BLUE LAB | ADVANCED INQUIRY 01

반도체의 발전사와 현재, 그리고 미래

더 작게 만드는 기술에서, 시스템을 함께 설계하는 기술로

물리 · 화학 · 정보 · 공학 · 수학을 연결하는 심화 탐구

탐구를 이끄는 질문

**반도체는 어떤 한계를 해결하며 발전했고,
앞으로는 무엇이 성능 향상을 이끌까?**

01

발전사

소자의 발명과 집적화는
어떤 문제를 해결했는가?

02

현재

전력·메모리·연결이
왜 중요한가?

03

미래

새 구조·소재·계산 방식은
어디까지 검증되었는가?

이 보고서의 결론

반도체의 발전은 크기 축소와 함께 **전류 제어, 전력 관리, 데이터 이동, 제조 가능성**을 개선해 온 과정이다. 미래 기술도 성능·전력·면적·비용·신뢰성을 함께 비교해야 한다.

구성 원리 3-4쪽 / 발전사 5-7쪽 / 현재 8-10쪽 / 계산 탐구 11-12쪽
미래 13-15쪽 / 탐구 설계·결론 16-17쪽 / 참고 자료 18-20쪽

대상: 고등학생 · 자료 확인 기준: 2026.10.31 · 문헌 분석 + 가상 계산

연표를 넘어, 해결한 문제를 추적한다

같은 기술을 물리 원리·제조 조건·실제 활용이라는 세 관점으로 읽는다.

이 보고서는 반도체의 역사를 **"새로운 기술이 어떤 제약을 줄였고, 그 뒤에 어떤 제약이 드러났는가"**로 재구성한다. 문헌에서 확인한 사실에 가상 수치 계산을 더해, 미래 발전 방향에 대한 주장을 검토한다.

검토할 가설	확인 방법
집적화는 크기뿐 아니라 회로의 연결·제조 문제를 바꾸었다.	트랜지스터, 집적회로, MOS·CMOS의 원리와 기술사 비교
소자가 작아져도 전압과 데이터 이동을 함께 개선하지 않으면 시스템 이득이 제한된다.	동적전력 비율과 메모리 대역폭에 따른 성능 상한 계산
미래 성능 향상에는 소자·연결·계산 방식의 공동 개선이 필요하다.	각 기술이 겨냥한 문제와 확인된 개발 단계 비교

"발전"을 평가하는 기준

PPA는 성능(Performance), 전력(Power), 면적(Area)을 함께 보는 틀이다. 탐구에서는 여기에 **비용·수율·신뢰성**을 더한다. 수율은 정해진 기준을 만족하는 제품의 비율이며, 시험용 소자의 동작 비율과 완성 칩의 양품 비율은 구분한다.

자료의 단계	읽는 방법
연구 시연	특정 소자·회로가 정해진 조건에서 동작했는가?
공정 통합·시험 생산	웨이퍼와 복잡한 회로로 확대해도 재현되는가?
양산·상용 출하	실제 생산·출하 또는 출시를 명시한 자료가 있는가?
로드맵	목표 시점·구조를 제시한 계획이며 달성 여부는 따로 확인한다.

자료 기준일은 **2026년 10월 31일**이다. 제조사의 공식 자료는 해당 기업의 발표 범위를 확인하는 데 사용하며, 기업 간 절대적 우열을 판정하는 자료로 사용하지 않는다. 계산 예시는 실측 실험이 아닌 교육용 모형이다.

반도체의 핵심은 전하의 움직임을 조절하는 것

전기가 “절반만 통하는 물질”이라는 설명으로는 소자의 동작을 이해하기 어렵다.

실리콘 결정에서는 원자들이 전자를 공유해 결합한다. 전자가 차지할 수 있는 에너지 상태 중 가전자대와 전도대 사이에는 **밴드갭**이 있다. 열·빛·도핑 등의 조건은 전류를 운반하는 입자의 수와 상태에 영향을 준다. [1]

개념	물리적 의미	소자와의 연결
전자	전도대 등에서 이동하며 전류에 기여하는 음전하 운반자	n형 영역의 다수 반송자
정공	가전자대에서 전자가 비어 있는 상태를 양전하 운반자로 기술한 것	p형 영역의 다수 반송자
도핑	결정에 소량의 원소를 넣어 반송자 농도를 조절하는 방법	전류가 흐르는 조건을 설계

실리콘에 다른 원소를 넣으면

구분	대표 도핑 원소	주된 변화
n형	인(P) 등	쉽게 이동할 전자를 공급한다.
p형	붕소(B) 등	전자 빈자리인 정공을 늘린다.

n형과 p형이라는 이름은 **다수 반송자의 종류**를 가리킨다. 벌크 물질은 보통 반송자와 이온화된 도핑 원자의 전하가 균형을 이루므로, n형 전체가 음전하이거나 p형 전체가 양전하라는 뜻이 아니다. [1]

> **소자 원리로 이어지는 질문**
>
> 전하 운반자를 늘리는 것에서 한 걸음 더 나아가, **외부 전압으로 전류가 흐르는 길을 만들거나 약화할 수 있을까?** 이 질문이 다음 쪽의 MOSFET 동작으로 이어진다.

학습 연결: 화학의 공유 결합·원자가 전자, 물리의 전기장·전류·에너지 준위. 반송자 농도뿐 아니라 얼마나 쉽게 움직이는지인 이동도도 전도도에 영향을 준다. [1]

MOSFET은 전기장으로 통로를 조절한다

게이트에 가한 전압이 소스와 드레인 사이의 전류를 제어한다.

MOSFET의 게이트와 반도체 사이에는 절연층이 있다. 아래 nMOS 예에서는 게이트 전압이 표면의 전하 분포를 바꾸어 전자 통로를 형성한다. 게이트 전류를 통로에 직접 공급해 켜는 방식과는 다르다. [2]

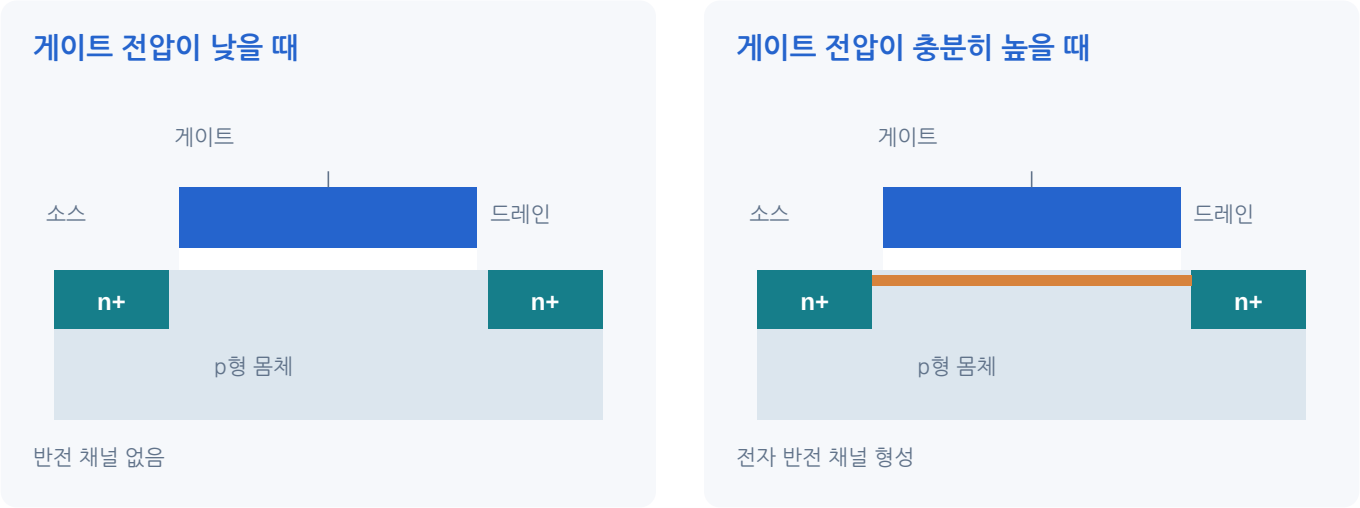

그림 1. nMOS의 동작을 단순화한 길이 방향 단면. 흰 띠는 게이트 절연층, 주황색은 전자가 모인 반전 채널이다. 소스와 드레인 사이에 전압이 걸려 있다고 가정한다. 실제 비율이 아니며 몸체 접점·배선은 생략했다. [2]

CMOS: nMOS와 pMOS를 상보적으로 사용한다

가장 간단한 예인 인버터는 입력이 낮으면 출력을 높게, 입력이 높으면 출력을 낮게 만든다. 이상적인 안정 상태에서는 두 소자 중 하나가 꺼져 있어 전원과 접지 사이의 지속 전류를 막는다. [3]

전력이 생기는 경로	이 탐구에서의 의미
전환할 때의 충전·방전	배선과 다음 회로의 정전용량을 충전하는 에너지가 필요하다.
누설·전환 중 전류	실제 소자는 완벽히 절연되거나 즉시 전환되지 않는다. 총전력은 단순 모형보다 복잡하다.

이 원리가 왜 중요한가

트랜지스터 수가 늘어나도 전력을 관리하려면 **얼마나 자주, 얼마나 큰 전압으로, 얼마나 큰 부하를 충전하는지**를 함께 봐야 한다. 11쪽에서 이 관계를 계산한다.

작은 소자에서 복잡한 회로로

각 사건은 다음 기술을 가능하게 한 조건을 바꾸었다.

시기·사건	확인한 사실	해결한 문제에 대한 해석
1947 트랜지스터	벨 연구소의 바딘·브래튼이 게르마늄 점접촉 소자의 증폭 작용을 구현했다. [4]	고체 소자를 이용하는 전자 회로의 길을 열었다. 이후 접합형 등 다른 구조로 확장되었다.
1958 킬비의 회로	반도체 재료로 회로 요소를 구현한 작동 회로를 시연했다. 요소 사이에는 가는 연결 도선이 남아 있었다. [5]	부품을 따로 조립하는 방식에서 한 기판에 집적하는 방식으로 질문이 옮겨 갔다.
1959 노이스의 IC 개념	평면 공정 위에 산화막 절연과 금속 연결을 결합하는 개념을 제안했다. [6]	소자를 만드는 일과 서로 연결하는 일을 제조 공정 안에서 다룰 기반을 마련했다.
1959-1960 MOS 트랜지스터	아탈라·강대원이 1959년 소자를 만들고 1960년 시연했다. 표면·산화막 제어가 중요했다. [7]	절연 게이트로 전류를 제어하는 구조가 집적 회로의 핵심 기반으로 발전했다.
1963 CMOS	완래스와 사가 p채널·n채널 MOS를 결합한 상보형 회로를 발표했다. [8]	정지 상태의 전력 소모를 줄이는 회로 구성이 가능해졌다.
1971 마이크로프로세서	Intel 4004는 CPU 기능을 한 칩에 모은 초기 상용 사례다. [9]	계산 기능을 집적한 칩을 여러 제품의 시스템 설계에 활용할 수 있게 되었다.

IC는 집적회로, CPU는 중앙처리장치다. CPU를 한 칩에 넣었다고 메모리와 입출력까지 갖춘 완전한 컴퓨터가 한 칩에 들어간 것은 아니다. 또한 최초 발명·공개 시연·특허·상용화는 서로 다른 사건이다. [9]

역사에서 읽을 수 있는 연결

물질과 표면을 제어하는 기술 → 소자를 안정적으로 만드는 공정 → 여러 소자를 연결하는 집적회로 → 복잡한 계산 기능. 발명 하나가 즉시 산업 전체를 바꾸었다고 보기보다 이 연결을 추적한다.

무어의 법칙과 데너드 스케일링은 다르다

집적도의 증가와 전력 밀도에 관한 조건을 구분한다.

구분	무어의 관찰·전망	데너드 스케일링
핵심 질문	경제적으로 집적할 수 있는 소자 수가 어떻게 늘어나는가?	소자의 치수와 전압을 함께 줄이면 전기적 특성이 어떻게 변하는가?
성격	산업의 경험적 추세와 전망. 1965년과 1975년 원문의 증가 주기도 다르다. [10]	이상적 조건을 둔 소자 축소 이론. 1974년 논문이 대표적이다. [11]
흔한 오해	컴퓨터의 실제 성능이 언제나 2년마다 정확히 2배가 된다는 뜻으로 읽기	크기만 줄여도 총전력이 자동으로 낮아진다고 읽기

무어는 1965년 원문에서 집적 부품 수가 대략 매년 두 배로 늘어나는 추세를 논의했다. 1975년에는 이후 약 2년 주기로 바뀔 가능성을 제시했다. 오늘날 자주 쓰는 “2년마다 두 배”를 1965년의 원문과 그대로 동일시하지 않는다. [10]

축소가 여러 이득을 함께 주던 조건

이상적인 일정 전기장 축소에서는 길이와 전압을 함께 줄인다. 더 많은 소자를 같은 면적에 배치하고 소자당 에너지를 낮출 가능성이 생긴다. 핵심은 **치수 축소와 전압 축소가 함께 진행된다는 조건**이다. [11]

전압은 왜 끝없이 같은 비율로 줄지 못했을까?

전압을 낮추면 스위칭 속도와 신호 구분에 제약이 생기고, 꺼진 상태의 누설도 관리해야 한다. 2000년대 중반 무렵에는 이상적 스케일링이 주던 전력 밀도 이득을 유지하기 어려워졌다. [3][11]

> **새로운 질문**
>
> “트랜지스터를 더 넣을 수 있는가?”와 함께 **“허용 전력 안에서 얼마나 유용하게 작동시킬 수 있는가?”**를 물어야 한다. 연산 구조, 메모리 연결, 소프트웨어까지 살펴볼 이유가 여기에 있다.

탐구자의 해석: 축소의 한계는 반도체의 발전이 끝났다는 뜻보다, 성능 향상을 설명하는 변수가 늘어났다는 뜻으로 이해하는 편이 타당하다. 11-12쪽의 계산으로 이 해석을 검토한다.

미세한 회로를 반복해서 만드는 기술

재료를 놓고, 패턴을 만들고, 필요한 부분을 바꾸는 과정을 거듭한다.

과정	하는 일	주의할 점
설계·검증	회로의 기능과 배치를 정하고 제조 조건에 맞춰 검토한다.	좋은 소자 하나만으로 복잡한 칩이 완성되지는 않는다.
막 형성·패턴 정의	재료 층과 감광막을 형성하고 노광으로 패턴을 정의한다.	빛이 회로 전체를 한 번에 완성하는 것은 아니다.
식각·도핑·배선	선택한 부분을 제거하거나 성질을 바꾸고 연결을 만든다.	여러 층과 목적에 맞춰 공정을 반복한다.
검사·절단·패키징	동작을 검사하고 외부 연결·보호·열 관리가 가능한 형태로 조립한다.	패키징도 성능과 제조 비용에 영향을 준다.

위 순서는 이해를 위한 묶음이다. 실제 공정은 수많은 반복 단계와 중간 검사로 구성되며, 공정 종류와 제품에 따라 순서가 달라진다. [12]

EUV의 13.5 nm와 “2nm 공정”의 2는 서로 다른 수치다

EUV 노광은 파장 약 **13.5 nm**의 빛을 사용한다. 구현 가능한 패턴은 파장 외에도 광학계와 공정 조건에 영향을 받는다. 모든 층을 EUV로 만드는 것도 아니다. [12][13]

표현	해석
파장	빛의 성질을 나타내는 물리량
공정 노드 이름	기술 세대를 구분하는 명칭. 오늘날 그 숫자가 모든 게이트 길이·배선 폭과 같지는 않다. [14]

> **비교의 기준**
>
> 공정 이름의 숫자만으로 회사나 제품의 성능을 순위 매기지 않는다. **같은 작업·정밀도·전력 조건의 성능, 밀도, 수율과 비용**을 확인해야 비교가 성립한다.

반도체는 서로 다른 일을 나누어 맡는다

“가장 작은 공정” 하나로 모든 제품의 가치를 설명할 수 없다.

종류	주된 역할	관찰할 지표
CPU	다양한 명령과 제어 흐름을 처리한다.	작업 처리 시간, 응답성, 전력
GPU·AI 가속기	많은 연산을 병렬로 수행하도록 설계한다.	정밀도별 처리량, 활용률, 메모리 제약
DRAM	작업 중의 데이터를 보관한다. 전원을 끄면 정보가 유지되지 않는다.	용량, 대역폭, 지연시간, 전력
NAND 플래시	전원이 꺼져도 정보를 보관하는 저장장치의 기반이다.	용량, 읽기·쓰기 특성, 수명
전력 반도체	전기에너지의 흐름과 전압·전류 변환을 제어한다.	손실, 내전압, 스위칭, 열·신뢰성

역할에 따른 개념 구분이다. 하나의 제품에 CPU·GPU·메모리·입출력 기능이 함께 포함될 수 있으며, 각 지표의 우선순위는 사용 목적에 따라 달라진다. 메모리 구조는 10쪽, 전력용 소재는 15쪽에서 다룬다. [18][19][26]

누가 설계하고, 누가 만드는가?

역할	의미
팹리스	주로 반도체를 설계하고 제조를 외부에 맡기는 기업
파운드리	다른 기업의 설계를 받아 제조하는 사업
IDM	설계와 제조를 함께 수행하는 종합 반도체 기업

설계 도구·소재·장비·검사·패키징 기업도 연결된다. 역할은 사업에 따라 겹칠 수 있다. 첨단 공정과 성숙 공정은 서로 다른 요구를 충족하며 공존한다. [12][15]

게이트가 채널을 더 잘 제어하게 만든다

평면형에서 입체 구조로: 크기 축소로 약해지는 전류 제어를 보완한다.

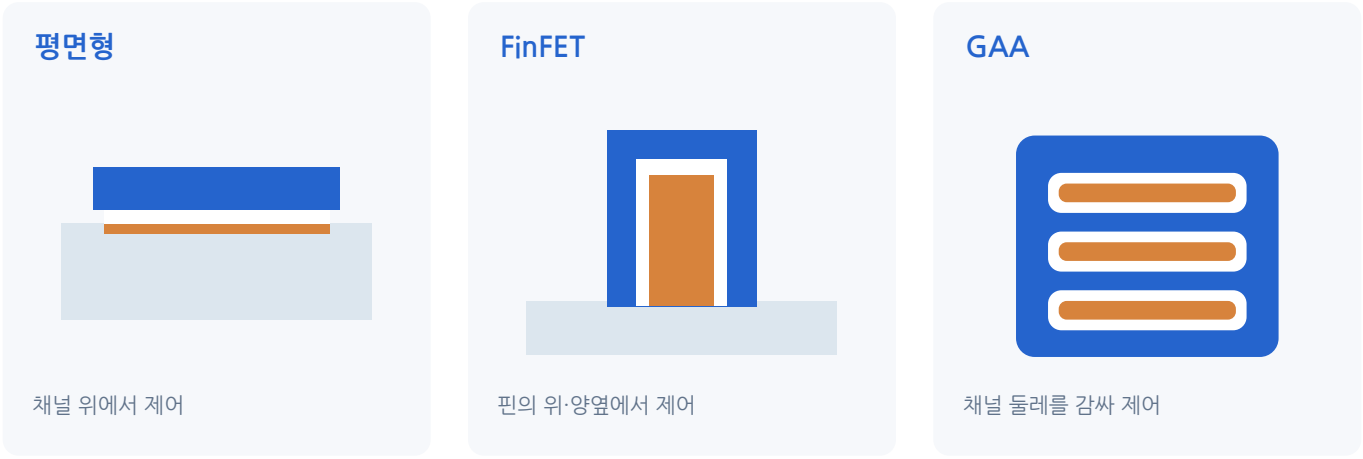

그림 2. 전류가 흐르는 방향에 수직인 개념 단면. 파랑은 게이트, 주황은 채널 영역, 흰 간격은 절연층이다. GAA의 여러 나노시트는 한 트랜지스터의 채널일 수 있으며, nMOS와 pMOS를 쌓는 CFET와 다르다. [16][22]

FinFET은 돌출된 핀의 여러 면에서, GAA는 채널의 둘레에서 게이트가 작용한다. 이 변화는 짧아진 채널의 전위를 더 잘 제어하기 위한 방향이다. 구조가 복잡해지는 만큼 정밀한 제조와 접촉·열 관리도 필요하다. [16][22]

2026.10.31까지 확인한 사례	자료로 확인한 범위
TSMC N2	2025년 4분기에 양산을 시작했다고 명시한다. 나노시트 트랜지스터를 적용한 공정이다. [15]
Intel 18A	RibbonFET(GAA)와 PowerVia(후면 전력 공급)를 결합한다. CPU에 18A를 사용한 386H 제품은 2026년 1분기 출시 상태로 표시된다. [16]

전원 배선도 설계 대상이다

후면 전력 공급은 전력 배선을 칩 뒷면 쪽으로 옮겨 신호 배선의 혼잡과 전압 강하를 줄이려는 방식이다. 이미 적용된 사례가 있으므로, 이를 모두 먼 미래의 연구로 분류하면 현재 기술을 잘못 설명하게 된다. [16]

회사마다 공정 도입 시점과 구조가 다르다. 이 표는 상용화 근거가 있는 대표 사례이며, 업계 전체의 동일한 전환이나 모든 제품의 같은 성능을 뜻하지 않는다.

데이터를 어디에 놓고 어떻게 옮길까

많은 연산을 하려면 필요한 데이터를 충분히 공급해야 한다.

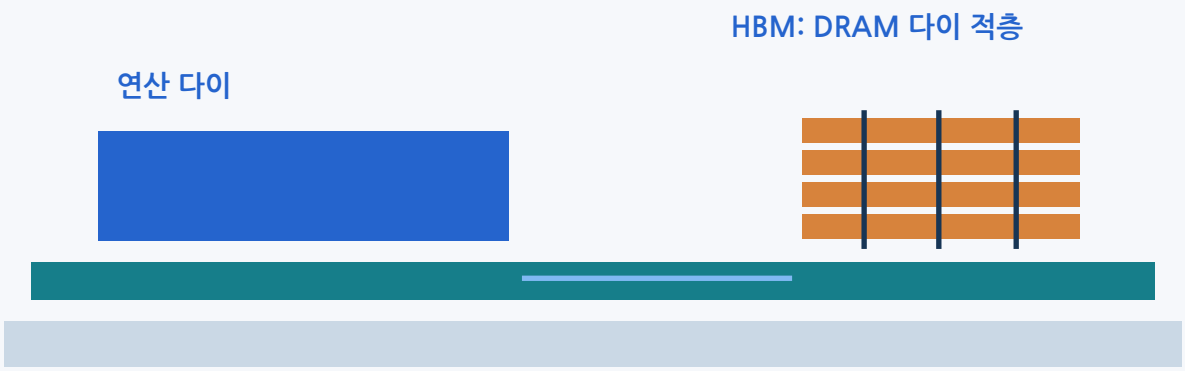

그림 3. 연산 칩과 HBM을 인터포저 위에 배치하는 개념도. HBM 안의 세로선은 TSV 연결을 뜻한다. 실제 다이 수·크기·배선·베이스 다이는 제품마다 다르며 일부 구조를 생략했다. [18][20]

기술	무엇을 쌓거나 연결하는가?	주된 목적
HBM	여러 DRAM 다이를 쌓고 TSV 등으로 연결한다. [18]	넓은 인터페이스를 통한 높은 데이터 대역폭
3D NAND	메모리 셀을 수직 방향의 여러 층에 배치한다. [19]	저장 밀도 확대와 평면 축소의 제약 완화
칩렛·첨단 패키징	기능을 나눈 다이들을 가깝고 촘촘하게 연결한다. [20]	기능·공정을 조합하고 다이 간 연결을 개선

HBM과 3D NAND는 모두 “쌓는다”는 표현을 쓰지만 구조와 역할이 다르다. 인터포저 위에 다이를 나란히 연결하는 2.5D 패키징과, 다이를 수직으로 연결하는 3D 적층도 구분해야 한다. [18][19][20]

> **현재 단계 확인: HBM4**
>
> 삼성전자는 **2026년 2월 12일** HBM4의 양산과 상용 고객 출하를 발표했다. 따라서 HBM4를 연구 아이디어로만 소개하지 않는다. 이 발표만으로 모든 고객의 실제 활용률이나 시스템 성능까지 알 수 있는 것은 아니다. [17]

용량은 저장할 수 있는 양, **대역폭**은 단위 시간당 옮기는 양, **지연시간**은 요청 후 응답까지 걸리는 시간이다. 한 지표가 좋아져도 다른 지표가 같은 비율로 개선되지는 않는다.

작아졌는데 전력은 왜 그대로일까?

전압을 함께 낮출 수 있는지가 동적전력의 이득을 바꾼다.

$$P_{dyn} \approx a\,C\,V^2\,f$$

a는 클록당 평균 0→1 충전 전이 수, C는 비교하는 회로의 부하 정전용량, V는 전원 전압, f는 클록 주파수다. 충전·방전의 단순 모형이며, 아래에서는 a가 같다고 가정한다. [3]

가상 설계	C/C_0	V/V_0	f/f_0	P_{dyn}/P_0
기준	1.00	1.00	1.00	1.00
A: 전압도 축소	0.80	0.80	1.25	0.64
B: 전압 유지	0.80	1.00	1.25	1.00
C: 주파수 더 증가	0.80	1.00	1.50	1.20

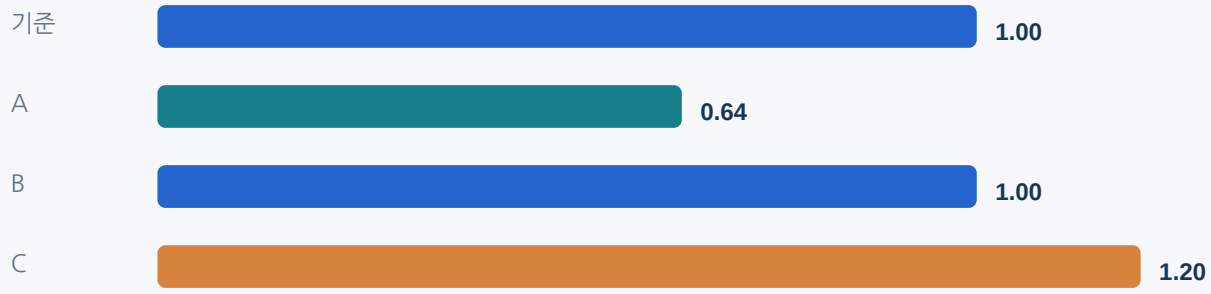

그림 4. 표의 조건을 계산한 상대 동적전력. 기준 설계를 1로 둔다. 실제 반도체의 실측값이 아니다.

A의 계산: $0.8 \times 0.8^2 \times 1.25 = 0.64$ → 동적전력 36% 감소.
B의 계산: $0.8 \times 1^2 \times 1.25 = 1$ → 정전용량 이득을 주파수 증가가 상쇄한다.
C의 계산: $0.8 \times 1^2 \times 1.5 = 1.2$ → 동적전력이 20% 증가한다.

해석의 범위

각 조건에서 동작 가능한 **별도 설계**가 있다고 가정했다. 임의로 전압을 낮춰도 같은 주파수로 동작한다는 보장은 없다. 누설 전력·전환 중 단락 전력과 메모리·입출력 전력은 포함하지 않았으며, 소자 수가 늘면 총 C와 활동량도 달라질 수 있다.

확장 질문: 처리 시간이 달라지면 에너지 E=P×t는 어떻게 달라질까? 전력과 작업당 에너지를 구분하고, 주파수 증가가 실행 시간을 같은 비율로 줄인다는 가정을 먼저 검토한다.

계산 능력이 2배면 작업도 2배 빨라질까?

Roofline 모형으로 연산과 데이터 공급 중 무엇이 상한을 정하는지 비교한다.

$R_{actual} \le \min(R_{peak}, B \times I)$

R_actual은 실제 처리량, R_peak는 최대 연산 처리량이다. B는 메모리 대역폭, I는 옮긴 데이터 1 byte당 유효 연산 수다. FLOP는 부동소수점 연산 1회다. B와 I는 **같은 메모리 계층**을 기준으로 정하고, 연산 정밀도도 맞춘다. [21]

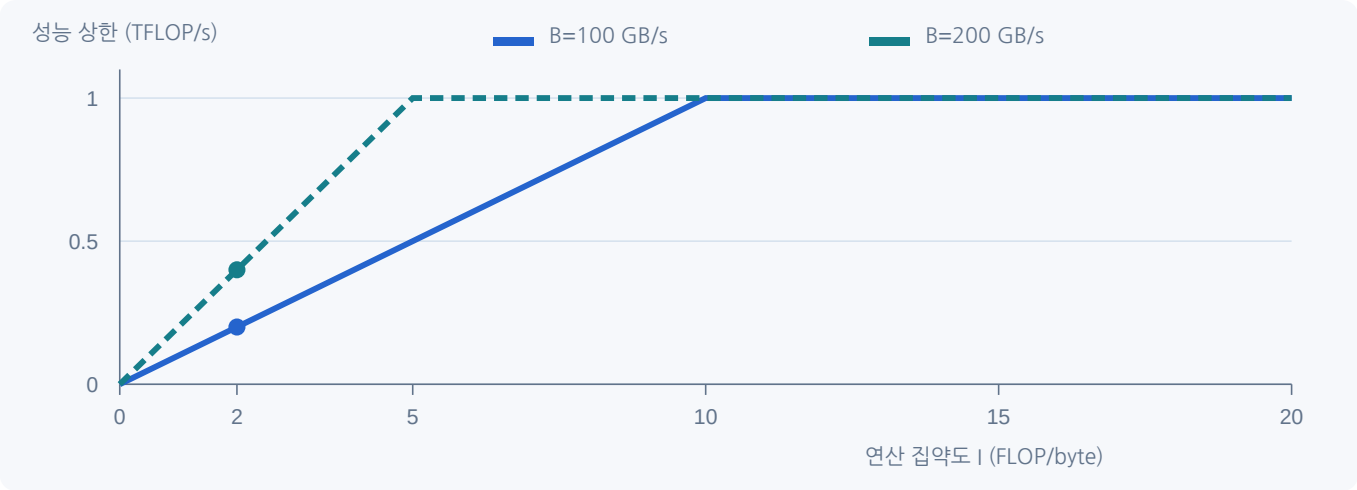

그림 5. 계산 상한 R_peak=1 TFLOP/s인 가상 장치. I가 작으면 대역폭이, 충분히 크면 연산 능력이 상한을 정한다. 선은 달성된 실측 성능이 아니다. [21]

가상 조건	연산 상한	B×I	전체 상한
기준: B=100 GB/s, I=2	1.0	0.2	0.2
연산 능력만 2배	2.0	0.2	0.2
대역폭만 2배	1.0	0.4	0.4
재사용 개선: I=10	1.0	1.0	1.0

표의 처리량 단위는 TFLOP/s다. 100 GB/s × 2 FLOP/byte = 200 GFLOP/s = 0.2 TFLOP/s이다. G=10^9, T=10^{12}로 계산했다. 모든 값은 실제 제품과 무관한 가상값이다.

결론: 데이터 공급이 병목이면 연산 장치만 강화해도 상한은 그대로다. 같은 유효 작업에서 데이터를 재사용하면 필요한 전송량을 줄여 I를 높일 수 있다. 의미 없는 연산을 더하는 것과 다르다.

실제 성능은 지연시간·병렬성·프로그램 구현 등에 의해 이 상한보다 낮을 수 있다. 대역폭 2배가 항상 실제 속도 2배를 보장하는 것도 아니다. [21]

더 얇은 채널, 더 촘촘한 배치

가능한 구조와 제조 가능한 제품 사이에는 여러 검증 단계가 있다.

CFET: nMOS와 pMOS를 위아래로 배치

CMOS 회로에 필요한 n형·p형 트랜지스터를 수직으로 쌓아 회로가 차지하는 바닥 면적을 줄이려는 접근이다. GAA 소자 안에 같은 종류의 채널 시트를 여러 장 넣는 것과 구분한다. imec의 2026년 9월 자료는 접점과 게이트 통합 등 공정 연구를 다룬다. [22]

2차원 소재 채널: 원자층 수준의 얇은 통로

MoS_2 같은 2차원 반도체는 매우 얇은 채널에서도 전기장으로 전류를 제어할 가능성을 제공한다. imec·ASML·TSMC는 2026년 6월 **지름 300 mm 웨이퍼**에서 2D 소재 nFET·pFET의 공정 통합 시연을 발표했다. 이것은 완성된 CPU의 양산 발표와 다른 단계다. [23]

접근	줄이려는 제약	검토할 과제
CFET	n/p 소자를 옆으로 놓을 때의 면적	아래 소자의 접점, 여러 공정의 간섭, 배선·열·수율
2D 소재 채널	매우 짧고 얇은 채널에서의 전류 제어	균일한 재료 성장·전사, 접촉 저항, 절연막 통합, 재현성

이 표의 과제는 보고서가 자료를 바탕으로 정리한 검토 항목이다. 일부 특성의 기록을 세운 것과 대량 생산에서 모든 요구 조건을 동시에 만족하는 것은 구분해야 한다. [22][23]

논문·보도자료를 읽는 질문

① 소자 하나, 시험 회로, 완성 칩 중 무엇인가?
② 어떤 조건에서 무엇을 측정했는가?
③ 웨이퍼 크기와 실제 제품의 양산·수율을 구분했는가?
④ 기존 기술보다 유리한 점과 남은 과제를 함께 제시했는가?

데이터 이동을 바꾸는 두 방향

이동 경로를 개선하거나, 이동해야 할 양을 줄인다.

광연결: 빛으로 데이터를 전달하는 구간을 확대

광연결은 전기 신호를 빛으로 바꾸어 전달하고 다시 전기 신호로 읽는 방식을 이용한다. 광소자와 전자 칩을 가까이 패키징하는 CPO도 이 방향이다. NVIDIA는 2026년 5월 Spectrum-X Ethernet Photonics의 생산을 발표했다. **일부 스위치의 상용 기술**과 다른 적용처의 연구를 구분해야 한다. [24]

이 사례가 모든 CPU·GPU 내부 연산을 빛으로 대체했다는 뜻은 아니다. 비교할 때는 광원, 전기·광 변환, 패키징과 열 관리까지 포함한 연결 전체의 이득을 따져야 한다. 이는 이 탐구에서 제안하는 시스템 평가 기준이다.

메모리 내부 연산: 저장 위치 가까이에서 계산

데이터를 매번 멀리 옮기는 대신 메모리 안에서 일부 계산을 수행하려는 접근이다. IBM 연구진의 2023년 논문은 상변화 메모리를 이용한 64개 아날로그 연산 코어와 디지털 처리·통신을 결합해 특정 신경망 추론 과제를 시연했다. [25]

비교 질문	검토해야 하는 이유
정확도를 같은 수준으로 맞췄는가?	처리량이 커져도 필요한 결과의 품질을 충족해야 한다.
주변 회로와 변환 비용을 포함했는가?	아날로그 연산의 장점만으로 전체 에너지를 판단할 수 없다.
어떤 작업에 유리한가?	특정 추론 실험을 범용 CPU나 모든 AI 작업으로 일반화할 수 없다.

두 계산 모형과 연결하기

11쪽은 **스위칭에 드는 전력**, 12쪽은 **데이터 공급이 정하는 성능 상한**을 보여 주었다. 미래 기술은 이 제약의 일부를 줄일 수 있지만, 다른 위치의 비용과 제약까지 없애 주는 것은 아니다.

미래는 용도에 따라 여러 갈래로 진행된다

같은 기술도 적용할 문제에 따라 가치가 달라진다.

사용 목적	우선 관찰할 조건	비교할 기술 방향
배터리 기반 기기	작업당 에너지, 대기 전력, 열, 필요한 응답 속도	저전력 회로와 작업 배치, 메모리 접근 감소
대규모 연산 시스템	계산·메모리·네트워크의 균형, 전체 전력	HBM·패키징·연결 개선, 작업에 맞는 가속
전력 변환 시스템	전압·전류·주파수, 손실과 신뢰성	Si·SiC·GaN 등 소재와 소자 구조 선택
센서·제어 장치	요구 성능, 가격, 수명, 안정적 공급	목적에 맞는 공정과 주변 회로의 조합

위 표는 기술 순위가 아니라, 목적에 따라 비교 지표를 정하는 탐구 틀이다. 최대 연산량이나 가장 작은 노드만으로 네 종류의 제품을 한 줄에 세울 수는 없다.

SiC·GaN은 이미 활용되는 전력용 소재다

탄화규소(SiC)와 질화갈륨(GaN)은 넓은 밴드갭을 갖는 반도체로, 전력 변환 분야에서 적용 조건에 따라 장점을 갖는다. 전력 정격과 스위칭 주파수 등 요구 조건을 맞춰 선택해야 한다. CPU의 실리콘을 모두 대체하는 단일한 다음 단계로 볼 수는 없다. [26]

효율이 좋아지면 전체 사용 에너지도 줄어들까?

토론용 가상 상황: 작업 1회당 에너지가 20% 감소했지만 작업 횟수가 2배가 되면, 다른 조건이 같을 때 연산 에너지는 0.8 × 2 = **1.6배**다. 이것은 미래 수요 예측이 아니라 단위당 효율과 총량을 구분하는 계산이다.

> **미래 기술의 평가 문장**
>
> "이 기술은 **어떤 작업**에서 **무슨 제약**을 줄이며, **무엇까지 검증**되었고, **어떤 비용과 조건**이 남아 있는가?" 이 네 항목을 채우면 전망의 근거와 한계가 함께 드러난다.

문헌 조사를 검증 가능한 탐구로 바꾸기

제안 수업: 4차시. 계산기 또는 표 계산 도구로 진행할 수 있다.

차시	수행할 일	남길 결과물
1 · 원리·역사	5쪽의 사건 3개를 골라 이전 제약과 변화한 조건을 비교한다.	사건·근거·해결한 문제·남은 제약 표
2 · 전력	11쪽을 재계산하고 C, V, f 중 하나씩 바꾼다. 나머지는 통제한다.	조건표, 상대전력 그래프, 해석
3 · 메모리	12쪽에서 I=1, 2, 5, 10, 20을 대입해 상한을 비교한다.	어떤 범위에서 대역폭이 병목인지 설명
4 · 미래 평가	CFET·2D 채널·광연결·메모리 내부 연산 중 하나를 선정한다.	근거 자료와 개발 단계를 포함한 결론

보고서에 사용할 연구 질문

주질문: 미세화가 성능 향상으로 이어지려면 어떤 조건이 함께 개선되어야 하는가?
하위 질문: 전압 축소 여부는 전력 이득을 얼마나 바꾸는가? 메모리 대역폭은 어떤 작업에서 상한을 정하는가? 선택한 미래 기술은 이 제약 중 무엇을 겨냥하는가?

비교 기준을 바꾸어 반례 찾기

① “이 모형에서 대역폭을 늘리면 성능 상한이 언제나 높아진다”는 주장을 I=20에서 검토한다.
② “C가 작아지면 동적전력이 항상 줄어든다”는 주장을 11쪽 B·C로 검토한다.
③ 연구 발표에서 “양산된다”라고 결론 내리려면 어떤 추가 자료가 필요한지 적는다.

해석 예시와 점검 기준

기준 장치에서 I=20이면 BI=2 TFLOP/s지만 연산 상한이 1이므로 전체 상한은 1이다. 이때 대역폭만 2배로 해도 상한은 그대로다.

점검: 단위와 계산이 맞는가? 사실·가정·해석이 구분되는가? 원문이 주장 범위를 뒷받침하는가? 모형이 설명하지 못하는 조건을 밝혔는가?

반도체의 미래는 시스템 전체의 문제다

기술의 이름보다, 줄이는 제약과 함께 늘어나는 과제를 본다.

탐구 결과를 연결하면

발전사: 트랜지스터와 집적회로, MOS·CMOS는 전류 제어와 회로 통합의 조건을 바꾸었다. 미세화는 더 많은 기능을 구현할 기반이 되었지만, 전압·전력의 이득이 항상 같은 속도로 따라오지는 않았다. [4]-[11]

현재: GAA·후면 전력 공급·HBM·첨단 패키징은 소자와 데이터 연결을 함께 개선하는 방향을 보여 준다. N2 양산, 18A 적용 제품, HBM4 출하처럼 시점이 명시된 근거로 현재 단계를 판단할 수 있다. [15]-[20]

미래: CFET·2D 채널·광연결·메모리 내부 연산은 서로 다른 제약을 겨냥한다. 일부는 이미 상용 적용이 있으며, 일부는 연구 시연 단계다. 이들을 한 묶음으로 “곧 모든 반도체를 대체할 기술”이라 부르는 것은 근거를 벗어난다. [22]-[26]

> **학생 보고서용 결론 예시**
>
> 본 탐구의 가상 계산에서는 정전용량 감소만으로 동적전력 감소가 보장되지 않았으며, 데이터 공급이 부족한 작업에서는 연산 능력을 늘려도 성능 상한이 변하지 않았다. 따라서 반도체의 발전을 설명하려면 소자 크기와 함께 전압, 메모리, 연결 구조, 작업의 특성을 고려해야 한다. 미래 기술의 가치는 이러한 제약을 실제 제조와 사용 조건에서 얼마나 줄이는지로 평가할 수 있다.

끝까지 구분해야 할 다섯 쌍

혼동하기 쉬운 표현	정확하게 구분하기
집적도 / 실제 성능	더 많은 소자와 더 빠른 작업은 같은 지표가 아니다.
전력 / 에너지	전력은 소비 속도, 에너지는 시간 동안 소비한 총량이다.
메모리 용량 / 대역폭	저장할 양과 시간당 옮길 양을 구분한다.
GAA 채널 적층 / CFET	한 소자의 여러 채널과 n/p 소자의 수직 배치를 구분한다.
연구 시연 / 양산	시험 조건의 성공과 대량 생산·상용 출하를 구분한다.

범위와 한계: 공개 문헌과 단순 모형을 사용한 탐구다. 독자적인 반도체 제작·실측이나 특정 기업의 수율·원가 비교를 수행한 결과는 아니다.

출처와 확인한 내용

본문의 [번호]를 누르면 해당 항목으로 이동한다. 링크는 원문 또는 원문을 제공하는 공식 페이지다.

[1] MIT OpenCourseWare · Clifton Fonstad (2009). 6.012 Lecture 1: Semiconductors.

실리콘 결합, 전자·정공, 도핑과 반송자 개념.

강의 자료 · https://ocw.mit.edu/courses/6-012-microelectronic-devices-and-circuits-fall-2009/resources/mit6_012f09_lec01/

[2] MIT OpenCourseWare · Clifton Fonstad (2009). 6.012 Lecture 10: MOS Capacitors and MOSFETs.

게이트 전압과 표면 반전, MOSFET 동작의 근거.

강의 자료 · https://ocw.mit.edu/courses/6-012-microelectronic-devices-and-circuits-fall-2009/resources/mit6_012f09_lec10/

[3] MIT OpenCourseWare · Jesús del Alamo (2005.10.27). 6.012 Lecture 14: Digital Circuits (III), CMOS.

상보형 회로, 부하 충전·방전, 동적전력. 계산 표의 수치는 본 보고서가 정한 가상값.

강의 자료 · https://ocw.mit.edu/courses/6-012-microelectronic-devices-and-circuits-fall-2005/resources/lec14/

[4] Computer History Museum. 1947: Invention of the Point-Contact Transistor.

연도는 사건의 연도. 바딘·브래튼의 점접촉 트랜지스터와 벨 연구소의 연구 맥락.

박물관 기술사 · https://www.computerhistory.org/siliconengine/invention-of-the-point-contact-transistor/

[5] Computer History Museum. 1958: All Semiconductor “Solid Circuit” is Demonstrated.

킬비의 게르마늄 회로 시연. 부품 사이의 가는 도선 연결이 남아 있었음.

박물관 기술사 · https://www.computerhistory.org/siliconengine/all-semiconductor-solid-circuit-is-demonstrated/

[6] Computer History Museum. 1959: Practical Monolithic Integrated Circuit Concept Patented.

노이스의 산화막 절연과 금속 연결 개념, 호어니의 평면 공정 기반.

박물관 기술사 · https://www.computerhistory.org/siliconengine/practical-monolithic-integrated-circuit-concept-patented/

[7] Computer History Museum. 1960: Metal Oxide Semiconductor (MOS) Transistor Demonstrated.

아탈라·강대원의 MOS 소자: 1959년 제작, 1960년 공개 시연.

박물관 기술사 · https://www.computerhistory.org/siliconengine/metal-oxide-semiconductor-mos-transistor-demonstrated/

[8] Computer History Museum. 1963: Complementary MOS Circuit Configuration is Invented.

완래스와 사의 상보형 MOS 회로. 낮은 대기 전력과 초기 도입의 제약.

박물관 기술사 · https://www.computerhistory.org/siliconengine/complementary-mos-circuit-configuration-is-invented/

[9] Computer History Museum. 1971: Microprocessor Integrates CPU Function onto a Single Chip.

4004 등 초기 마이크로프로세서의 맥락. CPU 칩과 완전한 컴퓨터 시스템을 구분.

박물관 기술사 · https://www.computerhistory.org/siliconengine/microprocessor-integrates-cpu-function-onto-a-single-chip/

출처와 확인한 내용

자료 확인 기준: 2026.10.31. 날짜가 있는 자료는 발표·사건 시점을 구분해 표기했다.

[10] Gordon E. Moore (1965; 1975). 집적도 증가에 관한 두 원문.

1965년에는 연간 증가 추세를, 1975년에는 약 2년 주기로의 변화를 논의. 성능의 물리 법칙과 구분.

1965 · Cramming More Components onto Integrated Circuits · https://www.cs.cmu.edu/~18742/papers/Moore1965.pdf

1975 · Progress in Digital Integrated Electronics · https://ethw-images.s3.us-east-va.perf.cloud.ovh.us/ieee/5/5e/Moore_1975_IEEE_Speech.pdf

[11] Robert H. Dennard 외 (1974); IBM History. MOSFET 축소 이론과 그 한계.

논문: IEEE JSSC 9(5), 256-268, DOI 10.1109/JSSC.1974.1050511. 역사 자료는 2000년대 중반의 한계를 설명.

Design of Ion-Implanted MOSFET's with Very Small Physical Dimensions · https://research.ibm.com/publications/design-of-ion-implanted-mosfets-with-very-small-physical-dimensions

Bob Dennard · IBM History · https://www.ibm.com/history/bob-dennard

[12] ASML. How Microchips Are Made; Six Crucial Steps in Semiconductor Manufacturing (2021).

반복되는 웨이퍼 공정, 설계·제조 역할, 검사·패키징의 개요.

제조 개요 · https://www.asml.com/en/technology/all-about-microchips/how-microchips-are-made

공정 단계 · https://www.asml.com/en/company/stories/2021/semiconductor-manufacturing-process-steps

[13] ASML. Light and Lasers; Rayleigh Criterion.

EUV 파장 13.5 nm, 광학계·공정 조건과 해상도의 관계.

노광 광원 · https://www.asml.com/en/technology/lithography-principles/light-and-lasers

해상도 원리 · https://www.asml.com/en/technology/lithography-principles/rayleigh-criterion

[14] Intel (2025.02.14 검토). 반도체 공정 노드 명칭 안내.

현대 공정 이름의 수치를 특정 구조의 실제 길이와 일대일로 대응시키지 않음.

공정 명칭 자료 · https://www.intel.com/content/www/us/en/support/articles/000097598.html

[15] TSMC (2026). 2025 Annual Report; 2nm Technology.

연차보고서와 공정 안내에서 N2의 2025년 4분기 양산, 나노시트 구조를 확인.

2025 연차보고서 · https://investor.tsmc.com/static/annualReports/2025/english/index.html

2nm 공정 안내 · https://www.tsmc.com/english/dedicatedFoundry/technology/logic/l_2nm

[16] Intel. Intel 18A; Intel Core Ultra 9 Processor 386H Specifications.

18A의 RibbonFET·PowerVia와 제품의 CPU 공정·2026년 1분기 출시 상태를 교차 확인.

18A 공정 · https://www.intel.com/content/www/us/en/foundry/process/18a.html

제품 사양 · https://www.intel.com/content/www/us/en/products/sku/245529/intel-core-ultra-9-processor-386h-18m-cache-up-to-4-90-ghz/specifications.html

[17] Samsung Newsroom (2026.02.12). Samsung Ships Industry-First Commercial HBM4 with Ultimate Performance for AI Computing.

본 보고서는 양산·상용 출하 발표만 인용한다. 제목의 우선권·최상급 표현을 평가 결과로 사용하지 않는다.

공식 발표 · https://news.samsung.com/global/samsung-ships-industry-first-commercial-hbm4-with-ultimate-performance-for-ai-computing

출처와 확인한 내용

자료 확인 기준: 2026.10.31. 날짜가 있는 자료는 발표·사건 시점을 구분해 표기했다.

[18] Micron. HBM.

여러 DRAM 다이의 적층, TSV와 넓은 데이터 연결. 제품별 수치와 전체 시스템 효과는 별도 검증 필요.

기술 개요 · https://www.micron.com/products/memory/hbm

[19] Kioxia. BiCS FLASH: 3D Flash Memory Technology.

NAND 셀의 수직 배치와 평면 미세화의 제약. HBM의 다이 적층과 구분.

기술 개요 · https://www.kioxia.com/en-jp/rd/technology/bics-flash.html

[20] TSMC 3DFabric. CoWoS.

연산 다이·HBM과 인터포저를 활용한 패키지 연결. 그림은 해당 구조를 단순화한 자체 제작 도식.

기술 개요 · https://3dfabric.tsmc.com/english/dedicatedFoundry/technology/cowos.htm

[21] Samuel Williams, Andrew Waterman, David Patterson (2008.10.17). Roofline: An Insightful Visual Performance Model for Floating-Point Programs and Multicore Architectures.

UC Berkeley, UCB/EECS-2008-134. 연산 집약도와 메모리 대역폭에 따른 성능 상한. 수치·그래프는 가상 예시.

원문과 서지 정보 · https://www2.eecs.berkeley.edu/Pubs/TechRpts/2008/EECS-2008-134.html

[22] imec (2026.09.07). Advancing the CFET-Based Device Roadmap: Novel Integration Modules and Standard Cell Configurations, Part 1.

n/p 소자 수직 적층, 접점·게이트 통합 연구. 장기 노드 로드맵을 양산 실적으로 해석하지 않음.

연구 개요 · https://www.imec-int.com/en/articles/advancing-cfet-based-device-roadmap-novel-integration-modules-and-standard-cell

[23] imec · ASML · TSMC (2026.06.15). 2D-Material Transistors: Breakthrough 300mm Integration.

지름 300 mm 웨이퍼에서 2D 소재 nFET·pFET 공정 통합 시연. 완성된 상용 CPU의 양산과 구분.

연구 발표 · https://www.imec-int.com/en/press/asml-tsmc-and-imec-bring-industry-ready-2d-material-transistors-closer-breakthrough-300mm

[24] NVIDIA Newsroom (2026.05.31). NVIDIA Vera Rubin Ramps Into Full Production to Power Agentic AI Factories Worldwide.

본문의 Spectrum-X Ethernet Photonics 생산 발표를 인용. 광연결 스위치와 광학 논리 연산은 다른 주장.

공식 발표 · https://nvidianews.nvidia.com/news/vera-rubin-full-production-agentic-ai-factory

[25] Manuel Le Gallo 외 (2023.08.10). A 64-Core Mixed-Signal In-Memory Compute Chip Based on Phase-Change Memory for Deep Neural Network Inference.

Nature Electronics 6, 680-693. DOI 10.1038/s41928-023-01010-1. 연구용 칩의 특정 추론 과제 실증.

논문 요약·원문 연결 · https://research.ibm.com/publications/a-64-core-mixed-signal-in-memory-compute-chip-based-on-phase-change-memory-for-deep-neural-network-inference

[26] Infineon Technologies. Wide-Bandgap Semiconductors (SiC/GaN).

전력·주파수 등 적용 조건에 따른 소재 선택. 제조사의 일반적 이점 설명을 모든 제품의 비교 우위로 일반화하지 않음.

기술·응용 개요 · https://www.infineon.com/technology/wide-bandgap-semiconductors-sic-gan

제작·분석 안내: 본문의 도식은 교육용으로 단순화해 직접 구성했다. 수치 예시는 공개 자료의 원리·모형에 가상의 조건을 대입한 것이며, 제품 성능 측정치로 사용할 수 없다. 기업 보도자료의 우선권·최상급 표현은 보고서의 독립적 판단으로 채택하지 않았다.