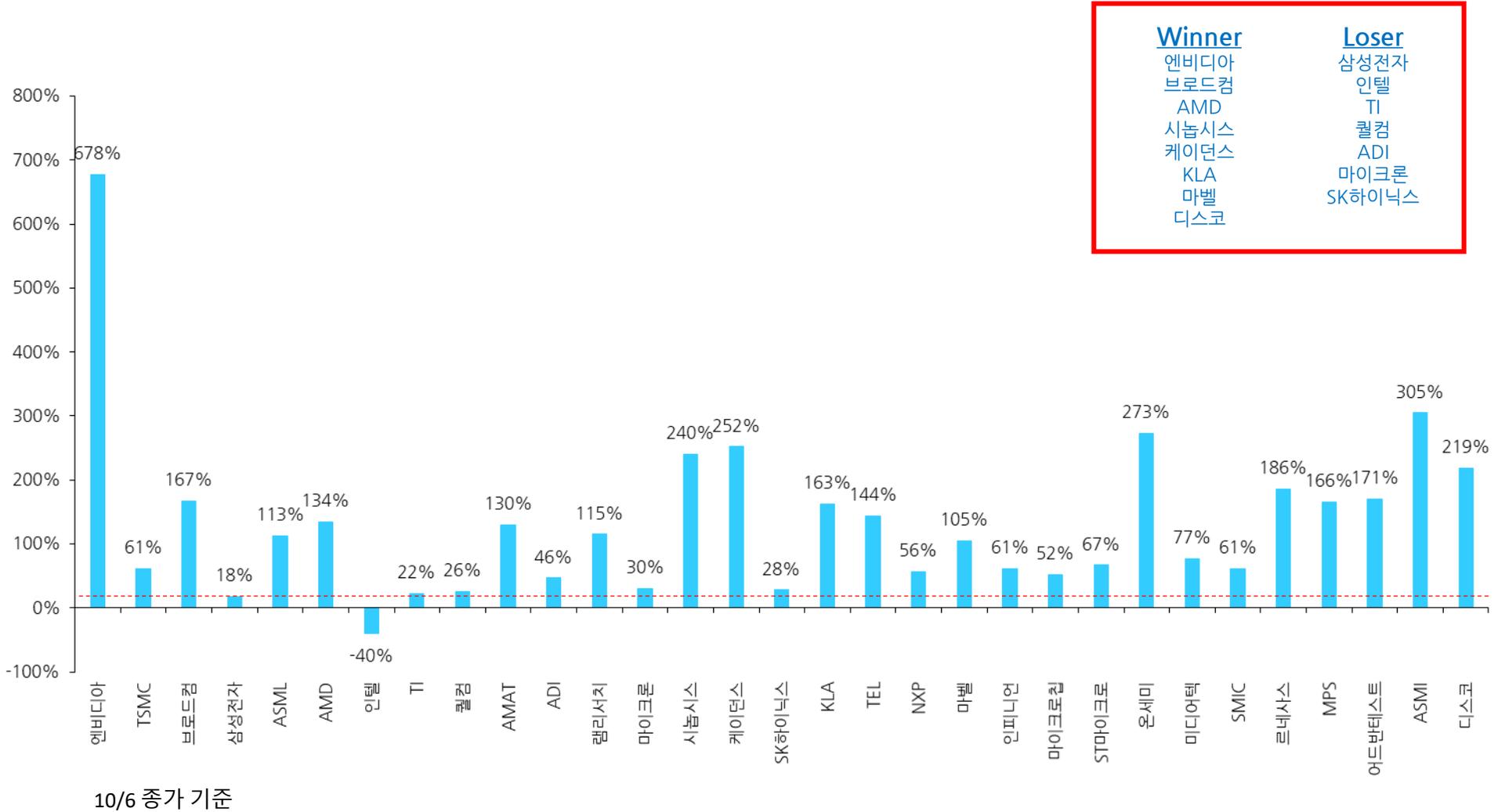


반도체

글로벌 AI 열풍 속 국내 소부장에 미치는 영향

이 승 우
유진투자증권 리서치센터장 / 반도체
swlee6591@eugenefn.com

팬데믹 이후 반도체 주가 상승률 비교 (2019년말 대비)

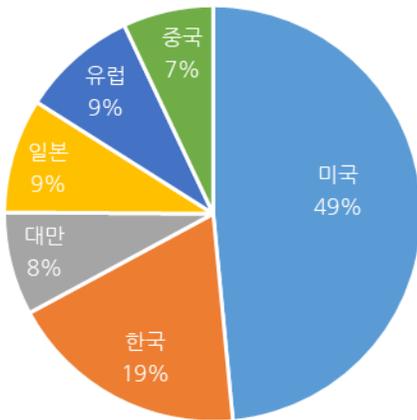


한국 반도체 산업의 냉정한 현실

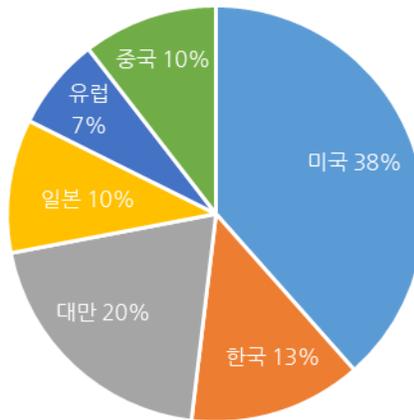
놀라울 정도로 허약한 한국 반도체의 현실

- SIA 기준 주요국 반도체 시장 점유율은 미국 49%, 한국 19%, 유럽, 일본 9%, 대만 8% 순임.
- 그러나, 이 공식 통계에는 예를 들어 TSMC 및 소부장 업체들의 데이터가 포함되지 않음. 따라서, 반도체 산업의 실질적 위상을 나타낸다고 보기 어려움.
- 상장 반도체 기업의 매출 합계 기준으로는 미국 38%, 대만 20%, 한국 13%로 대만에 열세를 보임
- 특히, 상장 반도체 시가총액 기준 점유율은 미국 56%, 대만 13%, 중국 10%, 한국 9%로 더욱 열세
- So What? → 안타깝지만, **한국 반도체 산업의 펀더멘털은 놀라울 정도로 허약** → 현실을 직시해야 올바른 대응이 가능

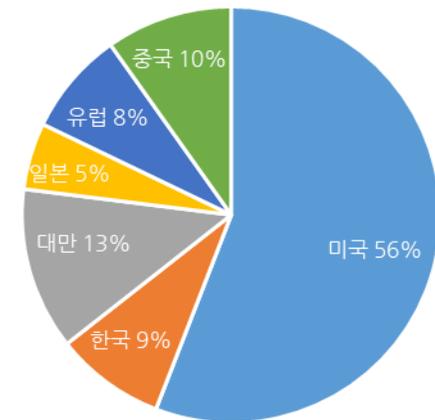
2022 SIA 기준 점유율



2022 서플라이체인 기준 점유율



2023 시가총액 기준 점유율



공격적 투자 + 충분하지 않은 수요 → 과도한 반도체 재고

2분기말 삼성 DS + SK하이닉스 재고 50조원 (삼성 재고는 더 증가)

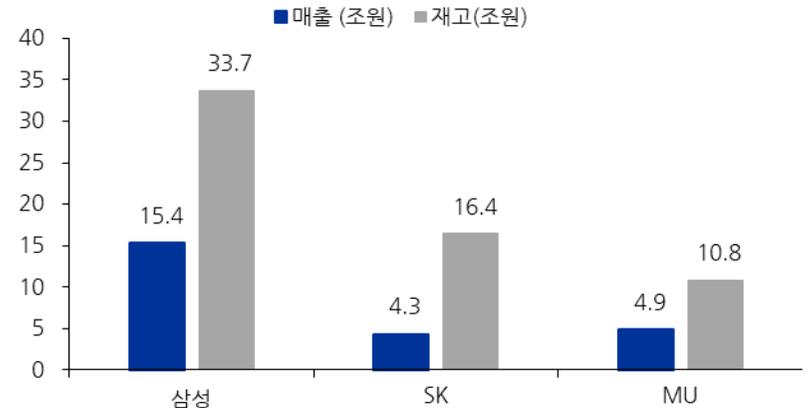
- 2분기 말 메모리 3사와 비메모리 17개 사의 반도체 재고자산 규모는 1,068억 달러로 역대 최고 기록을 계속 경신 중
- 삼성전자의 반도체 매출은 15.4조원이나 재고자산은 33.7조원 → 재고자산/매출 = 2.3X
- SK하이닉스는 재고자산이 1분기 매출의 3.4배. 상각전 기준으로는 3.5배

반도체 제조사 Top 20 재고 자산 규모 - 증가속도 둔화



자료: 유진투자증권

2분기 메모리 3사 재고 vs. 매출



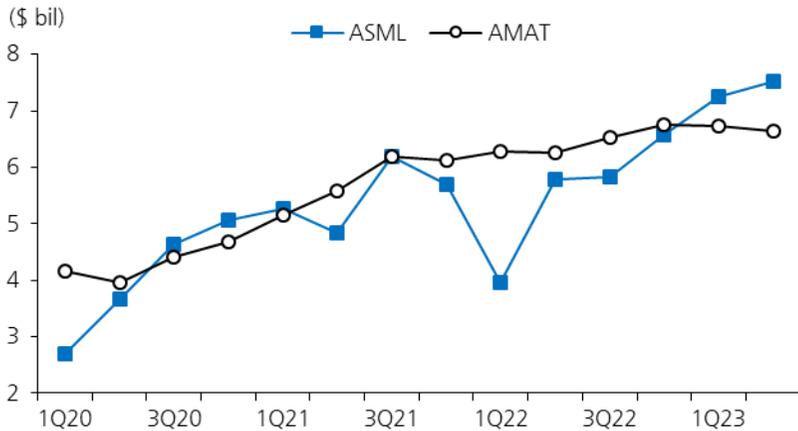
자료: 유진투자증권

막대한 반도체 투자의 수혜는 어디로 갔나?

대규모 메모리 투자의 수혜 → 1) 미국 장비사, 2) 애플, 3) 미국 빅테크

- 투자의 직접 수혜는 메이저 반도체 장비주들
- 애플의 FY 3분기(4~6월) 매출은 818억 달러로 세분기 연속으로 전년비 감소
- 애플의 모든 제품의 매출이 감소했지만, GP 마진은 44.5%로 역대 최고를 기록.
- 반도체 가격 하락이 결과적으로 애플의 수익성에 크게 기여

ASML과 Applied Material의 매출



자료: 유진투자증권

애플 GP 마진 추이



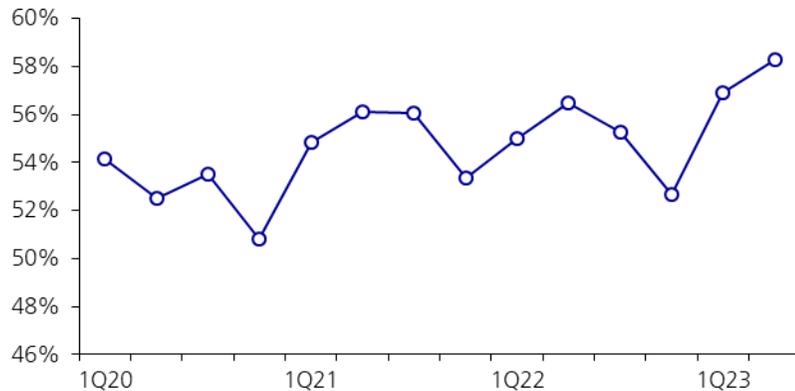
자료: 유진투자증권

- continued

수 십 조원을 들인 파멸적 투자

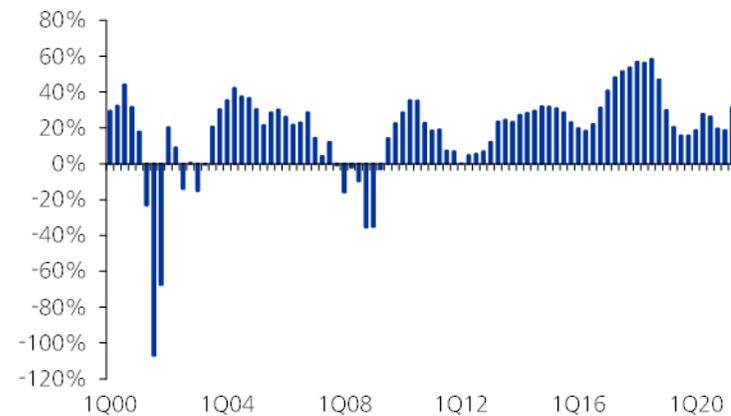
- 마이크로소프트, 알파벳 등 빅테크들의 매출총이익률도 오히려 더 높아져 있는 상황
- 그러나, 투자의 주체인 반도체 업체들은 최악의 실적을 기록: 삼성전자 메모리, SK하이닉스, 마이크론, 키옥시아의 1분기 영업이익률은 각각 -39%, -57%, -47%, -52% 역대급 손실을 기록

빅테크 매출총이익률 추이



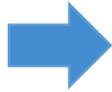
자료: 유진투자증권

메모리 3사 영업이익률



자료: 유진투자증권

이미지넷 프로젝트

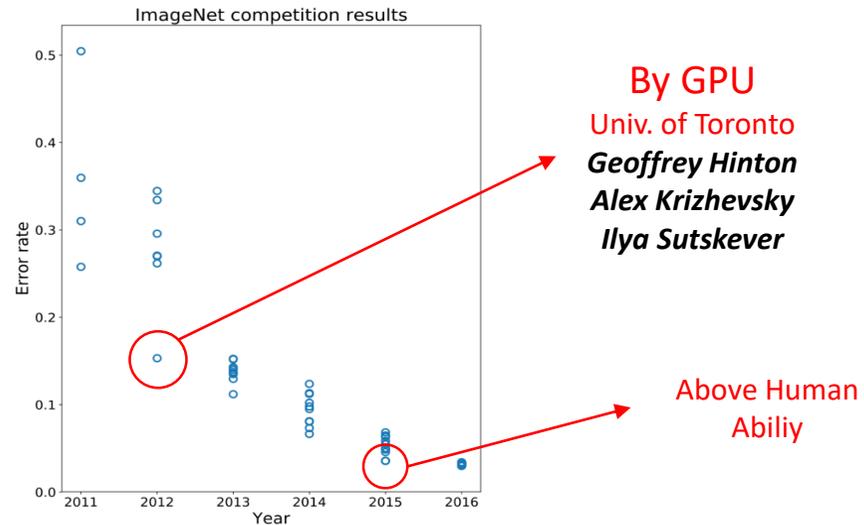


Cannot tell what is what

2007년 ImageNet 프로젝트 (by 프린스턴대 페이페이리)



ILSVRC (ImageNet Large-Scale Visual Recognition Challenge)



Cognition 및 AI 컴퓨팅 시대의 개막

1세대 : 1890~

Tabulating

데이터의 집계



Hollerith Desk

2세대 : 1950~

Programming

데이터의 탐색

기계어 명령어

결정론적 접근

정형 데이터 활용

단순 출력

IMAGENET

3세대 : 2010~

Cognition

데이터의 발견

자연어 명령어

확률론적 접근

빅데이터 활용

지능적 선택

1997 IBM Deep Blue



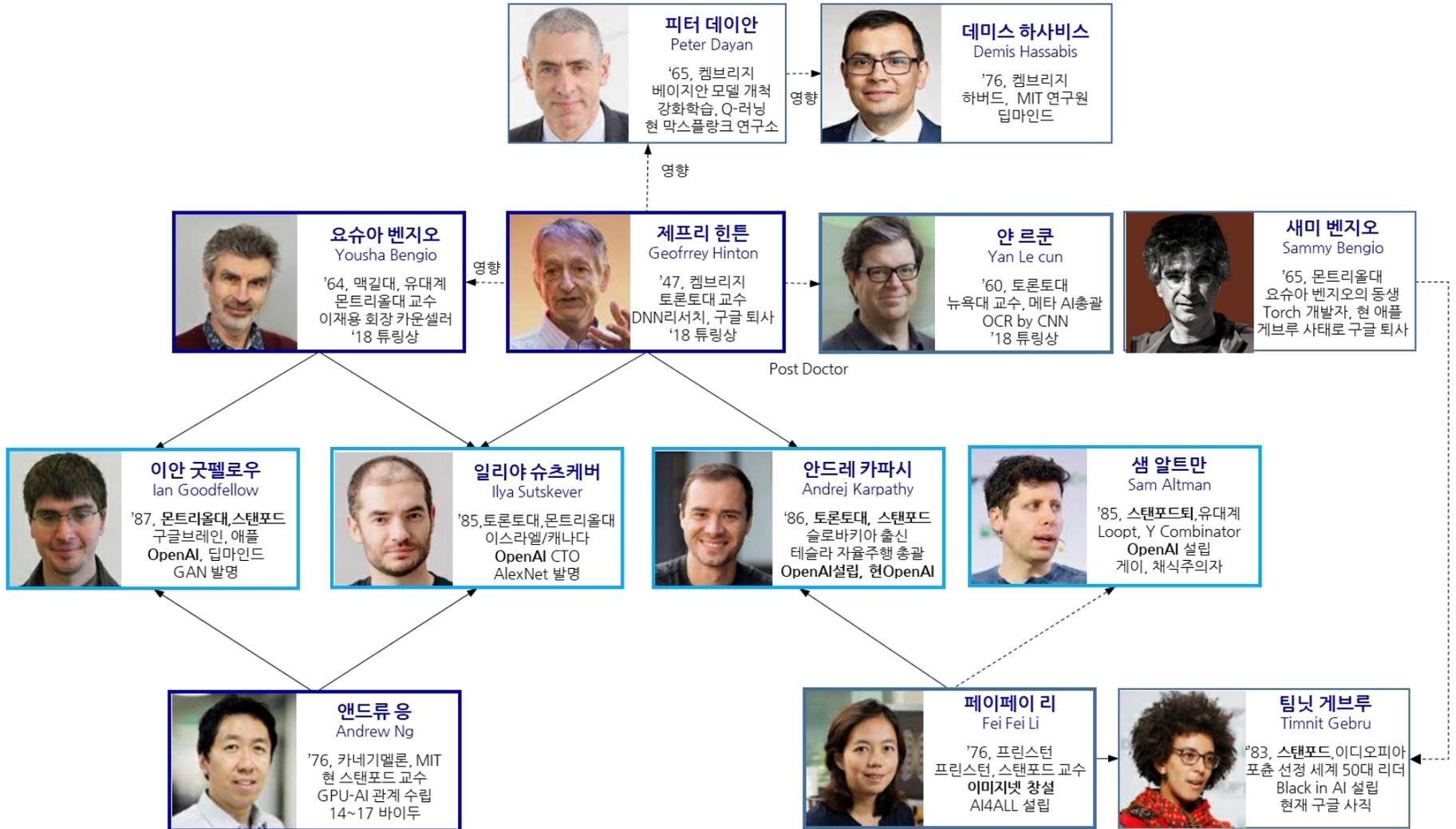
2011 IBM Watson



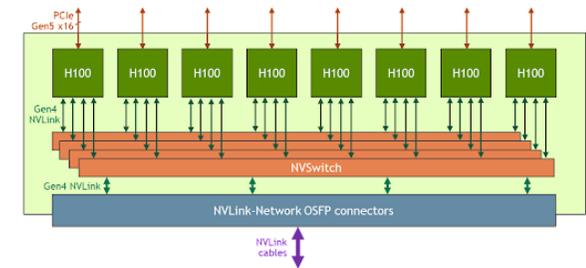
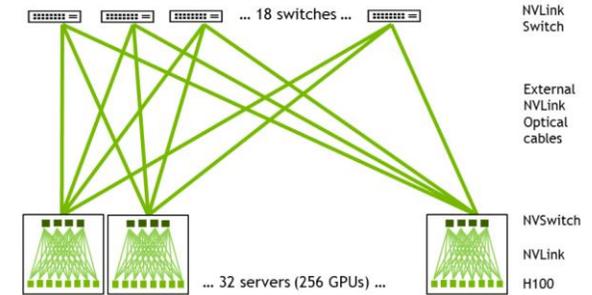
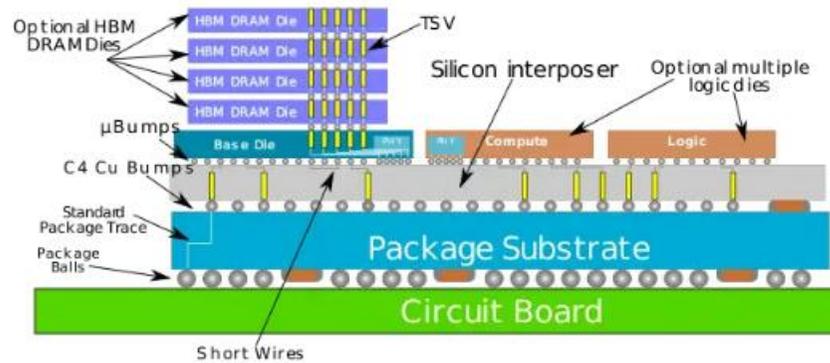
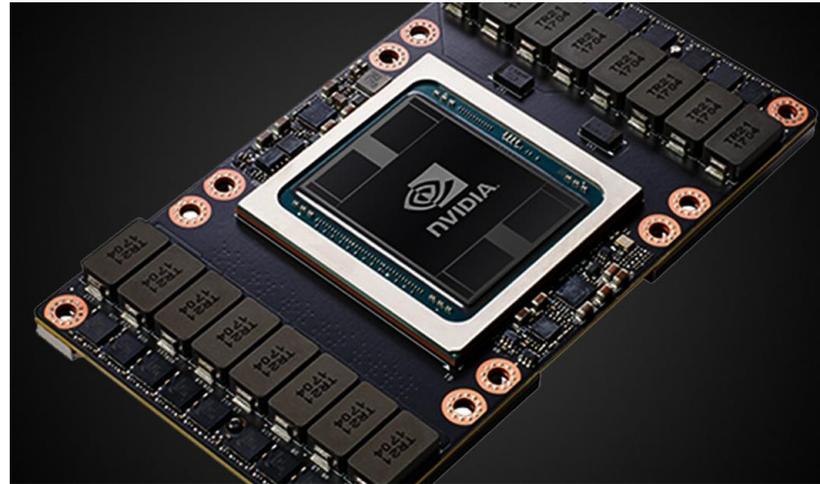
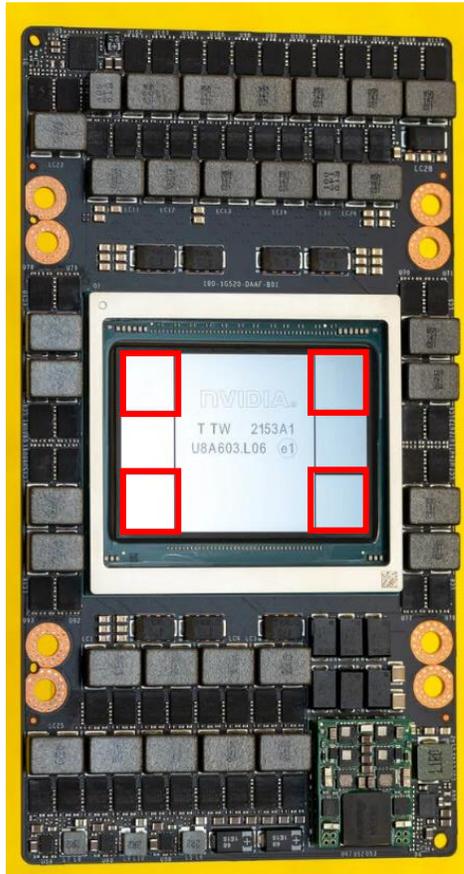
2016 Google AlphaGo



AI를 지배하는 사람들



최고의 히트 상품: 엔비디아 H100

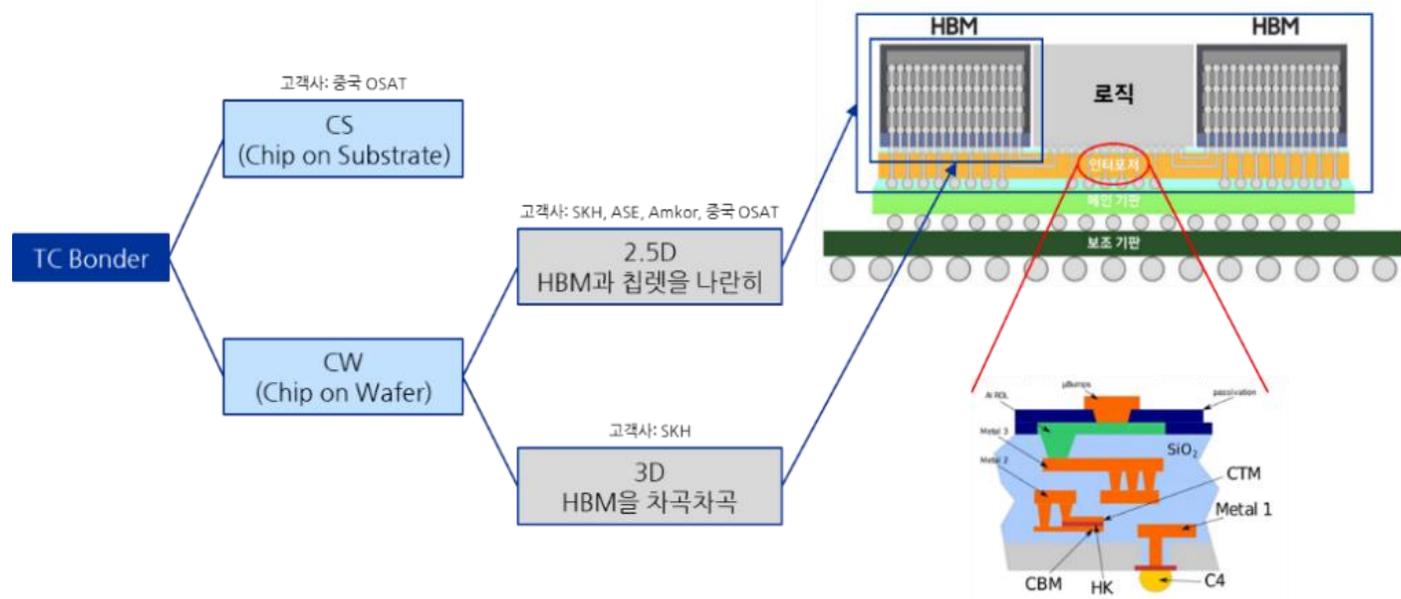


엔비디아 서버 라인업

- 엔비디아의 서버 라인업: DGX(데이터센터급), HGX (HPC급), EGX(엔터프라이즈급), IGX(산업용/엣지급)
- DRAM 스펙은 GPU용 HBM 2/3 640GB(8*80GB) + 시스템 메모리 2TB D4 or D5 DRAM
- D4 평균가격=\$1.5~1.6/GB, D5 평균가격 = \$2~2.2/GB
- HBM2e 가격 = \$10~11/GB, HBM3 예상 가격 = \$16~18/GB → HBM 80GB 가격 = 약 \$1300~\$1400
- HBM3 80GB x 8 = 11000볼 & 시스템(2TB RDIMM) 4000~5000볼 → DRAM BOM = \$15,000~\$16,000
→ BOM ratio = 3.75~4% (vs. PC or Server DRAM BOM ratio = 3~6%)
- DGX 데이터센터는 기존 AI 데이터센터 대비 전력소모 1/20, 스페이스는 1/25로 줄여 실제 COO은 1/10 수준이라 주장.



TSMC의 CoWoS 기술



	제조사	관련 장비사
HBM	SK하이닉스 삼성전자 마이크론	한미반도체 ASMPT Shinkawa
Chiplet	TSMC	BESI
HBM과 Chiplet를 인터포저에 분딩	ASE Amkor	한미반도체 ASMPT

(십억달러)	2022	2027(E)	CAGR
3D SoC	0.02	0.08	31%
3D NAND	0.23	0.81	28%
3D HBM	0.57	1.63	23%
2.5D EMIB	0.32	0.84	21%
3D Foveros	0.42	1.01	19%
3D Stacked Memory	0.59	1.3	17%
Si Interposer	0.19	0.29	9%
UHD FO	1.27	1.88	8%

제 2의 무어의 법칙: 헤테로 인테그레이션

1965년 <일렉트로닉스 매거진>에 실린 고든 무어의 'Cramming More Components onto Integrated Circuits' 를 보면 트랜지스터의 집적화 기술을 강조한 Transistor Focus와 Heterogeneous Integration의 가치, 즉 System Focus가 이미 소개됨.

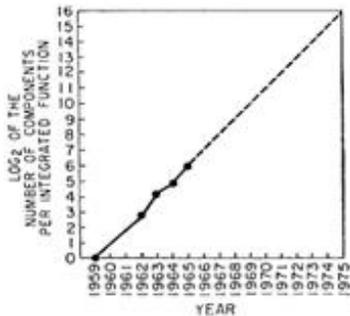


diagram to technological realization without any special engineering.

It may prove to be more economical to build large systems out of smaller functions, which are separately packaged and interconnected. The availability of large functions, combined with functional design and construction, should allow the manufacturer of large systems to design and construct a considerable variety of equipment both rapidly and economically.

IX. LINEAR CIRCUITRY

Integration will not change linear systems as radically as digital systems. Still, a considerable degree of integration will be achieved with linear circuits. The lack of large-

diagram to technological realization without any special engineering.

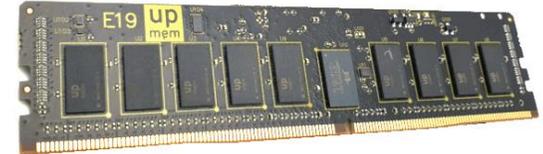
It may prove to be more economical to build large systems out of smaller functions, which are separately packaged and interconnected. The availability of large functions, combined with functional design and construction, should allow the manufacturer of large systems to design and construct a considerable variety of equipment both rapidly and economically.

IX. LINEAR CIRCUITRY

Integration will not change linear systems as radically as digital systems. Still, a considerable degree of integration will be achieved with linear circuits. The lack of large-

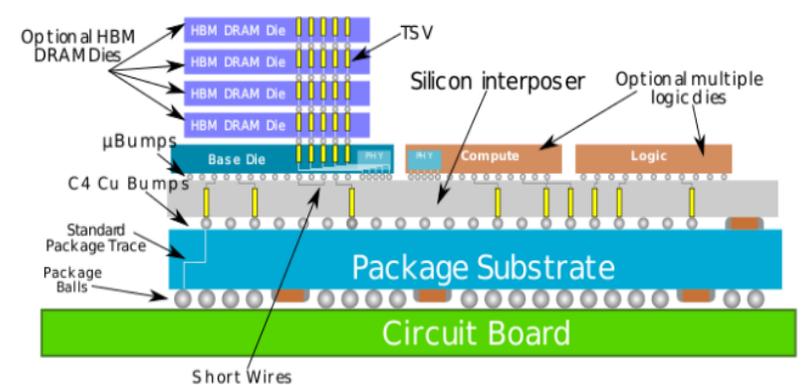
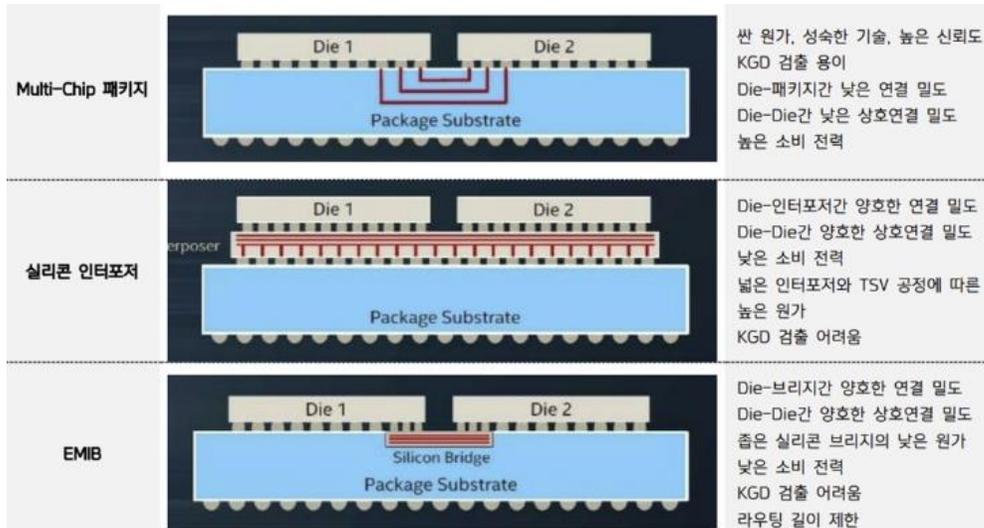
AI 반도체 기술동향

- **인공지능 반도체, 전력반도체 등 신소자 구현을 위한 저전력·고성능·소형화를 목표로 한 회로 설계기술 개발이 한창**
 - (소형화·지능화) 단일 소자에 집적도를 높인 SoC 설계 일반화, 딥러닝에 특화된 아키텍처 연구
 - IoT, 재난안전서비스 등 올웨이즈 커넥티드 기술 내장 추세
 - 0.5V 이하 동작 초저전압소자, 로직과 메모리가 통합된 PIM, 3차원 집적 및 스마트 배선 등 다양한 소재와 신소자 연구
 - SNN(Spiking Neural Network) 등 비폰노이만 구조 반도체 등 이머징 소재를 이용한 두뇌모사 소자/아키텍처 연구
- **고대역폭 메모리에서부터 연산기능 갖춘 메모리 반도체까지 기술 확장**
 - 인터포저 이용해야 하나 고대역폭, 짧은 레이턴시, 작은 폼팩터, 낮은 소모전력 등은 성능 상의 장점으로 작용.
 - (SK하이닉스) 2021년 10월 819GB/s 대역폭을 지원하는 'HBM3' 발표, 22년 6월 양산, 23년 4월 12단 적층 24GB 개발
 - (삼성전자) HBM3 16GB와 12단 24GB 제품 샘플 출하 및 연내 HBM3P(Plus) 선보일 계획
 - (마이크론) 16GB 용량과 410GB/s 대역폭의 HBM2E 제품 개발 완료
- **PIM: 반도체 설계 패러다임 전환을 위한 기술**
 - (AMD) HBM 기반 GPU에 PIM 기능을 HBM의 기본 로직 다이에 추가 → 7% 속도 향상과 85% 에너지 효율성 증가
 - (삼성전자) 2021년 2월 세계 최초로 HBM2 아쿠아볼트에 AI 엔진을 탑재한 HBM-PIM을 개발
 - (SK하이닉스) 2022년 2월 PIM이 적용된 'GDDR6-AiM' 샘플 개발
 - (프랑스 UPMEM) DIMM 형태의 DDR-4 기반 PIM 구조 및 칩을 개발
- **CXL: 고성능 컴퓨팅 시스템에서 CPU, GPU, 메모리 인터커넥트 기술**
 - CXL DRAM은 메인 DRAM과 공존하면서 DRAM의 용량과 대역폭을 확장
 - (삼성전자) 2022년 5월 세계 최초로 CXL 1.1 기반 CXL DRAM 개발 및 1년 만에 CXL 2.0을 지원하는 128GB CXL DRAM을 개발
 - (SK하이닉스) 2022년 8월 CXL 기반 메모리 샘플 개발 10월 CXL 메모리에 연산기능 통합한 CMS (Computation Memory Solution)개발



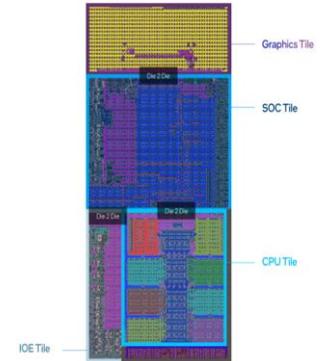
첨단 패키징 기술 경쟁

- TSMC: 2012년 처음으로 2.5D 패키징 기술을 도입. 전체 설비투자의 10%를 첨단 패키징에 사용**
 - 차세대 패키징 기술 핵심인 3D 적층 기술인 SoIC를 2022년 상용
 - 일본 이바라키현 쓰쿠바시에 370억엔을 들여 반도체 연구개발 센터를 설립 → 3D 집적회로의 재료 연구 수행
- 인텔: 2017 EMIB(Embedded Multi-die Interconnect Bridge) + 3D 포베로스**
 - EMIB는 서로 다른 공정과 아키텍처 기반 칩을 통합하는 기술, 포베로스는 컴퓨팅 타일을 수직으로 쌓아 올리는 기술
 - EMIB는 인터포저를 사용하는 점에서는 TSMC의 2.5D 패키징 기술과 비슷하나 패키지 면적을 줄이고 소비전력을 더 낮춤
 - 3D 패키징 기술인 포베로스(FOVEROS)와 EMIB 기술을 접목할 계획
- 삼성전자: 4개의 HBM 칩을 원패키지로 구현한 2.5D 패키지 기술 '아이큐브4'를 '24년 2Q부터 양산**
 - CPU, GPU 등 로직과 HBM 4개를 실리콘 인터포저 위에 배치한 이중집적화 패키지 기술
 - 복수의 칩을 1개의 패키지 안에 배치하여 면적은 줄이고 전송 속도는 높임.
 - 실리콘 인터포저로 초미세 배선을 구현해 구동에 필요한 전력은 안정적으로 공급 가능
 - '23년 로직칩과 8개의 HBM을 구현한 I-Cube8, '24년 12개의 HBM을 통합한 I-Cube12를 개발한다는 계획

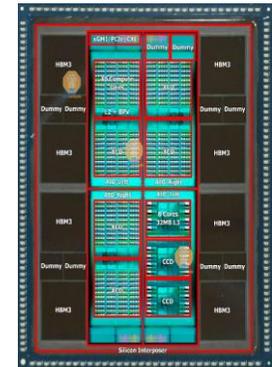


Chiplet Technology 개발

	AMD	인텔	삼성전자	TSMC
기술명	칩렛	MCM (Multi-Chip Module)	Cube	SoC (System on Integrated Circuits)
적용 (예정) 제품	GPU: RX7900, MI300X CPU: Ryzen 시리즈	CPU: 메테오 레이크, Stratix 10 FPGA, 사파이어 래피즈	I-Cube, H-Cube, X-Cube	CoWoS 기술 InFO_PoP 기술
특징	칩렛 기술 최초 도입 제품 생산	자체적인 칩렛 CPU 생산 이외 개방형 파운드리 운영	HBM(메모리)와 로직 반도체가 결합된 제품 제공 가능	2025년 칩렛 생산 캐파 2021년 대비 2배 이상 목표
차기 출시 제품	2024년 최대 13개의 칩렛을 탑재한 차세대 GPU 개발 중	애로우 레이크 (2024) 루나 레이크 (2025)	HBM 과 로직 반도체 결합한 I-Cube12 24Q4 출시 목표	인텔의 3nm 칩렛 타일
기술 발전 방향	Substrate 활용, 2D에서 3D로 적층, RDL 및 인터포저 층 추가			
UCIe	Universal Chiplet Interconnect express: 칩렛 생태계 활성화를 위한 컨소시엄으로 칩렛에 대한 새로운 규격을 책정하기 위해 설립. 서로 다른 다이 상호 접속 규격을 표준화하여 타사 실리콘 다이를 자사 제품에 넣거나 중소 기업의 독자적인 칩렛 설계 가능			
UCIe 주요 참여 기업				



Intel Meteor Lake

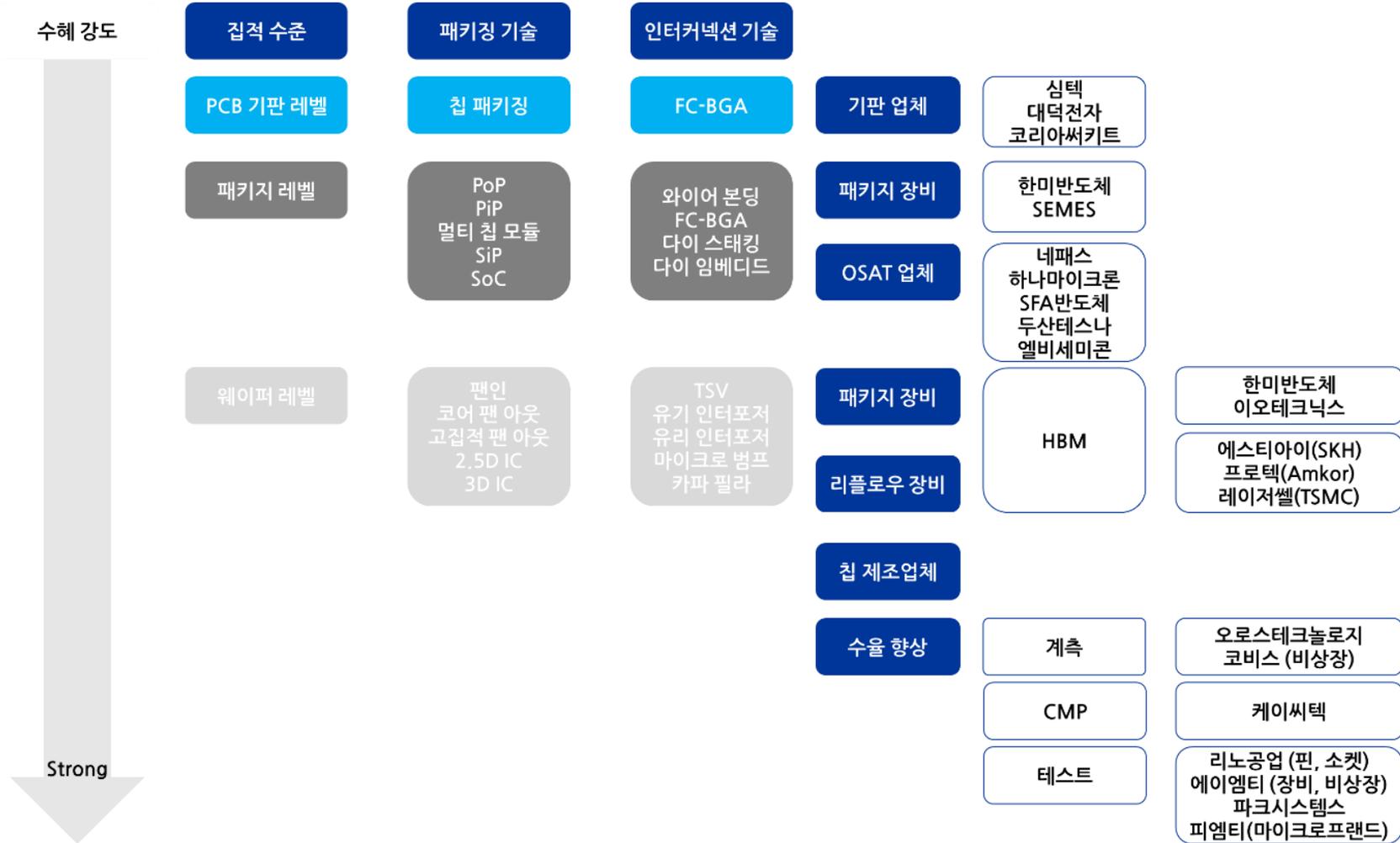


AMD MI300

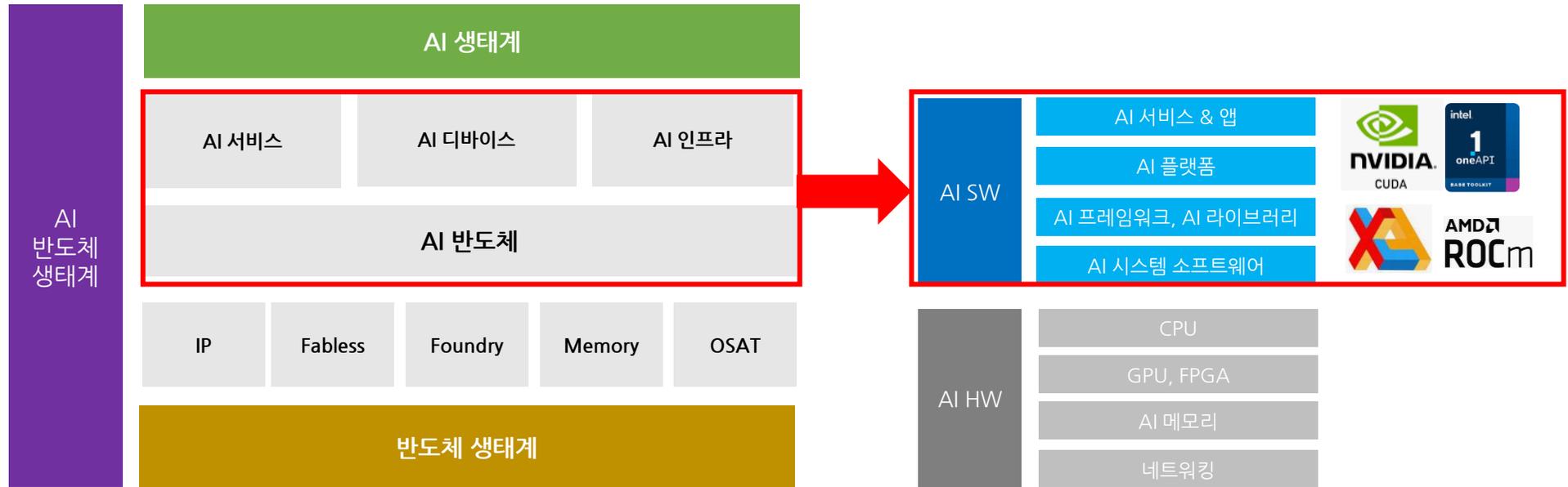
주요 반도체 기업들의 후공정 투자 계획

인텔	TSMC	삼성전자	ASE	Amkor
<ul style="list-style-type: none"> 테스트, 조립 중심으로 운영된 말레이시아 페낭 팍을 후공정 종합 거점화 목표 2032년까지 총 140억달러를 투자할 계획, 이 중 70억달러를 어드밴스드 패키징 시설에 투자 집행 사파이어 래피즈와 메테오레이크 등의 제품이 페낭에서 생산된 후, 쿨림에서 후공정 진행 내부와 외부 파운드리를 모두 활용하는 IDM 2.0 전략 추진의 일부 	<ul style="list-style-type: none"> 7월 대만 주커에 약 3.7조원 규모의 어드밴스드 패키징 공장 건설 계획 발표 엔비디아와 AMD 등 AI 반도체 수요를 대응할 수 있는 CoWoS 생산 능력 24년 말까지 월 2.8만 장으로 확대 계획 2024년까지 TSMC선단 공정 캐파의 40%를 엔비디아가 주문 어드밴스드 패키징 공장을 대만에만 이미 5개 보유하고, FO-WLP 기반의 InFO 기술로 첨단 패키징 시장을 주도하는 1위 기업 	<ul style="list-style-type: none"> 2023년 2조원 이상 패키징 라인 증설에 투자 천안 사업장 중심으로 어드밴스드 패키징 투자 예정이며, 올해 어드밴스드패키징(AVP) 사업팀 설립 일본 요코하마에도 어드밴스드 패키징 개발 거점 설치 계획 삼성전기의 기판 사업과 함께 메모리와 파운드리 사업 모두 영위하고 있는 바, 원가 경쟁력을 갖춘 턴키 솔루션을 제공할 수 있다는 강점 보유 	<ul style="list-style-type: none"> AI와 머신러닝의 도입 증가로 OSAT 업체 또한 어드밴스드 패키징 캐파 확장 노력 중 광 인터커넥트에 대한 관심으로 실리콘 포토닉스(SiPh)에 대한 R&D 진행 TSMC의 엔비디아 AI GPU 생산 병목으로 CoWoS 패키징 기술의 일부인 CoW 기술을 엔비디아로부터 인증받아 생산 예정(2024년 3월) TSMC가 영위한 후공정 일부를 고객사로부터 직접 주문 받아 공정 세분화 강화 	<ul style="list-style-type: none"> ASE와 마찬가지로 CoWoS 기술의 일부를 인증받아 엔비디아로부터 직접 주문 받는 방식으로 전환 예정 특히 엔비디아를 비롯한 TSMC의 고객사 중 일부는 2.5D 패키징 관련 공정을 Amkor에 주문 2024년 2.5D 패키징 캐파를 상반기 5,000장, 하반기 7,000장 규모까지 확대 예정 2.5D 패키징 제품 생산의 70~80%는 엔비디아가 차지할 것으로 전망

첨단 패키징 수혜 기업



AI 반도체 생태계 & 엔비디아의 강점



“많은 사람들이 AI 와 관련하여 가속 칩에 대해서만 생각하고 있지만, **사실 핵심 포인트는 스택과 소프트웨어에 있다.** 엔비디아가 구축해 놓은 네트워킹 스택인 **DOCA(Datacenter infra On Chip Architecture)** 와 가속 라이브러리인 **Magnum IO**가 바로 데이터센터 내의 수만 개 GPU를 연결할 수 있도록 하는 핵심 기술이다. **그것이 바로 엔비디아가 네트워킹에 집착하는 이유이자, 멜라녹스를 인수한 목적이다.**

엔비디아의 컴퓨팅 네트워크는 컴퓨팅 패브릭인 **NVLink**로부터 메모리를 참조해가며 확장되며 **인피니밴드**에 연결된다. 이러한 **최적화된 광섬유의 네트워크가 바로 엔비디아가 고객들에게 제공하는 인공지능 시스템이다.** 엔비디아는 전세계 모든 데이터센터에 이 같은 기술들을 각 컴포넌트로 구분해 제공한다. 그렇게 해야 모든 스타일의 아키텍처 위에 엔비디아의 소프트웨어 스택이 제대로 얹혀질 수 있기 때문이다. 이 과정은 놀라울 정도로 복잡하지만, 이를 통해 **엔비디아의 AI 솔루션은 경쟁하기 어려울 정도의 압도적인 확장성을 갖는다.”**

2023년 5월 엔비디아 실적발표 컨퍼런스콜에서...

The End of Presentation

감사합니다