本、韓国などが急落した。トランプ氏の銃撃事件後、米国株式市場ではトランプ氏の再選が濃厚と予測され、トランプ氏が力を入れている石油・ガス関連や金融などに資金がシフトしたが、他方バイデン大統領が推進してきた再生エネルギー分野が軒並み下落している。

次に2022年生成AI「ChatGPT」の登場により、世界中でAIブームが巻き起こり、その後AIブームに乗って更にAI半導体の需要が爆発的に上昇している。代表的なAI半導体メーカーであり、世界最大の恩恵を受けているのが、米国のNVIDIAである。AI半導体とは、「深層学習モデルの処理に適したプロセッサ」といい、具体的にはGPU (Graphics Process Unit) やFPGA (Field Programmable Gate Array)、AI特定用途向け集積回路の処理専用チップASIC (Application Specific Integrated Circuit) とCPUなどで構成されている。

マーケットレポートジャパンによると、AI半導体はAIに最適化された機能を持ちAIアルゴリズムが必要とする計算やコンピューティングタスクを高速化でき、小型のトランジスタに依存しているため、消費エネルギーも少ない。またデバイス上でAIプログラムを実行できるため、データ保存のためにクラウドに依存する必要がなく、自然言語処理 (NLP)、ロボット工学、ネットワーク・セキュリティを提供するAI半導体は、情報技術、防衛業界、ヘルスケア、小売業や自動車業界で応用されているという¹。

2022年 IMARC社の市場調査によると、世界の人工知能チップ市場規模は2022年の134億ドルから2028年に896億ドルまで到達し、2023年～2028年の予測期間中に年平均34.5%成長すると推測される²。

今後のAI半導体の需要予測については、様々な業界でAI技術の採用が大幅に増加している中で、特にアジア太平洋地域におけるAI半導体の需要が増えると予測されている。日本のみならず、中国や韓国、台湾、オーストラリアなどにおいて、製造業界や医療業界におけるAIサービスの導入が著しく、AI半導体のコスト削減や性能向上、スマートフォンを通した即時性の高いサービス提供への需要増加などの要因が背景にあると見られる。

さらに主要なAI半導体関連メーカーとして米国のNVIDIA、AMD、英国のARMや韓国のサムスンとSKハイニックス、オランダのASML、台湾のTSMCなどがあり、これらの企業は今後グローバル市場における地位やシェアを更に拡大していくと見られる。

本論文では、グローバルAI半導体企業 (NVIDIA、TSMC、サムスン、SKハイニックス) の動向について考察する。

2. AI 半導体を巡る諸問題

米国は2022年10月、中国に対する半導体輸出規制（CHIPS・科学法）を発動した。1つ目は、AI処理やスーパーコンピュータに利用される高性能半導体の輸出規制で、もう1つが最先端半導体の製造に利用される半導体製造装置の輸出規制である。

しかし半導体輸出規制下においても、米国半導体装置メーカーの対中国輸出が増加しつつあり、規制がかからない非先端品の輸出を中心に中国向け売上高比率は4割を超えている。

クリス・ミラー（2023）は、台湾を巡る米中の軍事的緊張は、かつてなく高まっており、TSMCの最先端半導体製造工場にミサイルが一発打ち込まれれば、携帯電話、データセンター、自動車、通信ネットワーク、その他テクノロジーの生産遅延が累積し、瞬く間に何千億米ドルという損害が生じかねないと指摘している。

コロナ後、原油を超えて世界最重要資源になっているAI半導体を巡って、米中対立の激化による緊張感が増幅している。世界最大手ファウンドリー企業TSMCの生産拠点が、台湾に集中するリスクが顕在化したことにより、TSMCは日本をはじめ米国やヨーロッパに生産拠点の分散化を推し進めている。またTSMCが生産している半導体は、半導体材料や製造装置を日本、米国やオランダなど外国から輸入されているため、台湾有事の際には半導体生産がオールストップされてしまい全世界的に半導体調達に壊滅状態に陥るリスクが高い。

近年膨大な計算が必要な生成AIの利用拡大により、世界各地でAI向けデータセンターの整備が急ピッチに進められている。例外なく日本でも成田空港の利便性や地盤の頑丈さから千葉県印西市に米国ビッグテック企業のデータセンターが乱立している。

米国のビッグテック4社（マイクロソフト、アルファベット、アマゾン、メタ）が発表した24年1月-6月のデータセンターの設備投資は、合計1,000億米ドルを超え、前年同比5割増となっており、半導体メーカーの景気に大きな影響を与えている。

さらにAI向けデータセンターの急拡大により、大量の電力消費や環境への負荷などが懸念されている。米国ムーディーズによると、世界のデータセンターの容量は今後5年で2倍となり、データセンターのAI向け電力使用の増加分はオランダ使用量に匹敵、国際エネルギー機関は2026年の電力消費量が22年から最大で2.3倍になると試算した。

AI半導体サプライチェーンの地政学的リスクと合わせて、世界各地で乱立するデータセンターの大量の電力消費による温暖化ガス排出量の増加、環境への悪影響及び米国ビッグテックによる過度な巨額投資競争などの課題が露呈されている。

3. 米台韓AI 半導体企業の競争と協働

3.1 世界のAI 覇者、米国の半導体ファブレス企業「NVIDIA」の戦略

2024年6月、NVIDIAは、一時時価総額3.3兆米ドルとなり、アップルやマイクロソフトを

抜き世界一になった。2022年11月に世界に登場したChatGPTによって世界中で生成AI半導体の開発が活発になったが、とりわけNVIDIAの存在感は圧巻である（図1）。



図1 米国大手企業の時価総額推移

出典：Yahoo ニュースから抜粋

NVIDIAは1993年台湾系米国人ジェンスン・ファンによって創業され、その後生成AIブームに乗って爆発的にGPU（画像処理装置）の需要が高まり、急成長を遂げている。同氏は、常日頃から「NVIDIAは市場を取りにいく会社ではなく、市場を作る会社だ」という経営信条に則り、同社を牽引してきた。

元々 NVIDIA は、パソコン向け GPU を設計して台湾 TSMC に生産してもらうファブレス企業で、他社が真似できない高い技術をいち早く開発、次々投入して競争に勝ち抜いた結果、高い市場シェアを獲得できた³。

NVIDIA が最も得意とする GPU は、AI や画像認識の高速データ処理に不可欠で、IT 企業の注目の的となっている。1 年で売上高が 2 倍、株価が 3 倍と波に乗る同社の勢いが、関連企業まで波及し、技術のアピール合戦となっている。とりわけ生成 AI 「ChatGPT」を開発した米国オープン AI の技術やアップルの最新型新製品 VR ゴーグル端末「ビジョンプロ」、米国アプロニックは、NVIDIA の最新ソフトを活用してヒト型ロボットを工場や物流向けに開発供給している。

NVIDIA は、主力の GPU 「H100」と開発ツールソフトウェア「CUDA」によって大きくリードしており、CUDA の 2023 年総ダウンロード数は 4,500 万回、開発者は 400 万人にのぼり、AI 開発における事実上のスタンダードとなるため、大口取引先にとって他社の代替製品への切り替えは困難である。

エヌビディアは相次ぎ新型GPUを投入	
製品名	概要
H100	現在の主力製品で業績拡大をけん引、需要急拡大で品薄に
H200	3月に出荷開始、性能テストでH100よりも最大45%処理速度を向上
B200	年内に投入、CPUなどと組み合わせで既存製品より性能30倍に

図2 NVIDIAの製品シリーズ

出典：日本経済新聞 2024年3月25日朝刊より転載

NVIDIAは、2024年3月18日米国のサンノゼで開催された「GTC2024」で、新型GPUの次世代AI半導体「B200」を年内投入すると発表、既存製品より30倍の処理能力があるとされる(図2)。特に最新型の「B200」は、2つのGPUダイ周辺に韓国SKハイニックス製のHBMを使用して並べて、1つにパッケージングされており、その価格は現行の「H100(約350万円)」よりも1.4～1.7倍(約500～600万円)ほど高額になると見られるが、AIサーバー向けに不可欠なため需要は高い。

しかし、NVIDIAにとって今後課題になるのは、最大の顧客と競い合わなければならないことかもしれない。オラクルを加えたビッグ4企業はNVIDIAの大口顧客であり、4社から得ている収益は全売上高の40%以上を占めている。とりわけ米国のアマゾンやグーグル、マイクロソフトなどビッグ3企業による、データセンターへの設備投資額が、2025年に2,300億米ドルと2024年比で5割増える見通しで、自社サービス向けAI半導体の独自開発を進め実用化を目指しており、NVIDIAへの依存を減らそうとしている。今後ビッグ3企業の半導体開発が加速すれば、NVIDIAにとって最大のライバルになる可能性が高くなるのは否めない。

NVIDIAは、これらの打開策として現在のAI半導体トップを固めながらも、スタートアップ企業への投資に積極的に乗り出している。本格化した2023年以降74件の資金調達に参加し、累計109億米ドルを投資してきた。最近では日本のスタートアップ企業サカナAIに200億円規模の資金を投じており、グローバルスタートアップ企業に投資することでAI半導体の寡占を再強化する狙いがある。

3.2 韓国半導体企業「サムスン」と「SKハイニックス」の戦略

日本は2019年安倍政権下で、韓国向け半導体材料の「輸出管理措置」を厳格化していたが2023年の岸田政権下においてようやく解除され、管理品目で半導体製造に不可欠なフッ化水素やフォトレジスト(感光剤)、フッ化ポリイミドの3品目が、韓国に向けて輸出が再開された。韓国は、2022年尹・錫悦(ユン・ソクヨル)新大統領が就任してから、2024年大規模半

導体産業支援策として、3兆円を支援すると発表した。

文・在寅（ムン・ジェイン）前政権は、日本政府の「輸出管理措置」を機に「日本の依存から脱却せよ」と呼びかけ、半導体材料の「国産化」を目指して急進させた。同時期に半導体サプライチェーンを巡って米中対立が激化し、各国政府は打開策として、各々の国産化を目指し進めていた。

韓国政府は、2030年までに半導体素材や部品、設備の国産化率を現在の30%から50%に引き上げることを目指し、2023年3月時点では、半導体素材の国産化率は約50%、製造装置の国産化率は約20%と推定される。

韓国を代表する世界半導体大手サムスン電子（以下サムスン）は、生成AI普及により爆発的に情報処理が増えたことでデータをより速く伝送できる、次世代半導体「CXL（Compute Express Link）⁴」の開発を急ピッチで進めており、2024年下半年期に出荷する予定だ。2024年7月ソウルで開かれた記者向けCXLの技術説明会で、メモリー事業部の崔・チャンソク常務は、「幅の太い道路をつなぎ合わせ、大量のデータを行き来させるイメージだ」と表現した。サムスンは、2021年から業界を先行してCXLを発明してきたが、2024年6月には米国の大手ソフトウェア企業のレッドハットと共同で初デモ機を立ち上げ、実用化に向けた研究は大詰めを迎えている⁵。

また同年8月6日から8日（現地時間）まで米国のサンタクララコンベンションセンターで開催されたFMSで、AI時代に増加する顧客需要を満たすためのメモリー及びストレージソリューションを披露した。Flash Memory Summitとして知られるこのイベントは、以前の範囲を拡張してFuture of Memory and Storageに名称を変更し、NANDだけでなくDRAMを包括するようになった。AIの拡散というパラダイム転換を迎えながら、サムスンは「The AI Revolution: Fueling New Demands for Memory and Storage」というテーマの基調講演を通じて、今後のメモリー及びストレージ需要に対応できるという自信を示した。この基調講演には、Jim Elliott副社長、呉・ファソク副社長、孫・テクサン常務が参加していた。サムスンは高容量、超高性能、低電力ソリューションの必要性を強調し、V-NAND製品の量産高容量SSD、そしてHBM、CXLなど多様なDRAM製品を披露した。米国NVIDIAが主管した「AIワークロードのためのストレージとメモリーイノベーション」に関するパネルディスカッションにも参加した趙・サンヨン副社長は、AIワークロードの増加に伴うメモリー性能と容量の増加と共に、信頼性が確保された低電力ソリューションの重要性を強調し、次世代製品競争力のためのカスタマイズが業界成功の主要因になる見通しだ⁶。

米政府は、2022年成立した「CHIPS・科学法」に基づき、総額527億ドルを米国内で半導体製造拠点を新增設する企業などに振り分ける。米商務省によれば米国の補助金が支給される主な企業は、インテル85億ドル、TSMC66億ドル、サムスン64億ドル、グローバルファンドリーズ15億ドルとなっている。サムスンは、最大ライバル企業で、世界最大手ファウンドリーメーカーであるTSMCに追従するように米国で2ナノ製品の量産に向けて新工場を整備する。そし

て米政府は2024年4月15日に、サムスンがテキサス州テイラーに建設する半導体新工場と研究開発拠点に、最大64億米ドル（約9,800億円）を補助すると発表した。新工場は、最先端2ナノ（ナノは10億分の一）メートル品や4ナノ品のロジックIC半導体などを26～27年にかけて生産開始を見込む⁷。

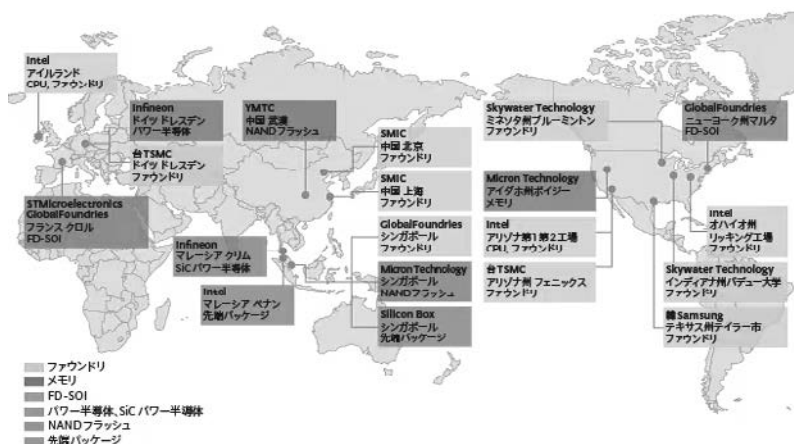


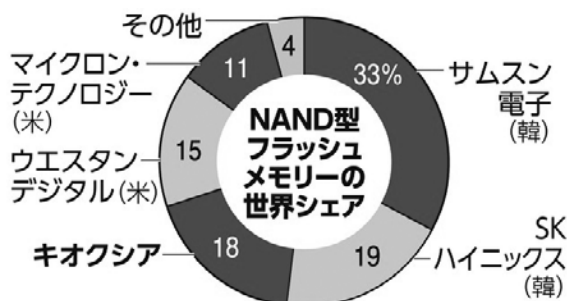
図3 世界の半導体工場の最新動向

出典： https://www.tel.co.jp/museum/magazine/report/202312_01/?section=3

米国は、これまであまり得意ではなかった半導体製造（ファウンドリー）を自国内での製造に切り替え、大手半導体ファウンドリー企業の国内誘致に力を入れている。またヨーロッパや日本、東南アジアなどでも新しい工場建設が活発に行われている（図3）。

サムスは、2024年4月にデータの長期保存に使用する第9世代のNAND型フラッシュメモリーで世界最先端となる286層の量産を始めたと発表した。これは現行の第8世代236層に比して1.5倍容量が高くなったことで、データの出入力の速度が33%速く、消費電力量は10%抑えられる。大量の情報を学習・処理する生成AIの普及で、記憶の役割を果たすNANDは、大容量化が求められる⁸。

2023年NAND型フラッシュメモリーの世界シェアをみると、サムスン33%、SKハイニックス19%、日本のキオクシア18%、米国のWD15%の順となっている。主に韓国メーカーの2社が、全体の52%を占めていることがわかる（図4）。



2023年、売上高ベース、英調査会社オムディア調べ

図 4 2023 年 NAND 型フラッシュメモリの世界シェア

出典：朝日新聞 2024 年 4 月 17 日より転載

SK ハイニックスは、AI ブームに上手くのれた 1 社でもある。同社の AI 向け広帯域幅メモリー「HBM (High Bandwidth Memory)」は、爆発的な量産が期待され、2024 年には DRAM に占める比率が二桁になり、AI チップ市場の 80 % を占める NVIDIA に唯一「HBM3」を供給していることから、SK ハイニックスの人気はにわかに高まっている。一方サムスンの HBM は、消費電力の問題から NVIDIA の AI 用プロセッサに使用するためのテストに合格出来なかった。半導体メモリーの王者であるサムスンの立ち遅れは予想もしなかったことである⁹。

HBM とは、高性能な DRAM で、一次記憶装置として大きな容量や読み書きスピードを満たすよう DRAM を何層にも重ね合わせて、1 つの超高性能 DRAM として使えるようにしたもので、主にデータセンターのサーバーで行われる AI 学習や推論に必要で、一般的にパソコンやスマートフォンなどで使用されている DRAM に比べて処理能力が非常に高く、高額であることが特徴である。

ようやくサムスは、NVIDIA に、第 4 世代の高帯域幅メモリー「HBM3」を供給できるようになる。HBM3 は、2024 年 8 月から中国市場に出荷予定の NVIDIA の「H20」GPU に採用される見通しだ。NVIDIA の従来の調達先である SK ハイニックスは、HBM3E の生産を増やすために HBM3 の供給制約が生じるとみられており、NVIDIA は HBM3 を供給する第 2 のサプライヤーを必要としている。但し、サムスの HBM3 が中国向け以外の AI プロセッサにも搭載できるかは、いまのところ不透明だ。サムスは 2023 年 7～9 月期に HBM3 を初めて量産し複数顧客への供給を開始すると同時に、NVIDIA とも取引を開始している¹⁰。

表1 SKハイニックスとサムスンのHBM量産ロードマップ

	SKハイニックス	サムスン電子
'24	3月にHBM3E/8層の量産 7～9月期にHBM3E/12層の量産	上期にHBM3E/8層量 産と12層の承認審査中
'25	HBM4/12層の量産計画	HBM4の開発計画
'26	HBM4/16層の量産計画	HBM4の量産計画

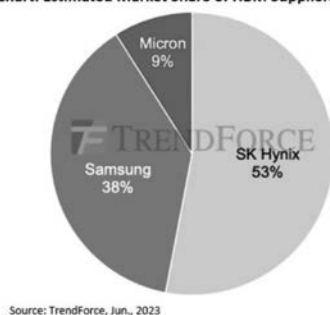
(電子デバイス産新新聞調べ)

出典：電子デバイス産業新聞 2024年8月9日より転載

サムスは、2024年7～9月期にAI向けの高付加価値製品であるHBM3EもNVIDIAの品質テストを通過できるのか注目されている。サムスがNVIDIAへのHBM供給を拡大できれば爆発的な成長を見せるAI半導体市場の波にのることができ、世界のHBM市場で独走を続けているSKハイニックスに巻き返すことができる。今後も第6世代HBM4を、2025年の本格量産に向けて計画を進めている（表1）。

2023年台湾の調査会社Trend Forceによると、世界のHBM市場のシェア率は独走中のSKハイニックス53%、追従してサムスンが38%、米国のマイクロンが9%となっている。これはNAND型フラッシュメモリー同様、HBM市場においても韓国勢が、90%以上を独占していることがわかる（図5）。

Chart: Estimated Market Share of HBM Suppliers in 2023



Source: TrendForce, Jun., 2023

図5 2023年HBMシェア

出典：韓国ネイバーニュース 2023年6月27日より転載

世界のHBMメモリー大手3社は、今後の研究開発と増産に向けて工場を新設する。SKハイニックスは、韓国の清州市（韓国中部地域）に2兆円、他地域にも計12兆円を投じHBM工場を新設するほか、TSMCと次世代品の技術提携を交わす。また米国インディアナ州に2028年の量産計画に向けて組立工場と研究開発施設を新設する。一方サムスは、HBM3Eの研究開発に集中投資する。同社はAI市場の大波に乗り遅れた教訓から急ピッチで研究開発を進める。米国のマイクロンは、2025年から広島工場でHBMの量産を始める。同社は、広島工場に最大5,000億円を投じ、現状10%のシェア率を2025年に25%までに高めていく（図6）。

メモリー大手がHBM量産を急ぐ	
SKハイニックス(韓)	・韓国で2兆円投じ工場新設 ・米国でHBM組み立て工場と研究開発施設を新設
サムスン電子(韓)	・24年にHBMの供給能力を前年比3倍以上に増やす計画
マイクロン・テクノロジー(米)	・広島工場で次世代技術を導入し最新型HBMを増産

図6 HBM大手3社の量産計画

出典：日本経済新聞 2024年5月1日朝刊より転載

今後もサムスンとSKハイニックスは、大口取引先であるNVIDIA向けにカスタマイズ高性能メモリーを開発生産することに注力していくと見られるが、半導体の後工程において不可欠な日本企業の存在が大きく関わっているため、今後の日韓関係に左右されることもあり、更に注視すべきである。

3.3 台湾の世界最大ファウンドリー企業「TSMC」の戦略

TSMCは、1987年台湾系米国人のモリス・チャンによって創業された世界最大のファウンドリーメーカーである。前年1986年に日本は、日米半導体協定を結んだが、東芝、日立、NECなどの日本のメモリーメーカーは、当時世界の半導体市場を席巻していた。

日米半導体協定により、日本の半導体メーカーは高関税を課せられたため、価格面から米国や韓国半導体メーカーに太刀打ちできなかった。その結果日本のメーカーは、サブコントラクター（半導体後工程）に転落したものの、近年爆発的な需要が伸びているAI半導体の量産に、日本の半導体後工程関連メーカーが不可欠であるため、世界半導体市場において改めて日本のメーカーの重要性が再認識されている。

近年世界中でコロナ禍による空前の半導体不足が続き、その重要性に気づいた日米欧は、TSMCの工場誘致に躍起になっている。TSMCは、ソニーとデンソーの出資により2024年2月日本の熊本県菊陽町で半導体第1工場が完成した。

九州には、既に日本の主な半導体関連企業が多く進出しており、TSMCの熊本進出は、グローバルな生産拠点の1つとして重要な役割を果たすと見られる（図7）。



図7 九州の主な半導体関連日本企業の工場

出典：西日本新聞 2021年10月29日より転載

TSMCは、日本のみならずドイツのドレスデンや米国のアリゾナ州などで工場建設を盛んに進めている。米国のアリゾナ州工場は、熊本工場建設よりも早期（2021年4月）に着手されたが、先端技術の半導体製造向け工場ということで稼働までにまだ時間がかかる。遅れの最大要因は、現地の半導体エンジニア不足である。米国では半導体エンジニアが不足気味で台湾から多くのエンジニアを派遣しているのが現状だ。ただ台湾からのエンジニアに対して現地労働組合からの反発や残業、AS、災害対処などの労働文化の構造的な相違、高い生産コストなど課題山積である。また先端半導体生産には、多大な電力を必要とするが、台湾の電気代は安い。装置や材料のサプライチェーンがTSMCのために集積し、数十年かけ効率化されたがアリゾナ州はそうではなく、人件費も割高だ¹¹。

米国の工場と比べて熊本工場建設が、計画から完工までスムーズに進んだ理由として、ソニーなどのパートナーの存在が大きかった。とりわけソニーは立ち上げから多くの有能なエンジニアを派遣し、惜しみなく支援してきた。それは両社の大口取引先である米国アップルの存在があるからだ。すなわち、アップルのための相互扶助パートナーシップである。

TSMCは、他社で設計された半導体を製造し、2022年メモリーを除く世界の半導体生産額の30%を占めており、高性能な半導体製品を生産できる技術力に強みがある。具体的には3nmプロセス技術と呼ばれる最先端技術を保有し、主要顧客にはNVIDIA、アップル、ブロードコム、クアルコム、AMDなどがあり中でもアップルが売上高の約25%を占めている。

また5ナノデバイスでTSMCに製造委託している主な大口取引企業は、米国企業6社、台湾企業1社、中国企業1社であり、その委託内容も多種多様である（表2）。

表2 5 ナノデバイスでTSMCに製造委託企業

企業(国)	委託内容(製品、事業)
アップル(米)	スマホやタブレットPC用プロセッサを製造委託中、また、パソコン用CPU(中央演算処理装置)をインテル製からアップルの自社設計に切り替え(アップル・シリコン)という名称で製造委託
インテル(米)	買収した旧アルデラ社が手がけていたFPGA(回路を自由に書き換えられる半導体)事業 GPU(グラフィックス・プロセッシング・ユニット=画像処理用半導体)、CPU、インテル社内の生産トラブルで新たにTSMCへ製造委託を交渉中か
クアルコム(米)	5Gスマホアプリケーションプロセッサ、モデムなど
AMD(米)	GPU、CPU
エヌビディア(米)	GPU
メディアテック(台)	5Gスマートフォンアプリケーションプロセッサ、モデムなど
ビットメイン(中)	仮想通貨ビットコインのマイニング(採掘)用チップ

出典：週刊エコノミストオンライン 2020年10月14日より転載

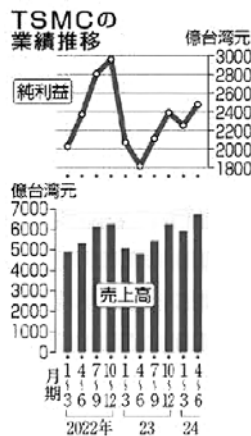


図8 TSMCの業績推移

出典：沖縄タイムス 2024年7月19日より転載

2024年7月18日に発表されたTSMCの4～6月期決算をみると、売上高は前年同期比40.1%増の6,735億台湾元、純利益は36.3%増の2,478億台湾元で、共に過去最高を記録し、7～9月期も増収を見込む(図8)。その理由として上記の表2で示しているようにAI関連からパソコンやスマートフォン、仮想通貨マイニングチップまで多種多様な半導体製造を一括受託することで半導体を総取りしている点にある。

モリス・チャンは、数年前からTSMC「グランドアライアンス(台積大同盟)」という概念を提唱してきた。これは第二次世界大戦中に英米がドイツ、イタリア、日本の枢軸国に対抗するために結んだ連合国の概念を取り入れたものだ。今、TSMC「グランドアライアンス」はNVIDIA、AMD、クアルコム、ブロードコム、台湾メディアテックといった古馴染みの顧客だけでなく、アップル、マイクロソフト、テスラ、アマゾン、アルファベットといったシステムメーカー、IPベンダーのARMやEDAベンダーのシノプシス、ASMLのような半導体装置メーカーも包括している。分業と「グランドアライアンス」の協力構造を通じて、TSMCは半導体産業をひっくり返すほどの力を手に入れた¹²⁾。

半導体市場の成長は、ChatGPTの出現によってAIが具現化されたことで、半導体業界に巨大なビジネスチャンスを与えた。とりわけChatGPTによって、GPUを生産するNVIDIAやAMDなどが恩恵を受けたが、最終的な勝利者はTSMCである。

またマイクロソフトとオープンAIの提携、Googleの生成AIグーグルバード、メタの大規模言語モデルLLaMAなどがリリースされることによって、自社設計が進むにつれて、半導体を確保する必要性から、最終的に全てTSMCに発注することになる。

更にTSMCの最大の強みは、健全かつ潤沢な資金を有する財務力、社内政治や株主に左右されないコーポレートガバナンス、社員の持株制度、社内に米国型企業文化を取り入れたマネジメントスタイル、長年旧知の協力関係であるオランダのASML、台湾系米国人企業であるNVIDIAやAMDとの緊密な信頼関係、NANDフラッシュメモリのマイクロンとの提携など枚挙にいとまがない。TSMCにとって最大のライバルはTSMC自身であり、今後も数十年間世界最大ファウンドリー企業としてトップを走り続けるに違いない。

4. おわりに

近年中国経済の急激な失速により、以前よりは米中の経済摩擦が微小になったように見られるが、トランプ氏の大統領再選が確実となり、今後のグローバル経済の地殻変動は、以前に勝っても劣ることはないと予想される。

アフターコロナ後、世界中で米国の生成AI「ChatGPT」によって、AI技術は米国1強主義となり、米国主要企業による寡占状態に陥っている。とりわけAI半導体の開発が急速に進むにつれて、爆発的にNVIDIAのGPU需要が高まり、米国大手クラウド企業やテクノロジー企業をはじめ世界中の企業にとって、今後のAI半導体の安定的な供給が不安視されている。日本でも米国の最先端技術やサービス導入が広がり、以前にも増して米国への依存度が高まり経済安全保障上のリスクになりかねない状況に陥っている。

米国企業によるAI技術の寡占は、グローバル企業にとって危機感を増幅させ、AI半導体の調達先としても最大のリスクが内在するため、安定的な供給を目指し、TSMCの自国内工場誘致にしのぎを削っている。

最後に、TSMC、SKハイニックス、サムスンは、半導体関連原材料の供給先として、その多くが日本の企業によって支えられているため、東アジアに偏重していることがわかる。世界中のAI半導体の製造とその原材料の最供給地は東アジアに集中していることから、不安定な朝鮮半島情勢や台湾有事などが勃発した際、半導体地政学的リスクを回避することが喫緊の課題となる。そのリスクヘッジの1つとして、これらの企業は積極的に欧米において工場建設や計画を推進しており、さらなる半導体の国際分業体制が再構築されつつある。そのためには次世代の高度な人材育成と安定的な人材供給の確保と合わせて、安定的な政治情勢と国家安保体制が不可欠である。

【注】

- 1 マーケットレポートジャパン
<https://www.marketreport.jp/artificial-intelligence-chip-market-global-imarc23jul0118>
(2024年8月13日閲覧)
- 2 マーケットレポートジャパン
<https://www.marketreport.jp/artificial-intelligence-chip-market-global-imarc23jul0118>
(2024年8月13日閲覧)
- 3 週刊東洋経済 2024年8/10-17 p45
- 4 CXL (Compute Express Link) : 高性能サーバーシステムで、CPUと共に使用されている加速器、メモリ、貯蔵装置などをより効率的に活用するために新しく設計された相互接続。
<https://semiconductor.samsung.com/kr/news-events/news/>
(2024年7月10日閲覧)
- 5 日本経済新聞 2024年8月14日朝刊
- 6 <https://semiconductor.samsung.com/kr/news-events/tech-blog/>
(2024年8月16日閲覧)
- 7 日本経済新聞 2024年4月16日朝刊
- 8 日本経済新聞 2024年4月26日朝刊
- 9 電子デバイス産業新聞 2024年8月9日及びアジア日報 2024年3月19日
- 10 電子デバイス産業新聞 <https://www.sangyo-times.jp/article.aspx?ID=13503>
(2024年8月10日閲覧)
- 11 週刊東洋経済 2024年8/10-17 p56
- 12 林 宏文 (2024年)『TSMC世界を動かすヒミツ』p137, CCCメディアハウス.

【参考文献】

太田泰彦 (2022年)『2030半導体の地政学』日本経済新聞出版.

林 宏文 (2024年)『TSMC世界を動かすヒミツ』CCCメディアハウス.

クリス・ミラー (千葉敏生訳) (2023年)『半導体戦争』ダイヤモンド社.

塩地 洋・田中 彰編 (2020年)『東アジア優位産業—多元化する国際生産ネットワーク』中央経済社.

センス・アンド・フォース (2019年)『半導体業界の動向とカラクリがよくわかる本』秀和システム.

週刊ダイヤモンド (2023年5/27)『半導体・EV&電池』ダイヤモンド社.

週刊東洋経済 (2024年8/10)『エヌビディアの猛威』東洋経済新報社.

AI半導体関連 <https://www.semiconportal.com/archive/blog/insiders/nagami/240129-pickup764.html>

津田健二 (2024年)『エヌビディア半導体の覇者が作り出す2040年の世界』PHP研究所.

Trends on Global AI Semiconductor Companies

Mie Yoshioka

Abstract

The emergence of the 2022 generation AI “ChatGPT” has led to an explosion in demand for AI semiconductors worldwide. The U.S. semiconductor giant NVIDIA accounts for 80% of these AI semiconductors.

Behind the spread of NVIDIA’s GPUs in the development and use of generative AI is a software platform called CUDA, which the company began offering in 2006. Initially intended for image processing, it became the standard bible for AI developers around the world because it enabled general-purpose computation. NVIDIA’s main products are designed by NVIDIA, with SK Hynix and Samsung supplying the high-performance memory HBM, TSMC the world’s largest foundry, and Japanese companies managing the post-processing and packaging of the final product.

In addition, since expensive GPUs and huge power consumption are the biggest weaknesses of the AI data centers that are being built around the world to handle the huge number of calculations required, U.S. big tech companies have begun to develop their own power-saving semiconductors specialized for AI processing. Although the performance of these models has not reached the level of “ChatGPT,” they can be used for everyday applications such as text summarization and image search. These new developments in U.S. Big Tech pose a potential threat to how long NVIDIA’s one powerhouse can continue.

This paper therefore addresses the corporate strategies of global AI semiconductor companies NVIDIA, TSMC, Samsung and SK Hynix.

Keywords: AI-Semiconductor, Graphics Processing Unit, High Bandwidth Memory