



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

석사학위논문

반도체 초진화론

-세계를 지배하는 기술의 미래-

(半導体超進化論-世界を制する技術の未来-)

한유경

제주대학교 통역번역대학원

한일과

2024년 2월



반도체 초진화론

-세계를 지배하는 기술의 미래-

(『半導体超進化論-世界を制する技術の未来-』 翻譯論文)

지도교수 이 예 안

한 유 경

이 논문을 통역번역학 석사학위 논문으로 제출함

2023년 12월

한유경의 통역번역학 석사학위 논문을 인준함

심사위원장 _____ (인)

위 원 _____ (인)

위 원 _____ (인)

제주대학교 통역번역대학원

2023년 12월

목차

역자서문	1
국문초록	2
일러두기	3
I. 일양내복(一陽來復) Prologue	4
1. 만찬회-스테이지가 바뀌다	4
2. 도쿄대학이 움직인다-Agile!	7
3. More People-세계의 두뇌들을 끌어들여라	10
4. 반도체의 숲-공생과 공진화(共進化)	12
II. 권토중래(捲土重來) Game Change	15
1. 반도체 전략-선선의 선을 치다	15
2. 범용 칩에서 전용 칩으로-반도체 산업의 게임 체인지	21
3. 산업의 쌀에서 사회의 뉴런으로-포스트 코로나 시대의 반도체	25
4. 댐에서 반도체로-디지털 사회의 인프라	30
III. 구조개혁(構造改革) More Moore	35
1. 뇌와 컴퓨터와 집적회로의 짧은 역사-그리고 하나의 미래	35
2. 스케일링 시나리오-지수함수의 경이	39
3. 칩의 구조 개혁-누전을 줄여라	43
4. AI 칩-뇌에서 배우다	48

역자서문

이 책의 저자 구로다 다다히로(黒田忠広, 1959)는 도쿄대 전기공학과를 졸업하고 도시바 연구소 재직 후 게이오대, 캘리포니아 버클리대, 도쿄대 교수를 거쳐 현재 도쿄대 대학원 교수이자 산학연계 반도체 연구센터인 d.lab(디랩) 연구센터장 겸 기술연구조합 RaaS(라스) 이사장이다. 연구자이자 반도체 전문가이기도 한 일본 반도체 산업의 핵심 인물이 일반인을 대상으로 쓴 첫 번째 책이라는 것이 이 책의 특징이다.

반도체를 둘러싸고 급격히 변해가는 정세 속에서 반도체 기업 간, 국가 간의 경계를 초월한 협력이 무엇보다도 중요해졌다. 반도체 산업은 특정 국가나 기업이 주도하기 어려워졌으며 일본은 반도체 분야에서 세계 각국의 기업들과 연계하고 있다.

이 책은 일본 반도체 전략의 지향점을 제시한다. 저자는 검토에서 ‘선선의 선을 치듯’ 다음 경쟁의 무대를 예견해 이에 대한 역량을 집중적으로 키워야 한다고 주장한다. 바로 미세화 기술에 대한 투자이다.

일본은 인구 감소가 진행되고 있는 상황 속에서 단지 자율주행 자동차 개발 또는 AI, 양자 등과 같은 최첨단 기술을 위해서만이 아니라 노동력 부족에 대처하기 위해 디지털 전환을 추진하고 있다. 한국도 인구가 감소하고 있는 것은 마찬가지이기에 반도체 산업에 더욱 주력하기 위해서는 한국과 일본의 협력도 중요할 것으로 보인다.

반도체는 뉴스에 자주 나오고, 일상생활 속 곳곳에 들어가 있다. 본서에서는 저자의 인사이트를 통해 반도체 산업에 대해 관련 업계 종사자뿐만 아니라 전반적인 지식을 교양적인 차원에서 습득하려는 일반인도 이해할 수 있도록 설명하고 있다. 총 6장으로 구성되어 있으며 그중 제1장, 제2장, 제3장을 번역하였다. 본 번역 논문이 반도체와 일본의 전략에 대한 이해를 넓히는 데에 보탬이 되었으면 한다.

마지막으로 본고의 번역에 귀한 조언을 해 주신 교수님들, 도움과 피드백을 준 동료들에게 감사를 드린다.

국문초록

본서는 반도체 정세와 격변하는 환경에 대한 대응책을 제시한다.

1장에서는 반도체 산업 무대의 변화에 본격적으로 대처하는 일본의 상황을 저자의 시점에서 보여준다. 또한 기술 개발과 인재 양성에 초점을 맞춘 일본의 전략을 제시한다.

2장에서는 일본 반도체 산업의 지난 30년간을 돌아보고, 주도권을 되찾기 위해 선행투자가 중요하다는 일본의 반도체 전략을 제시한다. 또한 현 정세를 풀어내기 위해 이해해야 하는 세 가지 징조(산업 주역 교체, 시장 변화, 기술 패러다임 변화)를 설명한다. 디지털 사회의 기반이 되는 반도체의 중요성 및 일본의 전략과 연계를 소개한다.

3장은 반도체 집적회로(IC)의 발전 과정을 짚어보며 기술의 한계와 극복을 위한 해결책을 제시한다. 칩 구조 개혁의 필요성과 현 기술들을 소개하고 인간의 뇌가 발전해왔듯이 실리콘 뇌에 대한 저자의 통찰을 이야기한다.

일러두기

1. 인명, 지명의 외래어 표기는 국립국어원 외래어 표기법을 따르되, 국내에서 널리 통용되는 표기가 있을 경우 그에 따랐다.
2. 동음어의 경우 괄호 안에 한자를 표기하였다.
3. 원 텍스트의 주석은 저자가 독자의 이해를 돕기 위한 추가적인 설명이 기재되어 있으며 본문 내용과 독립적인 내용이므로 번역을 생략하였다.
4. 영어 원문 인용의 경우, 원문을 기본으로 하되 가독성을 높이기 위해 일부를 수정했다.
5. 저자가 강조한 부분은 볼드로 표기했다.

I. 일양내복(一陽來復) Prologue

1. 만찬회—스테이지가 바뀌다

2022년 12월 14일 밤.

오쿠라 호텔 대연회장 ‘헤이안 룸(平安の間)’은 400명의 열기에 휩싸여 있었다. 사회자 아바타가 거대한 스크린에 등장했고 사람들은 점차 자리를 채웠다.

맨 앞줄 중앙에는 VIP들이 나란히 앉았다.

아마리 아키라(甘利明) 자민당 반도체전략추진의원연맹 회장, 니시무라 야스토시(西村康稔) 경제산업상, 고노카미 마코토(五神真) 이화학연구소 이사장. 경제산업성의 노하라 사토시(野原諭) 상무정보정책국장, 니시카와 가즈미(西川和見) 총무과장, 가나자시 히사시(金指壽) 과장, 오기노 요헤이(荻野洋平) 실장.

산업계에서는 사와다 준(澤田純) NTT 회장, 고시바 미쓰노부(小柴滿信) JSR 명예회장, 쓰네이시 데쓰오(常石哲男) 도쿄 일렉트론(TEL) 전 회장과 가와이 도시키(河合利樹) 사장, 요시다 요시아키(吉田芳明) 아드반테스트 사장, 히로에 도시오(廣江敏朗) 스크린 홀딩스 사장, 시미즈 데루시(清水照士) 소니 세미컨덕터 솔루션즈 사장, 하야사카 노부오(早坂伸夫) 키옥시아 사장, 시바타 히데토시(柴田英利) 르네사스 일렉트로닉스 사장, 가와하라 노부아키(川原伸章) 미라이즈 테크놀로지스 이사, 데라마치 아키히로(寺町彰博) THK 사장, 호리바 아쓰시(堀場厚) 호리바제작소 회장 겸 CEO가 참가했다.

여기에 다리오 길(Dario Gil) IBM 연구소장과 무케시 카레(Mukesh Khare) 부소장, 조데벡(Jo De Boeck) imec 최고전략책임자, 아짓 마노차(Ajit Manocha) SEMI(국제반도체장비재료협회) CEO, 준 허(Jun He) TSMC 부사장. 오노데라 마코토(小野寺誠) TSMC 재팬 사장과 에모토 유타카(江本裕) TSMC 재팬 3 DIC 연구개발센터장의 얼굴도 보인다.

그리고 오늘 밤의 주인공인 히가시 데쓰로(東哲郎) 라피더스 회장과 고이케 아쓰요시(小池淳義) 사장이 가운데 자리에 앉았다.

모두가 밝은 미래를 마음에 그리고 있었다.

반도체 산업은 성장산업이다.

1982년에 150억 달러(약 3.7조 엔)였던 반도체 시장은 2021년에 5,000억 달러(약 65조 엔)에 달했다. 평균 연 9.4%의 고도성장이 40년 동안 이어지고 있다.

당초에 반도체 시장 규모는 명목 GDP의 0.2% 정도였다. 그랬던 것이 1990년대 중반에 0.4%로 급성장했다. 무슨 일이 일어난 걸까?

1990년대 중반에 윈도우 95가 세계적인 히트 상품이 된 일을 떠올리는 사람들도 많을 것이다. 그전에 반도체는 TV나 비디오 플레이어 등 물리 공간을 풍요롭게 하는 가전제품에 많이 사용되었지만, 윈도우 95 출시 이후에는 개인용 컴퓨터(PC)나 스마트폰에 많이 사용되었다.

그리고 PC는 가상 공간을 만들어냈으며 스마트폰은 그 공간을 휴대할 수 있게 했다.

즉, 반도체 무대가 물리 공간에서 가상 공간으로 확대되면서 반도체 시장은 GDP의 0.2%에서 0.4%로 성장한 것이다.

최근 반도체 시장이 GDP의 0.6%를 목표로 재도약하려는 움직임을 보이고 있다.

코로나19로 인한 특수 수요가 있었기 때문에 조금 더 주의 깊게 살펴볼 필요가 있을 것이다. 이미 시장은 조정국면에 들어갔다. 하지만 조정 이후에 다시 한번 큰 성장이 기다리고 있다면 반도체 시장은 제3의 성장기를 맞이하게 될 것이다.

반도체가 만들어내는 새로운 가치는 물리 공간과 가상 공간의 고도 융합을 통한 데이터 구동형 사회 창출이며, 사회적 과제 해결과 경제 발전의 양립이다.

자율 주행이나 로봇공학, 스마트 시티가 그 예다. 센서가 모은 물리 공간의 실시간 데이터를 AI(인공지능)가 가상 공간 속 디지털 트윈을 활용해 분석하고, 즉시 물리 공간으로 돌아가 모터를 제어한다. 이렇게 사람들은 목적지까지 최단 시간과 최소 에너지를 소모해 안전하고 쾌적하게 이동할 수 있게 된다.

세계의 반도체는 이렇듯 건실하게 성장하고 있다.

한편 국내를 살펴보면 일본의 반도체 산업은 근 25년 동안 휴면 상태였다. 한국, 대만, 중국이 급속히 성장하는 가운데 일본만이 성장하지 못했다.

일본 반도체 산업의 조락(凋落) 요인이 여러 가지로 지적을 받았다. 미일 무역 마찰과 엔화 강세라는 경영 환경적 요인. 디지털화와 수평분업이 늦어지는 등 전략적 요인. 외부에 의존하지 않고 스스로 해결하는 일본 자전주의(自前主義)나 한국, 중국, 대만 정부의 기업 육성에 대항하지 못했던 산업정책적 요인 등이 있다.

그러나 상황이 바뀌었다.

경영 환경은 미일 연대와 엔화 약세로 호전되고 있다. 또한 디지털 산업 및 파운드리 투자 전략이 공격적으로 바뀌었다. 산업정책도 국제 연대를 추진하는 쪽으로 바뀌었다.

국가가 각오를 하고, 말하자면 국운을 걸고 반도체 산업 재생에 본격적으로 도전하고 있다.

이번에야말로 변할 것이다.

SEMI가 주최한 전시회 ‘세미콘 재팬 2022’의 주제는 ‘미래를 바꾸다. 미래가 바뀌다’였다.

기시다 후미오(岸田文雄) 일본 내각 총리가 개회식에 참석해 다음과 같이 결의를 다졌다.

“반도체는 말할 것도 없이 디지털화, 탈탄소화와 경제 안전보장 확보 등을 뒷받침하는 핵심 테크놀로지입니다.

그린, 디지털 등 분야에서의 사회적 과제를 성장 동력으로 전환하고, 지속 가능한 경제사회 실현을 위한 ‘새로운 자본주의’를 떠받치는 가장 중요한 물자입니다.

코로나로 인한 어려움을 극복하고 사회 경제 활동을 정상화해 나가며, 엔화 약세의 이점을 활용할 것입니다. 이를 위해서라도 사회를 떠받치는 반도체에 대한 공격적인 일본 국내 투자 확대를 지원하고 경제 구조를 강화해 나가겠습니다.

구마모토(熊本)에 유치한 TSMC 반도체 공장은 해당 지역에 10년간 4조 엔을 넘는 경제 효과와 7,000명 이상의 고용을 창출할 것으로 시산하고 있습니다.

지역 활성화로도 이어지는 이러한 투자를 더욱 촉진하기 위해 얼마 전 통과된 추가 경정 예산에 1조 3천억 엔을 편성했습니다. 이를 통해 일본 국내 반도체 투자를 전국적으로 확대하고 차세대 반도체 개발을 추진하겠습니다.

반도체 공급망을 한 나라가 담당하기는 어렵다는 현실을 직시해야 합니다. 정부가 지원하는 반도체 개발 프로젝트에서도 글로벌 연계를 강화해 나가겠습니다.

미래의 차세대 반도체 양산 거점이 될 기업 라피더스가 어제 IBM과의 공동 개발 파트너십 체결을 공표했습니다. 더불어 유럽 imec와도 협력해 나가며 2020년대 후반에 양산을 실현하는 것을 목표로 하고 있습니다.

AI 및 양자(量子) 같은 고도 계산 시스템이나 자율 주행, 차세대 로봇 등 앞으로 세계적으로 크게 진화해 나갈 디지털 경제를 뒷받침하는 최첨단 반도체를 일본에서도 공급하고자 합니다.”

2. 도쿄대학이 움직인다—Agile!

‘헤이안 룸’ 대연회장 무대에 고노카미 이사장이 올랐다. 고노카미 이사장은 도쿄대학교 총장을 역임한 후 이화학연구소 이사장으로 취임하면서 반도체와 양자 컴퓨팅 전략을 제시했다.

필자가 고노카미 총장과 만난 것은 2019년 5월이었다. 신록이 눈부신 도쿄대학 혼고(本郷) 캠퍼스의 야스다(安田) 강당은 정문이 3층에 있다. 경비원에게 신분증을 보여주고 들어가 계단을 내려와서 반시계 방향으로 돌자 비밀 기지 같은 공간이 나타났다.

고노카미 총장으로부터 일본의 반도체 부흥을 위해 무엇이 필요한지라는 질문을 받았다.

나는 “에너지 효율이 높은 전용 칩을 효율적으로 개발하고 3D 집적(集積)하는 기술입니다”라고 대답했다.

그리고 다음 설명을 덧붙였다.

“데이터 구동형(驅動型) 사회 Society 5.0 창출에 필요한 것은 고도의 컴퓨팅입니다. 에너지와 함께 일본이 가장 필요로 하는 자원입니다.

컴퓨팅 과제는 에너지 효율을 높이는 것입니다. 데이터 센터가 소비하는 전력은 이대로라면 10년 후에 10배로 급증할 것입니다. 에너지 위기의 해결 없이 데이터 구동형 사회의 지속 가능한 발전은 없습니다.

에너지 위기의 원인은 사실 AI에 있습니다. 폭발적으로 증가하는 데이터를 더 고도로 분석하기 위해 근 10년 동안 AI 계산량은 4자릿수나 늘었습니다. 반면 계산을 담당하는 범용 프로세서의 전력 효율은 10년 동안 1자리밖에 개선되지 않았습니다.

에너지 효율을 개선하는 기술은 반도체 미세화와 3D 집적에 있습니다. 미세화에 있어 일본은 세계 최첨단에서 크게 뒤쳐져 있으므로 이 부분은 해외에서 배워야 합니다.

한편, 일본에는 3D 집적에 필요한 소재나 제조 장비 분야의 뛰어난 기술이 적지 않습니다. 3D 집적의 초크 포인트(choke point)를 장악해야 한다고 생각합니다. 3D 집적을 통해 데이터의 이동 거리를 월등하게 단축할 수 있다면 데이터 이동에 쓰이는 에너지 소비를 크게 줄일 수 있을 것입니다.

3D 집적을 수행하는 패키지 공정은 미세화를 진행하는 웨이퍼 공정에 비해 훨씬 적은 투자로도 가능하기 때문에 3D 집적은 투자 효과가 높습니다.

한편, 설계기술에 관해서는 불필요한 회로를 제거한 전용 칩이 범용 칩에 비해 탁월하게 에너지 절약이 가능합니다. 전용 칩을 개발하려는 움직임은 GAFA나 테슬라 등에서는 이미 시작됐습니다. 범용 칩 시대는 자본 경쟁의 시대였지만 전용 칩 시대는 지(知)의 경쟁의 시대입니다. 바꿔 말하자면 설계 개발의 혁신이 필요합니다.

전용 칩 개발은 점점 어려워지고 있으며 최근에는 설계자 100명을 모아도 1년이라는 기간과 100억 엔이라는 비용이 듭니다. 이렇게 긴 시간과 큰 비용이 요구되면 전용 칩에 대한 관심이 줄어들고 설계자도 감소합니다. 이미 일본은 관심과 능력을 잃어가고 있으며 이러한 상황 속에서는 새로운 공장을 건설하고

제조 능력을 보유할 수 있더라도 즉시 산업력 강화로 이어지지 않습니다.

AI 기술은 나날이 발전하여 소프트웨어는 매달 업데이트됩니다. 디지털 경제에서는 하드웨어와 소프트웨어를 고도로 융합시켜 혁신을 창출하고 고속 사이클로 개량을 거듭하는 것이 핵심이지만 두 요소의 개발 속도가 이렇게나 차이가 나면 어려워집니다. 그러나 프로그램을 컴파일하고 칩을 자동으로 설계할 수 있는 실리콘 컴파일러가 있다면 재빠르게(Agile) 하드웨어를 개발할 수 있습니다.

물론 자동 설계된 회로의 성능은 설계자가 시간을 들여 최적화한 회로와 비교하면 80점 정도겠지만 이로 충분합니다. 이른바 ‘80점 주의’입니다. ‘80 대 20 법칙’을 살려 개발 효율을 5배 높이는 부가가치를 찾아냅니다.

이에 더해 설계 자산을 재이용함으로써 설계 규모의 폭발적인 증대를 억제할 필요도 있습니다. 칩렛이 앞으로 중요해질 것입니다. 칩렛을 구성해 패키지 내에서 시스템이 완성되기 때문에 3D 집적은 중요한 초크 포인트가 될 것입니다.”

도쿄대학의 움직임은 빨랐다.

우선 사회 연계를 캠퍼스에서 공개적으로 수행하는 센터인 d.lab을 2019년 10월에 설립했다. d.lab의 ‘d’는 디지털 기술로 각 개인이 빛나는 시대(digital inclusion)에 데이터(data)를 기점으로 소프트웨어에서 장치(device)까지 일관성 있는 영역 특화형(domain specific) 시스템을 디자인(design)하는 연구를 수행한다는 취지가 담겨 있다. 또한 기술을 젊은 세대에 전수하기 위해 사무실은 학생 기숙사(dormitory)에 만들어졌다.

다음으로 11월에는 TSMC와 반도체 기술 공동 연구를 세계적으로 앞장서서 모든 대학과 모든 회사 수준에서 수행하겠다고 발표했다. 공동 기자 회견에는 마크 리우(Mark Liu) TSMC 회장과 연구 책임자인 스탠퍼드대학 필립 왕(Philip Wong) 교수가 참석했고, 도쿄대학 총장이었던 고노카미 마코토와 당시 부총장이었던 후지이 데루오(藤井輝夫) 현 총장이 굳게 손을 잡았다.

그리고 이듬해인 2020년 8월에는 산학관 연계를 철저한 정보 관리하에 실시하는 기술 연구 조합 RaaS를 설립했다. RaaS는 첨단 시스템 기술 조합(Resea

rch Association for Advanced Systems)의 약자로 Research as a Service를 목표로 하며 ‘라스’라고 부른다.

현재 d.lab의 협찬 회원은 49개 사이며 RaaS 참여 기업은 누적 12개 사이다.

d.lab와 RaaS의 목표는 에너지 효율을 10배로 높이며 개발 효율도 10배로 키우는 것이다.

이 목표는 라피더스와 동일하다.

달성하기 위한 수단은 라피더스와 상호보완적이다.

즉, 에너지 효율 개선을 위해 라피더스는 미세화를 추구하고, 도쿄대학은 3D 집적을 추구하고, 그리고 개발 효율을 개선하기 위해 라피더스는 제조 기간을 단축하고, 도쿄대학은 설계 기간을 단축하는 것이다.

3. More People—세계의 두뇌들을 끌어들여라

기술뿐만이 아니라 인재 양성이 시급한 과제이다. **기술은 사람이다.**

그렇기에 2022년 4월 반도체 민주화 거점인 Agile-X(애자일X)가 첫발을 내딛는 것이다.

Agile은 ‘매우 빠른’, ‘민첩함’을 의미한다.

전용 칩 개발에 필요한 기간과 비용을 10분의 1로 단축할 수 있는 개발 플랫폼을 구축하여 세계의 두뇌들을 끌어들이고, 그 결과 전용 칩을 설계하는 사람이 10배로 늘어나면 반도체를 민주화할 수 있다. 이것이 AgileX의 목표다.

민주화를 추구하게 된 배경에는 ‘집단 두뇌’라는 사고방식이 있다. 더 많은 사람들의 아이디어가 교차하는 곳에서 혁신이 탄생한다는 발상이다.

예를 들자면 남태평양의 섬들에서 어업에 쓰이는 도구 종류와 섬 주민 인구 간에는 강력한 상관관계가 있다. 바꿔 말하자면 인구가 많을수록 많은 도구가 사용되고 있는 것이다.

마찬가지로 호모 사피엔스가 네안데르탈인과 비교했을 때 뇌 용적은 작았음에도 여러 가지 도구를 발명해 이용한 이유도, 호모 사피엔스 쪽이 더 큰 집단을

형성할 수 있었기 때문이라고 한다.

더 많은 사람들이 자신의 칩을 개발할 수 있게 된다면 더 많은 혁신이 생겨날 것이다. 이러한 사상 하에 반도체 민주화 운동이 세계에서 조용히 시작되고 있다.

마크 리우 TSMC 회장은 2021년 반도체 국제회의인 ISSCC(국제고체회로소자학회)에서의 기조 강연을 다음과 같이 마무리했다.

“혁신은 아이디어의 자유로운 흐름 속에서 탄생합니다. 아이디어는 사람에게서 나옵니다. 더 많은 사람들이 자신의 반도체를 소유함으로써 혁신이 민주화됩니다.”

1959년 물리학자 리처드 파인만(Richard Feynman)이 강연에서 “바닥에는 충분한 공간이 있다(There’s plenty of room at the bottom).” 즉 “나노 스케일 영역에는 아직 많은 흥미로운 것들이 있다”라고 말한 것을 계기로 세계는 미세 디바이스를 추구했다. 곧이어 마이크로일렉트로닉스가 탄생하고 나노일렉트로닉스가 발전하게 되었다.

미세화의 한계가 가까워진 가운데 미세화를 더 발전시켜 무어의 법칙을 계속 이어가는 More Moore 연구가 현재도 이어지고 있다.

최근에는 미세화 대신 새로운 가치 창출을 목표로 하는 More than Moore 연구도 시작되었다. 3D 집적 연구는 투자 효과가 높기도 해서 전 세계적인 투자를 모으고 있다.

나는 파인만의 말을 되새기며 “There’s plenty of room at the TOP”라고 하겠다.

집적도가 트랜지스터 1,000억 개를 넘는 칩 실현이 머지않았다. 팻 겔싱어(Pat Gelsinger) 인텔 CEO는 2030년까지 패키지 안에 트랜지스터 1조 개가 집적될 것이라고 예언했다.

더 많은 사람이 참가해 혁신 창발이 가능해지도록 ‘More people 연구’가 중요해진다.

거대 기업만이 전용 칩을 개발할 수 있는 이유는 산업 시스템이 공업화 사회의 대량 생산에 최적화되어 있기 때문이다.

지식이 가치를 창출하는 지식사회(知識社會)에서는 비용 대비 성능보다 시

간 대비 성능이 더 중요해질 것이다. 시간은 돈이기 때문에 시간 대비 성능은 비용 대비 성능을 내포한다.

빠름(Agile)의 가치를 누구보다도 잘 아는 사람은 연구자다. 논문 발표는 일각을 다투는 경쟁이다. 반도체가 연구자의 접근성을 높여 과학의 발전에 공헌하는 것은 중요하다.

반도체의 민주화를 추진하고 세계의 두뇌들을 끌어들이는 것. 파이 쟁탈이 아닌 파이 확장을 목표로 한다.

인재 양성은 아카데미아의 역할이다. 인재야말로 일본의 자본이며 일본의 미래를 개척한다.

일본은 More Moore를 세계로부터 배워 More than Moore와 More People로 세상에 공헌할 수 있을 것이다.

또한 일본에 고도의 컴퓨팅 인프라를 구축해야 한다. 기존 비트에 양자비트와 뉴런을 하이브리드 방식으로 결합하고, 여기에 소프트웨어와 하드웨어를 융합하여 고도의 계산 능력을 갖추어 전국에서 접속할 수 있는 통신망을 정비해야 한다. 이것이 일본이 총력을 다해 추진해야 할 디지털 사회 인프라이다.

4. 반도체의 숲—공생과 공진화(共進化)

반도체는 전략 물자이다. 기술 패권을 둘러싼 거대한 게임의 승자는 누구일까. 지정학적 리스크가 한층 더 높아지며 세계의 앞날은 불확실성이 커져가고만 있다.

패권을 쥐는 것이 아니라 민주화를 추진하고 반도체를 세계의 공동 자산, 인류 공유 재산(Global commons)으로 삼아 다양한 칩을 만들어 세상을 변영으로 이끌 수는 없을까?

그 대답에 대한 힌트는 지구의 다양성에 있는 것 같다.

백악기(약 1억 4,500만 년 전~6,600만 년 전) 이전, 생물의 종류는 현재의

1/10밖에 없었다.

그러나 꽃의 탄생이 지구를 일변시켰다.

꽃가루를 제공하는 대가로 꽃가루를 곤충이 운반케 했다. 이전까지는 일방적으로 먹혀왔던 식물이 곤충을 이용하는 대전환이 일어났다.

꽃은 곤충을 불러들이기 위해 화려한 색을 뽐내고 곤충은 꽃의 형태 변화에 맞춰 비행 능력을 향상시켰다. 서로가 서로를 진화시키는 진화의 상호작용, 공진화가 반복된 것이다.

이렇게 숲이 풍요로워지고 꽃에 모인 곤충을 먹는 포유류가 다양해지며 꽃에 열리는 열매를 통해 영장류가 진화했다.

드디어 꽃은 새로운 능력을 습득했다.

세대교체의 속도를 높인 것이다. 수분에서 수정까지 필요한 시간을 1년에서 수 시간으로 줄였다. 이것이 모든 생물의 진화를 가속화했다.

$$y=a(1+r)^n$$

이는 복리 계산 공식이다. r은 이율이고 n은 운용 횟수이다. 자본 a가 적어도 오래 운용하면 미래 가치가 커진다.

n을 1/t로 치환하면 디지털 경제의 기본 식이 된다. t는 개발 사이클 타임이다. 이 식은 칩의 성능 향상뿐만이 아니라 회사의 성장에도 적용된다.

바꿔 말하자면 고속 사이클로 개량을 여러 번 반복하는 것이 디지털 경제의 성장 전략이다. 개선율(r)보다도 개선 횟수(n)를 크게 하는 것, 즉 개발 사이클 타임(t)을 짧게 하는 것이 가장 중요하다.

그래서 애자일(Agile)인 것이다.

생명체들은 힘겨운 생존경쟁을 벌이면서도 종을 초월해 복잡하게 연결되어 서로 도우며 살아왔다. 공생과 공진화. 이 열쇠를 애자일이 쥐고 있다고 할 수 있다.

식물을 칩으로, 곤충을 칩 사용자로, 숲을 생태계로 바꿔 읽어보면 어떨까? 경쟁에서 공생과 공진화로 바꾸는 반도체의 ‘꽃’은 무엇일까?



‘헤이안 룸’에서 연회가 한창이다.

사회자 아바타가 바이올리니스트 하카세 다로(葉加瀬太郎)를 무대로 불러냈다. 사람들의 스마트폰 안에서 수천억 개의 반도체 스위치가 수억 번 동작했다.

“<어나더 스카이>를 연주하겠습니다. 비행기에도 반도체가 많이 쓰이고 있습니다.”

하카세 다로는 이렇게 말하며 관객들의 웃음을 자아냈다.

나는 <랩소디 인 블루>를 들으면 하늘 위 여행이 생각나는데, 이 순간만큼은 두 달 동안 정신없이 바빴던 해외 출장의 나날들을 떠올리고 있었다.

9월 19일 : 벨기에 imec에서 루크 반 덴 호브(Luc Van den hove) CEO와 협의

9월 24일 : 뉴욕 IBM 연구소에서 다리오 길 소장과 협의

9월 26일 : 올버니 나노텍 연구단지(Albany Nanotech Complex)에서 IBM의 무케시 카레 등과 의견·정보교환

9월 28일 : 프린스턴대학에서 연구 교류

10월 5일 : 캘리포니아대학 버클리 캠퍼스에서 연구 교류

10월 6일 : 로렌스 버클리 국립연구소에서 연구 교류

10월 10일~14일 : 일본 정부의 미국 사절단에 합류해 미국 상무부(DOC) 등을 방문

10월 31일 : SEMI가 개최한 국제회의 ITPC에서 인재 양성 관련 패널 토론 참가

11월 7일 : imec 테크놀로지 포럼에서 d.lab과 imec의 연계 발표

11월 29일 : 쓰쿠바시(つくば市)에 위치한 TSMC 재팬 3D IC 연구개발 센터를 d.lab 협찬 회원과 방문

오늘 밤 헤이안 룸에 모인 사람들과 최근 2달 동안 몇 번이나 만나 의견을 나눴다. 협조를 위한 네트워크 구축이 급선무였다.

옆자리에 앉아 있던 가나자시 히사시 경제산업성 과장이

“잠시 후에 DOC 회의가 있어서 먼저 들어가 보겠습니다.”

라며 자리에서 일어났다.

“수고가 많으십니다.”

바로 앞에서 하카세 다로가 <정열대륙>을 연주한다.

드디어 시작이다.

II. 권토중래(捲土重來) Game Change

1. 반도체 전략—선선의 선을 치다

게임 체인지

2021년 6월에 일본 경제산업성이 반도체 전략을 발표했다. 그 안에 『일본의 조락(凋落)』이라는 제목의 한 페이지 분량의 자료가 있다. 1988년에 50%였던 일본 기업의 세계시장 점유율이 그 이후에 굴러떨어지듯 일직선으로 하강하며, 현재에 이르러서는 10%에 불과하다는 내용이 나와 있다. 여기에 국민들의 관심이 집중되었다.

최근 30년간 세계의 반도체 시장은 연율 5% 이상의 고도성장을 이어왔지만 일본은 전혀 성장하지 못했다. 이대로 가면 ‘일본 점유율은 거의 0%로?!’라는 내용처럼 될 수도 있다. 반면에 세계 시장은 앞으로 디지털 혁명에 힘입어 연율 8%로 급성장하여 2030년에는 현재의 2배인 100조 엔을 돌파할 기세이다.

반전 시나리오는 있는 것일까?

반도체 전략의 요체는 한마디로 미세화 기술에 대한 적극적인 투자이다.

다만 정석적인 방법만으로 잃어버린 30년을 되찾기는 어렵다. 경쟁 무대의 제 2막을 예견해 선행 투자를 해야 할 필요가 있다. 검토에서 말하는 ‘선선의 선을 치는’ 것과 같다.

현재의 복잡한 정세를 읽어내려면 그 흐름을 만들어 내는 세 가지 변화를 이해할 필요가 있을 것이다.

하나의 산업의 주역 교체이다. 로직 반도체는 인텔 등 칩 제조업체에서 개발하는 범용 칩에서 GAFA 등 칩 사용자가 개발하는 전용 칩으로 주요 무대가 옮겨가려 하고 있다.

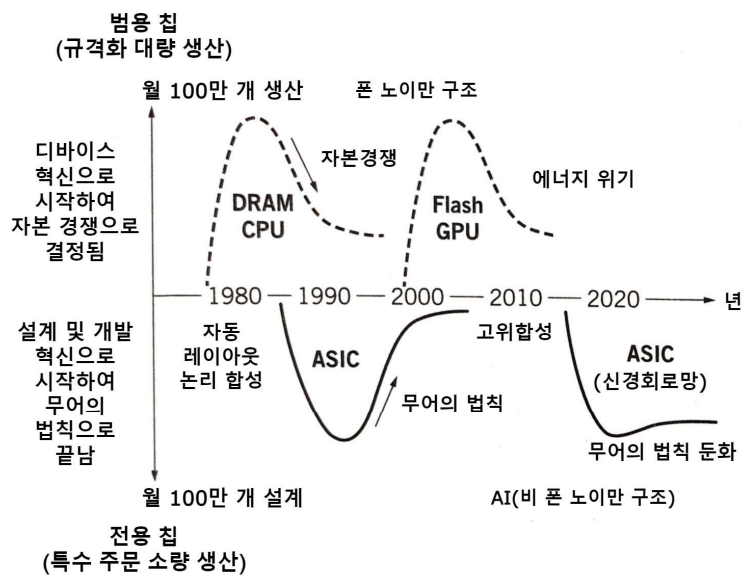
미국의 유력 벤처 캐피털 25개사가 2017년부터 3년 동안 투자한 안건을 살펴보면 무려 메모리 투자 금액의 9배나 되는 투자가 전용 칩과 AI 칩에 집중되어 있다.

전용 칩 시대가 도래한 것이다(그림 2-1).

사실 반도체 비즈니스의 왕도는 범용 칩을 규격화해 대량 생산하는 것이지만 전용 칩을 특수 주문 소량 생산했던 시대 또한 과거에는 있었다. 1985년부터 2000년 경이다. 범용 칩 사이에 분산되어 있던 로직을 하나의 전용 칩에 모음으로써 제품 제조 비용을 절감할 수 있었다.

하지만 전용 칩은 개발 비용이 든다. 그래서 컴퓨터를 이용한 자동 설계 기술이 미국의 대학에서 잇달아 개발됐다.

그림2-1 데이터 사회의 에너지 위기와 무어의 법칙 둔화로 전용 칩 시대가 도래하다



(주) ASIC : Application Specific Integrated Circuit
(출처) T.Kuroda, ISSCC 2010 Panel Discussion, "Semiconductor Industry in 2025"

하지만 15년이 지나고 무어의 법칙에 따라 집적도는 3자리 늘었고 어느새 설

계가 이를 따라갈 수 없게 되었다. 이리하여 전용 칩 시대는 끝났다.

현재 다시 한번 게임 체인지가 일어난 배경에는 에너지 위기가 있다. 폭발적으로 증대하는 데이터를 AI로 분석하기 위해서는 방대한 양의 에너지가 필요하다. 불필요한 회로를 제거하고 범용 칩에 비해 훨씬 많은 에너지를 절약할 수 있는 전용 칩이 요구되고 있다.

AI 처리를 전용 칩(하드웨어)으로 가속하고 다양한 기능은 범용 칩(소프트웨어)으로 처리한다. 즉 두 요소의 적절한 안배가 녹색 성장에 필수적인 것이다.

패러다임 전환

두 번째 변화는 시장의 파도이다.

25년마다 거대한 파도가 반도체 시장으로 밀려온다. 지금이 그때이다.

1970년부터 1995년까지는 가전, 1985년부터 2010년까지는 PC였고, 2000년부터 2025년까지는 스마트폰이다. 일본은 첫 번째 파도는 잘 포착했지만, 두 번째와 세 번째 파도에는 올라타지 못했다. 네 번째 파도에 대한 대비가 중요하다.

가전은 아날로그 기술로 ‘물리 공간’을 더욱 편리하게 한다. 한편 PC는 디지털 기술로 ‘사이버 공간’을 창출하고 스마트폰은 무선 네트워크 기술로 사이버 공간을 휴대할 수 있도록 한다.

지금 밀려오는 네 번째 파도는 센서와 AI, 모터를 활용해 사이버 공간과 물리 공간을 고도로 융합시켜 경제 발전과 사회적 과제 해결을 목표로 한다. 즉 ‘디지털 트윈’을 이용한 인간 중심 사회 ‘Society 5.0’ 창출이다.

예를 들면 자동차나 드론 등 이동 로봇을 포함한 로봇 공학 분야이다.

로봇 공학의 지능 수준은 미래학자 한스 모라벡(Hans Moravec)에 따르면 현재는 쥐 정도 수준이지만 2030년에는 원숭이 수준으로 발전하고 2040년에는 인간 수준에 도달한다. 지능형 로봇이 이동·물류·서비스부터 의료·돌봄·엔터테인먼트까지 일신한다.

바로 이것이 ‘과제 선진국’ 일본이 세계를 선도할 수 있는 시장이다. 일본의 강점인 물리 공간 최적화에도 역량을 발휘할 수 있다.

물론 네 번째 파도는 여기서 멈추지 않는다. 발상 능력이 시험받을 뿐만 아니라 아이디어를 곧바로 칩에 구현할 수 있는 신속(Agile)한 개발 능력이 필요하다.

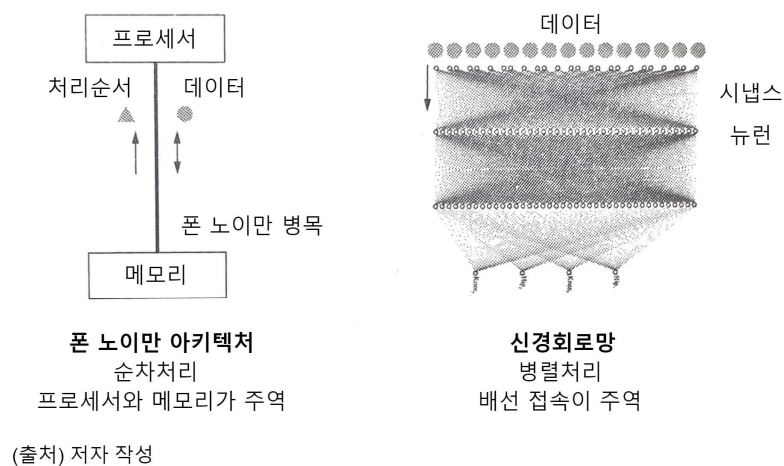
그리고 세 번째 변화는 기술적 패러다임 전환이다.

1950년대 컴퓨터는 연산 장치 사이의 배선을 바꿔 프로그래밍하는 ‘포선 논리 방식’이었다.

이 방식에는 결점이 두 가지 있다. 처리 가능한 프로그램의 최대 규모가 미리 준비한 하드웨어의 규모로 제한되는 ‘규모 제약 문제’와 시스템이 대규모화되면 접속 수가 너무 많아지는 ‘대규모 시스템 접속 문제’이다.

그래서 수학자 폰 노이만은 처리 대상인 데이터와 데이터의 이동 및 연산을 지시하는 명령을 메모리에 기억해 두고, 프로세서가 이 명령을 순차적으로 해석해 연산 처리를 하는 ‘프로그램 내장 방식(폰 노이만 아키텍처)’을 발명했다. 복수 연산기를 준비하고 이들을 물리적으로 배선하는 대신, 하나의 연산기에 사이클마다 다른 명령을 실행시킴으로써 규모 제약 문제를 해결하고자 한 획기적인 방식의 전환이었다.

그림2-2 폰 노이만 구조에서 신경회로망으로



한편 ‘대규모 시스템 접속 문제’를 다양한 각도에서 검토하는 과정에서 발명된 것이 전자기술자 잭 킬비(Jack Kilby)가 1958년 발명한 집적회로(IC)였다. 포

토리소그래피를 이용해 한 장의 칩에 소자를 집적하고 일괄 배선함으로써 이 문제는 훌륭히 해결되었다.

반세기 이상 이어진 이 두 가지 기본 방식에 패러다임 전환이 지금 일어나려 하고 있다.

하나는 폰 노이만 아키텍처에서 신경회로망(Neural Network)으로의 전환이다(그림 2-2).

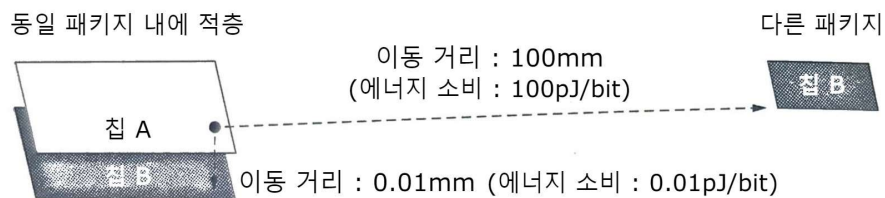
프로세서와 메모리 사이를 데이터가 오가며 차근차근 순차 처리하는 대신 신경회로망에서는 데이터가 막힘없이 흘러 한 번에 병렬 처리된다. 그 결과 에너지 효율이 크게 개선된다.

컴퓨터가 폰 노이만 아키텍처를 기반으로 할 때는 프로세서와 메모리가 대량으로 팔렸지만, 앞으로는 AI 처리용 신경회로망을 탑재한 전용 칩 시장이 발전할 것이다. 주역이 프로세서와 메모리에서 신경회로망의 배선 접속으로 옮겨간다. 이는 마치 뇌간, 소뇌에서 대뇌로 생물이 진화한 것과 유사하다.

인간의 뇌에는 태어날 때 50조 개 정도밖에 없던 시냅스가 초등학교에 들어갈 무렵까지 20배로 증가한다. 그 후 학습을 거듭하며 별로 사용되지 않는 시냅스가 제거되어 최종적으로 낭비 없이 효율적인 뇌 회로가 완성된다. 즉, 미완성 상태로 태어나 놀이를 통해 자라나며 배움을 통해 효율을 향상시키는 것이다.

신경회로망도 이와 비슷한 과정을 거친다. 기계 학습을 통한 제거 방법이 현재 활발히 연구되고 있다.

그림2-3 미세화에서 3D 집적으로



데이터 이동에 필요한 에너지를 3D 집적으로 훨씬 낮출 수 있다

(출처) 저자 작성

또 하나의 패러다임 전환은 미세화에서 3D 집적으로의 전환이다(그림 2-3).

미세화가 이제는 한계에 가까워지고 있다. 3D 집적은 데이터 이동에 필요한 에너지를 훨씬 절감할 수 있다. 이는 마치 국회 도서관까지 가지러 가던 데이터가 손을 뻗으면 닿는 위치에 놓이게 된 것과 같다.

이렇게 우리는 1950년대의 2가지 근원적인 문제에 다시 도전하게 된다. 무어의 법칙이 점차 종말을 맞이하는 가운데 기존 기술의 연장에는 없는 파괴적 기술(disruptive technology)에 상용화 기회가 주어진다.

녹색 성장 전략

지금까지의 고찰에서도 알 수 있듯 다양한 변화의 근원에는 에너지 문제가 있다. 에너지 효율을 높이기 위해 범용 칩에서 전용 칩으로 산업 주역이 교체되고, 폰 노이만 아키텍처에서 신경회로망으로 아키텍처가 쇄신되어 미세화에서 3D 집적으로 기술 체계가 바뀌려 하고 있다.

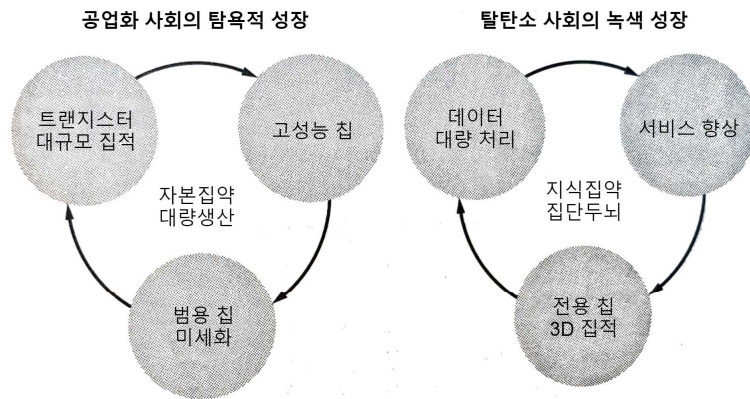
동시에 사회는 자본 집약적 공업화 사회에서 지식 집약적 지가사회로 진화한다. 이제는 트랜지스터를 대규모 집적한 저렴한 칩이 가치를 갖는 것이 아니라 대량의 데이터를 높은 에너지 효율로 처리할 수 있는 능력과 이를 활용해 만드는 뛰어난 서비스로 가치가 이동하는 것이다.

그러나 앞으로 탈탄소(Carbon-free) 규제가 큰 부담이 될 것이다. 에너지 소비를 적극적으로 감축해야 한다. 당연히 지금까지의 탐욕적 성장 전략에서 녹색 성장전략으로의 대전환이 필요하다(그림 2-4).

녹색 성장 전략의 '3개의 축'은 3D 집적의 초크 포인트 기술 창출과 전용 칩을 애자일하게 개발할 수 있는 플랫폼 구축, 그리고 일본 국내에 뿌리내려 군생하는 산업 생태계 보전이다.

에너지 효율 개선 없이 성장 없고, 개발 효율 개선 없이 전용 칩 없다. 즉 앞으로는 시간 대비 성능을 추구하는 것이 최우선 과제가 된다. 물론 시간은 돈이므로 기존의 비용 대비 성능을 내포한다.

그림2-4 탐욕적 성장에서 녹색 성장으로



(출처) 저자 작성

윈스턴 처칠(Winston Churchill) 전 영국 총리는 국난에 대처하며 다음과 같이 말했다.

“눈앞에 닥쳐온 어려움이나 큰 문제에 똑바로 맞서야 합니다. 그렇게 하면 그 어려움과 문제가 생각보다 훨씬 작다는 것을 알 수 있습니다. 하지만 거기서 도망가면 어려움은 2배로 커져 나중에 덮쳐올 것입니다.”

인텔을 창업한 로버트 노이스(Robert Noyce)는 이렇게 말했다.

“혁신을 만들어 내려면 낙천적이어야 합니다. 위험을 두려워하지 말고 변화를 추구하며, 안주하지 말고 모험에 나서야 합니다.”

일본 점유율의 반전은 각오와 낙관과 함께 오늘부터 시작된다.

2. 범용 칩에서 전용 칩으로

—반도체 산업의 게임 체인지

범용 칩 시대와 전용 칩 시대

범용품은 규격화한 제품을 대량 생산을 통해 저렴하게 만들 수 있기 때문에 널리 보급된다. 한편 전용품은 가격이 높지만 우수한 성능과 품질, 신뢰성을 제공할 수 있다.

반도체 비즈니스는 범용 칩이 주역이다. 연간 50조엔 규모의 시장에서 2조

개의 칩이 생산되며 평균 단가는 25엔에 불과하다.

1조 엔을 투자해 지은 최신예 공장에서 출하되는 최첨단 칩도 몇백 엔에 팔리고 있다. 박리다매(薄利多賣) 방식의 비즈니스인 것이다.

범용 칩이 대량으로 팔리는 주된 이유는 **컴퓨터가 폰 노이만 아키텍처를 채택하고 있기 때문이다.**

처리 순서와 데이터를 메모리에서 읽어와 프로세서가 해당 순서에 따라 데이터를 처리한 후 메모리로 돌려보낸다. 이것을 순차적으로 반복하면 아무리 복잡한 처리도 실행할 수 있고 처리 순서 즉 프로그램을 변경하면 어떠한 처리라도 실행할 수 있다.

요컨대 컴퓨터 발전 시나리오는 프로세서와 메모리를 대량생산해 하드웨어를 보급하고 소프트웨어로 다양한 용도에 이용하는 것이며, 반도체 비즈니스의 왕도는 프로세서와 메모리를 저렴하게 대량 공급하는 것이다.

빅데이터 활용이 시작되면 여기에 센서가 추가될 것이다.

이 비즈니스의 경쟁 방식은 자본 경쟁이다. DRAM이나 플래시 메모리 혹은 CPU나 GPU 등의 칩이 발명되고 이것이 커다란 비즈니스가 되리라 인식하면 거대한 자본이 투입되고 금세 과당경쟁이 일어나, 업계 재편 후에 과점 시장이 형성된다.

일본은 디바이스 혁신에서는 이겼지만, 자본 경쟁에서 패했다.

한편으로는 전용 칩이 성공했던 시대도 있었다. 1985년부터 2000년까지 ASIC(특정 용도용 집적회로, Application Specific Integrated Circuit)가 거대한 시장을 형성했다.

프로세서와 메모리를 서로 연결하기 위한 논리회로는 시스템에 따라 다르다. 당초에는 표준 로직 칩을 구성해 구현했지만, 이를 ASIC에 집적함으로써 시스템의 비용과 면적을 절감할 수 있었다.

더불어 컴퓨터를 이용한 설계기술(CAD, 캐드)을 활용해 개발 비용과 시간을 크게 절감할 수 있었던 것이 ASIC로 수익을 낼 수 있었던 큰 요인이다. 복잡한 칩의 경우 100명 이상의 설계자가 1년 이상을 필요로 하지만, CAD를 사용하여 1명의 설계자가 1달 만에 설계할 수 있게 된 것이다.

1980년대에 레이아웃이나 논리를 자동 생성하는 기술이 캘리포니아대학 버클리 캠퍼스를 중심으로 연구 및 개발되었으며, 툴 공급업체도 탄생했다. 더불어 반제작으로 의류를 만드는 것처럼 반환성품 칩을 제조해 두고 마지막에 배선을 커스터마이징하는 세미 커스텀 제조 방식도 개발되었다.

이러한 설계 및 개발 혁신 덕분에 개발 효율은 한 번에 3자리 향상되었다.

하지만 15년 후에는 무어의 법칙으로 집적도가 3자리 늘어났고 컴퓨터를 활용해도 이전보다 더 많은 인력과 시간이 필요하게 되었다. 이 결과 ASIC 비즈니스는 재산이 맞지 않아 종언을 고했다.

이렇듯 범용의 시대는 디바이스 혁신으로 막을 올리고 자본 경쟁 끝에 막을 내린다. 한편 전용의 시대는 설계 및 개발의 혁신으로 막을 올리고 무어의 법칙으로 막을 내린다.

게임 체인지—GAFA가 전용 칩 자사 개발에 나서다

요즘 게임 체인지가 일어나고 있다. 인텔이나 퀄컴 같은 반도체 전문 제조업체에게 범용 칩을 공급받아서 경쟁에서 승리할 수 없다고 느낀 GAFA 등 대형 IT기업들이 전용 칩 자체 개발에 착수했다.

이 배경에는 세 가지 이유가 있다.

첫 번째 이유는 **데이터 사회 특유의 ‘에너지 위기’**이다. 데이터가 급격히 증가하고 AI 처리가 고도화되며 에너지 위기가 가속화되고 있다.

현재 기술로 에너지 절약 대책이 전혀 이루어지지 않는다고 가정하면 2030년에는 현재 총 전력 소비량의 거의 두 배가 IT 관련기기에서 소비되며, 2050년에는 약 200배가 될 것으로 예상된다.

디지털 트랜스포메이션이 막대한 에너지를 소비해 지구 환경을 파괴하게 된다면 지속 가능한 미래는 기대할 수 없다.

칩 소비전력은 예전에는 0.1와트 정도였다. 이상적인 스케일링(미세화) 시나리오에 따르면 전력 밀도를 일정 수준으로 유지하며 성능 대비 비용을 개선할 수 있었을 것이다.

그러나 실제로는 이 이상으로 성능 개선을 우선시한 결과, 전력이 15년 동안

1,000배 증가하여 2000년에는 100와트에 달했다. 칩 전력 밀도는 전기 그릴의 30배를 초과하며 클라우드 서버 냉각에 막대한 전력이 소비된다.

냉각 한계를 초과하면 집적은 가능하지만 동시에 사용할 수 없는 트랜지스터가 증가한다. 7nm(나노미터) 세대에서는 전체의 3/4가, 5nm 세대에서는 전체의 4/5에 해당하는 트랜지스터를 동시에 사용할 수 없다. 이러한 제약 아래서는 에너지 효율을 10배 높인 사람만이 컴퓨터 성능을 10배 높일 수 있고 스마트폰을 10배 오래 사용할 수 있게 된다.

모든 태스크를 수행할 수 있는 범용 칩과 비교하면 불필요한 회로를 제거한 전용 칩은 에너지 효율을 10배 이상 향상할 수 있다.

전용 칩 자체 개발에 착수하는 두 번째 이유는 AI의 등장이다. 신경회로망과 심층학습(딥러닝)은 데이터 소유자에게 정보 처리의 새로운 방식을 제공했다.

신경회로망은 우리 뇌와 마찬가지로 배선 연결이 기능을 부여하는 포선 논리 방식이다. 순차 처리를 하는 폰 노이만 아키텍처에 비해 병렬 처리로 전력 효율을 10배 이상 높일 수 있다.

세 번째 이유는 분업화가 진행된 산업구조이다. TSMC 등의 파운드리가 세계의 공장이 되어 사용자 스스로가 AI 성능을 최대한 끌어낼 수 있도록 비즈니스 모델에 맞는 반도체 칩을 자사 개발할 수 있게 되었다.

대량의 칩을 사용하는 IT 플랫폼 업체라면 자체 개발을 하는 편이 반도체 벤더로부터 조달하는 것보다 빠르고 저렴하며 더 높은 성능의 칩을 조달할 수 있게 된다.

지식 집약적 사회에서의 제조업을 생각하다

예전에 미국의 컴퓨터 과학자 앨런 케이(Alan Kay)가 “소프트웨어를 중요하게 여기는 사람은 스스로 하드웨어를 만들게 된다”고 했다. 시스템 개발에는 하드웨어와 소프트웨어가 모두 필요하다.

다양한 제어가 필요한 논리적이고 계산적인 정보 처리에는 폰 노이만 아키텍처에 따른 범용 칩을 사용하고, 고도 AI가 필요한 직관적이고 공간적인 정보 처리에는 전력 효율이 높은 전용 칩을 사용하는 새로운 아키텍처에 관한 탐구가

시작되고 있다.

물론 범용 칩과 전용 칩은 저렴함과 고성능 사이의 상충 관계는 항상 존재한다.

예를 들면 정보 통신에서는 비교적 수량이 많지 않은 인프라 측에서는 가상화 기술을 활용함으로써 가능한 한 범용 하드웨어로 기능을 구현하려고 한다. 한편 비교적 수량이 많은 엣지 측에서는 전용 칩으로 성능을 높여 데이터를 생산된 곳에서 소비하는 지산지소(地產地消)를 촉진하려고 한다.

전용 칩에 필요한 것은 자본력이 아니라 학술이다. 예전에 캘리포니아대학 버클리 캠퍼스에서 레이아웃이나 논리의 자동 생성 기술을 창출했듯이 **기능이나 시스템을 자동 생성하는 학술의 창출이 필요하다. 대학이 맡은 역할이 커지고 있다.**

20세기는 ‘범용’의 시대였다. 세계대전 이후 물량송배와 경제 효율 예찬 하에 규격화 대량 생산이 경제 성장을 이끌었다.

사회가 성숙해짐에 따라 전체의 성장에서 개인의 성장으로 가치가 변했다. 그 결과 공업 사회는 끝나고 지식사회가 시작되었다.

이 변화가 선진국에서 개발도상국으로 퍼지는 과정에서 일본은 규격화 대량 생산을 계속하며 일시적으로 번영했지만 결국 아시아 국가들에 우위를 내어주게 되었다.

이번 세기는 ‘전용’의 시대가 될 것이다. 자본 집약에서 지식 집약으로, 규모에서 지혜로, 양적 확대에서 질적 발전으로, 물질에서 정신으로, 편리함에서 즐거움으로, 제품에서 서비스로, 대량에서 다양으로, 획일에서 개성으로, 누구나 할 수 있는 것에서 다른 사람은 할 수 없는 것으로 가치는 옮겨 갈 것이다.

그때 제조업은 어떻게 변해있을까? 이 질문에 대한 답을 찾는 것이 우리의 사명이다.

3. 산업의 쌀에서 사회의 뉴런으로 —포스트 코로나 시대의 반도체

방대한 에너지를 소비하는 리모트 사회

미국에 사는 친구가 숲속에 집을 짓고 원격 근무를 한다. EDA(전자 설계 자동화) 툴 개발 일을 하고 있기 때문에 컴퓨터와 인터넷이 있다면 어디에서든 일할 수 있으리라 생각했다. 그런데…….

코로나19 확산은 원격 사회로의 문을 열었다. 온라인 회의가 생각보다 유용하다는 것을 알게 되었으며, 3명 이상이 모여 대화하기에 편리하다.

3,000명이 모이는 국제회의도 온라인으로 전환되었다.

우연히도 필자는 2005년에 교토에서 열린 반도체 기술 학회인 VLSI 심포지엄의 저녁 만찬회에서 프로그램 위원장을 맡아 다음과 같이 인사말을 했다.

“여러분, 상상해 보십시오. 미래에 우리들은 인터넷을 통해 국제회의를 개최하게 될지도 모릅니다.”

“연구 발표와 토론회도, 복도에서의 짧은 대화도 온라인으로 진행될 것입니다. 여러분들께서는 자택에서 참가하시게 될 것입니다.”

“저녁 만찬도요……?”

“피자를 배달시켜서 냉장고에서 맥주를 꺼내고……. 별로 기대되진 않네요.”

“오늘 밤은 교토 요리와 사케를 음미하며 옛 친구들과 마음을 나누며 회담을 즐겨주십시오.”

“건배!”

지금 국제회의 주최자들은 그야말로 판도라의 상자가 열린 것 같으며 우려스러워한다. 온라인 회식까지 등장했기 때문이다.

디지털 트랜스포메이션이나 데이터 구동형 서비스의 기반이 되는 것은 빅데이터의 급증과 AI 처리의 고도화이다. 그리고 이런 요소들이 사회의 에너지 소비를 폭발적으로 증가시킨다.

앞서 언급했듯이 **2030년에는 현재의 총 전력 소비량의 두 배 가까운 전력이 IT 관련기기에서만 소비될 것으로 예측된다. 나아가 2050년에는 총 전력 소비**

량이 현재의 약 200배에 달할 것이라는 예측도 있다.

그 이유 중 하나가 통신 데이터의 급증이다. IP 트래픽이 2016년에는 연간 4.7ZB(제타바이트) 였지만 2030년에는 4배 늘어난 17ZB로, 2050년에는 4,000 배에 해당하는 2만 200ZB로 늘어난다고 한다. ZB는 10의 21승 바이트를 의미한다. 2050년에는 4GB(기가바이트) DRAM 칩이 실제로 5,000조 개나 필요할 것이다.

정보의 자산지소는 에너지 효율적인 관점에서 중요하지만 한편에서는 거대 IT 기업들에 의한 정보의 집적과 독점이 진행될 것이다.

여기에 더해 AI 처리가 고도화되고 있다. 데이터에 숨겨진 의미를 이해하고 이를 서비스로 승화시켜 사회에 도움이 되려면 방대한 양의 계산이 필요하다.

실제로 심층학습의 등장 이후 AI 처리의 계산량은 10년 동안 4자리 증가했다. 반면 이를 처리하는 범용 프로세서의 전력 효율은 1자리밖에 개선되지 않았다.

즉 통신 기기나 컴퓨터의 에너지 효율을 현격하게 개선하지 않으면 사회의 지속 가능한 성장은 기대하기 어렵다.

소비 에너지가 급증한 원인은 반도체에 있다. 그리고 그 해결책도 반도체가 쥐고 있는 것이다.

산업의 쌀에서 사회의 뉴런으로

2019년에는 전 세계에서 반도체 칩 1조 9,000억 개가 생산되었다.

시장의 세부 내용을 보면 제조업이 15%, 헬스케어가 15%, 보험이 11%, 은행 및 증권이 10%, 도매·소매가 8%, 컴퓨터가 8%, 정부가 7%, 교통이 6%, 공공사업이 5%, 부동산·업무 서비스가 4%, 농업이 4%, 통신이 3%, 기타 4%를 차지한다.

실로 사회의 구석구석까지 반도체가 쓰이고 있다. 한편으로 놀라는 분들이 계실 수도 있겠지만 통신 시장은 아직 작다.

그러나 방금 언급했던 것처럼 가까운 미래에는 통신량이 폭발적으로 증가할 것이다. 차세대 반도체 수요를 견인하는 것은 차세대 통신 ‘포스트 5G’이다.

포스트 5G는 높은 주파수를 이용한다. 주파수가 높아질수록 전파는 직진성이 강해지지만 먼 거리까지 도달하지 못한다. 따라서 더 많은 기지국이 필요하게 된다.

게다가 지연 시간이 짧아 고도의 서비스를 제공하게 될 것이다. 즉 고성능 데이터 처리가 기지국에 필요하게 된다.

포스트 5G가 차세대 반도체 수요를 주도할 것으로 보는 이유는 이러한 상황이 예상되기 때문이다.

앞으로는 사물 인터넷(IoT), 원격의료 등 디지털 의료·헬스케어, 그리고 모빌리티를 포함한 서비스가 반도체 분야의 거대한 시장을 형성할 것이다. 이러한 것들은 사회의 신경계라고 할 수 있다.

말하자면, 반도체는 산업의 쌀에서 사회의 뉴런으로 발전한다고 할 수 있다. 반도체는 그야말로 인류 공유재산인 것이다.

공업 사회의 부품에서 지식사회를 지탱하는 전략 물자로 성장함으로써 반도체의 가치 지표는 비용에서 성능, 특히 전력 성능으로 이동하게 될 것이다. 이에 더해 인프라에 사용되므로 타임 투 마켓과 신뢰성도 중요해질 것이다.

사회의 에너지 문제를 해결하려면 반도체 에너지 효율을 향상해야만 한다. 전용 칩을 사용함으로써 범용 칩보다 두 자리 정도 전력효율을 높일 수 있다. 왜냐하면 사용자와 이용 시나리오가 명확한 전용 칩에는, 불특정 사용자의 다양한 요구에 응하기 위한 범용성이나 계속성에서 기인하는 낭비가 없기 때문이다.

그러나 전용 칩은 개발 비용이 많이 들고 누구나 만들 수 있는 것이 아니다. 그래서 전용 칩을 개발하는 일본 국내 분위기가 수그러들며 공동화가 시작되고 있다.

그렇기에 **전용 반도체의 개발 비용을 1/10로 낮추고 시스템 아이디어를 가진 사람은 누구라도 전용 칩을 설계할 수 있도록 하고, 나아가 최첨단 반도체 기술을 사용하여 에너지 소비를 1/10로 줄이는 것이 데이터 구동 사회를 실현하기 위해 필수적이다.**

반도체가 산업의 쌀에서 사회의 뉴런으로 진화하기 위해, 산업 구조를 지난 세기의 자본 집약적 구조에서 이번 세기의 지식 집약적 구조로 변혁시키지 않

으면 안 된다.

디지털 문명을 만들기 위해서는

『사피엔스』¹⁾ 저자 유발 노아 하라리(Yuval Noah Harari)는 기술이 인간 스파이를 대신하고, ‘피부 아래의 정보’도 그대로 드러났다면 경종을 울린다.

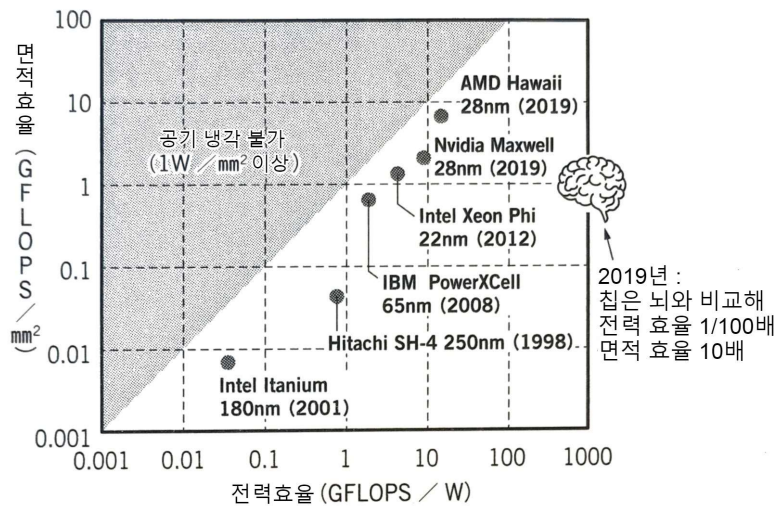
코로나19 확산 방지 대책 속에서 감시 사회가 형성되고 있다. 기술이 사회에 미치는 영향은 매우 커졌다. 우리 문명은 어떻게 되는 걸까? 그야말로 벼랑 끝에 서 있다는 의견도 들려온다.

지혜가 있다면 기술이 이를 구현할 수 있다. 보안과 개인 정보를 위협하는 것이 반도체라면 이를 해결할 수 있는 것도 반도체이다.

하지만 고도의 보안과 개인 정보 보호는 당연히도 반도체의 에너지 소비를 늘린다. 결과적으로 반도체 에너지 문제로 귀결된다.

그리고 그 너머에는 ‘마음’의 문제가 있다.

그림2-5 반도체 전력 효율을 20년 동안 3자리 개선해
2030년에는 뇌의 전력 효율에 접근한다



(출처) 저자 작성

1) 역주. 한국어 번역서 제목. 원제는 『Sapiens: A Brief History of Humankind』이며 일본에서는 『サピエンス全史』(河出書房新社)로 출판됨.

디지털은 논리적인 면에서는 우수하지만 감성은 아날로그이다. 디지털을 통해 사람들을 행복하게 하는 연구도 앞으로 시작될 것이다.

오감과 디지털을 상호 변환하는 센서와 액추에이터, 감각을 피드백하는 제어 기술, 가치를 교환하는 공학, 기술이 사회를 위협하지 않는 법체계, 이러한 것들에 대한 논의 없이 ‘뇌를 인터넷에 연결’하는 것을 추진해서는 안 된다.

태고의 옛날, 뇌는 사회를 만들고 마음을 탄생시켰다. 사람들은 자신의 의도를 알고 이를 전달하는 언어와 논리적 사고를 통해 인지 능력을 확장시키는 수학을 습득했다. 수학은 결국 주관적 직감을 초월하는 추상적인 기호 체계로 승화되었으며 **마침내 두뇌에서 넘쳐나와 컴퓨터가 탄생했다. 컴퓨터는 칩을 만들어 내고, 칩은 스케일링을 통해 지수함수적으로 성장하며 컴퓨터를 소형화했다. 그리고 마침내 극소형 컴퓨터가 다시 우리 몸 안으로 되돌아오려 하고 있다(그림 2-5).**

4. 댐에서 반도체로—디지털 사회의 인프라

햏타 댐과 TSMC

지금으로부터 100년 전, 대만에 우산터우(烏山頭) 댐이 건설되었다. 미국 후버(Hoover) 댐이 완성되기 전까지 이 댐은 세계에서 가장 큰 댐이었다. 건설을 감독한 사람은 일본의 토목 기술자 햏타 요이치(八田与一)였다. 그의 공로를 기리는 햏타 댐이라는 이름으로도 알려져 있다.

햏타 요이치는 1910년에 도쿄제국대학 공학부 토목과를 졸업하고 대만 총독부 토목과 기술자로 부임했다. 그리고 대만 남부에 펼쳐진 불모의 대지, 자난(嘉南)평원을 조사했다. 이 지역은 관개 시설이 부족해 농민들은 평소에 가뭄과 폭우 및 배수 불량으로 인해 힘들어했다.

햏타는 이수 사업을 통해 이 황량한 지역을 곡창 지대로 바꾸자는 제안을 했고 이는 국회에서 승인되었다. 수익자가 조합을 결성하여 사업을 시행했고 비용의 절반은 국비로 충당되었다. 햏타는 국가 공무원이라는 직업을 버리고 조합

소속 기술자가 되어 댐의 건설을 진두지휘했다.

대만의 핫타 요이치 동상



(사진 제공 요미우리신문/AFLO)

총공사비는 5,400만 엔이었다. 당시로서는 일본 역사상 유례없는 대규모 공사였다. 상류의 우산링(烏山嶺)을 3,078m나 파내고 본류인 청원강(淸文溪)에서 댐으로 물을 끌어오는 공사에서는 많은 희생자가 발생했다. 핫타는 이 대규모 공사가 끝날 때까지 공사 현장에 지은 30평 정도의 소박한 일본 가옥에서 아내와 8명의 자녀와 살았다.

10년이라는 세월 끝에 핫타 댐이 완성되었다.

또한 자난평원 일대에 16,000km에 걸쳐 수로가 세밀하게 조성되었다. 이는 물의 만리장성이라고 할 수 있다. 만리장성의 길이는 2,700km밖에 되지 않기 때문에 수로의 길이는

이를 훨씬 넘어선다.

이 수로에서 물이 흘러나오기 시작하자 현지 농민 60만 명은

“신의 물이 왔다.”

라고 기뻐하며 감격의 눈물을 흘렸다. 자난평원은 그 후 크게 풍요로워졌다고 한다.

자난에는 지금도 핫타 동상이 남아 있으며 전쟁 중에도 지역 주민들에 의해 소중히 보존되었다. 자난 지역의 농민들이 이 동상 앞을 지날 때 언제부터인가 손을 모아 합장을 하는 모습이 보이기 시작했다. 핫타의 기일인 5월 8일에는 자난 사람들이 요이치와 부인의 무덤을 찾아가 무덤 앞에서 제를 올린다. 핫타는 대만을 사랑한 일본인이며 자난대수(嘉南大圳)의 아버지로서 대만에서 사랑받는 일본인이 되었다.

핫타 댐 건설로부터 100년이 지난 지금, 일본과 대만이 협력하여 역사적인 대규모 사업이 다시 시작되었다.

이번에는 댐 대신 반도체이다.

흙과 모래는 고순도의 실리콘으로 바뀐다.

물은 데이터로, 물길 이용은 데이터 이용으로 바뀐다.

사회는 농경사회 Society 2.0에서 인간 중심 사회 Society 5.0으로 진화하며 물의 관리장성은 데이터의 관리장성으로 바뀐다.

작은 칩 안에 무수한 배선이 집적되어 있고 여기를 데이터가 이동한다. 이 배선을 칩에서 빼내면 총길이는 10km에 달한다.

칩이 1,000개 정도 나열된 웨이퍼를 TSMC 공장에서는 매달 수백만 장 제조한다. 만약 모든 칩의 배선을 모두 연결한다면 총길이는 10억 km에 달해 데이터가 지구를 2만 5천 바퀴나 돌 수 있는 길이이다.

일본 구마모토현 기쿠요마치(菊陽町). 20헥타르를 넘는 방대한 부지에 많은 건설 크레인들이 늘어서 있다. TSMC 구마모토 공장의 건설 현장이다. 신규 공장에서는 1,700명이 일하게 될 것이다. 그중 300명은 TSMC에서, 200명은 소니 그룹에서 파견하며 나머지는 신규 채용할 예정이다. 현지에서는 인재 확보 경쟁이 시작되어 기술자들의 임금 수준이 올랐다.

이 공장에서는 28nm 및 22nm 로직 반도체를 제조한다. 일본에서 가장 필요로 하고 생산량이 많은 볼륨 존에 해당하는 반도체이다. 나중에는 더 미세한 세대로 수요가 이동하는 것에 맞춰 16nm 및 12nm FinFET을 제조할 예정이다. 일본은 28nm 이후 세대에 대한 반도체 투자를 계속할 수 없었다. 그야말로 ‘신의 데이터가 온다’며 기뻐하게 될 것이다.

뿐만 아니라 TSMC는 이바라키현(茨城県) 쓰쿠바시에 3D IC 연구개발 센터를 개소했다. 칩을 3차원으로 구현해 데이터 이동 거리를 줄이기 위해서는 일본 소재의 힘이 필요하다. 우리는 새로운 소재를 발굴하고 3D IC 연구개발 센터와 연계해 일본의 소재를 활용할 길을 모색할 것이다.

디지털 사회의 인프라

인프라라고 하면 도로, 항만, 철도, 공항 등 교통 인프라와 상하수도 같은 도시 인프라, 그리고 발전·송전과 같은 에너지 인프라가 있다. 이것들은 20세기

인프라이다. 21세기 인프라는 반도체와 이를 이용한 고도의 컴퓨팅 및 통신망이 될 것이다.

일본은 세계대전 후 경제 부흥기에 자본집약형 공업 사회 및 공업 입국으로써 성공을 거머쥐었다. 이윽고 대량 생산·대량 소비를 통한 성장의 한계가 떠오르면서 앞으로 지향해야 할 사회는 공업사회나 정보 사회를 넘어선 인간 중심의 사회가 되어야 한다고 여기게 되었다. 이는 지식 집약적 사회이며 데이터를 활용해 지혜를 나누는 사회이다. 즉 자본 집약적 공업사회에서 지식 집약적 지식 사회로의 전환이다.

자본 집약적 사회에서 지식 집약적 사회로 패러다임 전환이 일어나면 산업 구조도 변한다. 자본 집약적 사회에서는 물질이 가치를 창출한다. 재료가 자원이고 이를 조합해 부품을 만들고 부품을 모아 제품을 만들었다. 반도체 칩은 부품이며 제품에 속하는 서비스나 디자인과 같은 지식과 정보가 사용자를 만족시키고 가치를 창출한 것이다.

그러나 지식 집약적 사회로 바뀌면 가치 창출의 주객이 전도된다.

가치는 물질에서 지식이나 정보로 이동한다. 자원은 재료에서 데이터로 바뀐다. IoT로 모은 데이터를 AI로 분석해 이를 서비스 솔루션으로 만들어 사용자에게 전달한다. 이를 처리하는 반도체는 단말기 배터리의 수명을 늘리거나 처리를 빠르게 해서 사용자를 만족시켜 가치를 창출할 것이다.

따라서 재료를 운송하는 도로, 항만, 철도, 공항을 대신해 데이터를 전송하는 IoT나 5G, AI가 디지털 사회의 인프라가 될 것이다.

반도체가 부품 사업이었을 때에는 비용이 최우선이었다. 규격화된 유사한 부품이라면 조금이라도 저렴한 쪽을 선호했다. 하지만 미세화가 어려워지면서 성능 면에서도 경쟁하게 되었다. **최근 자주 사용되는 지표로 PPAC(전력 소모 Power·성능 Perfomance·집적도 Area·비용 Cost)가 있다. 바꿔 말하자면 비용 대비 성능이다.**

디지털 사회의 인프라를 지탱하는 반도체에는 시간 대비 성능을 요구하게 될 것이다. 인프라 시장은 교체 수요가 적고 먼저 시장에 투입된 제품이 장기간 사용되기 때문이다. **타임 투 마켓이 추가된 PPACT가 지표가 된다.**

일본의 20세기 인프라는 우수했다. 도로는 어딜 가든 포장되어 있고 대중교통

은 정시에 운행된다. 21세기의 우수한 인프라는 일본 국내 어디서든 고속 무선망에 연결되어 고도의 컴퓨팅 자원을 이용할 수 있게 될 것이다.

1929년에 일어난 대공황을 극복하기 위해 미국은 뉴딜 정책을 펼쳐 다목적 댐 건설 등의 공공사업을 진행하여 인프라를 정비했다. 반도체는 디지털 사회의 인프라를 지탱하는 기반 기술이다. 지금 우리 사회에는 말하자면 디지털 뉴딜 정책이 필요한 것은 아닐까.

히로이 이사미의 가르침

다시 100년 전 이야기로 돌아가자.

당시 도쿄제국대학에서 토목공학을 가르치던 히로이 이사미(広井勇)는 ‘무엇을 위해 공학이 존재하는가’에 대해 다음과 같이 말했다.

“만약 공학이 단순히 인생을 번잡하게만 한다면 아무 의미도 없다. 공학을 통

히로이 이사미 동상



(사진 제공 ImageMart)

해 며칠이 걸리는 곳을 몇 시간 거리로 줄이고 하루 노동을 1시간으로 하고, 거기에서 얻은 시간으로 조용히 삶을 사유하고 반성하며 신에게 돌아갈 여유를 얻지 못한다면, 우리 공학에서 아무런 의미도 찾을 수 없을 것이다.” (다카하시 유타카(高橋裕) 『현대 일본 토목사(現代日本土木史)』 쇼코쿠샤(彰国社)에서 발췌)

핫타 요이치도 틀림없이 히로이 이사미의 가르침을 받았을 것이다.

지금 다시 한번 핫타의 정신과 히로이의 가르침을 생각해 보지 않을 수 없다. 일본과 대만의 민족과 국경을 초월한 역사적 사업이 구마모토 쓰쿠바시에서 시작된다. 필자는 큰 기대감으로 가슴이 부풀어 오르고 경외감으로 온몸이 떨리며 마음 깊은 곳에서 용기가 넘쳐 난다.

Ⅲ. 구조개혁(構造改革) More Moore

1. 뇌와 컴퓨터와 집적회로의 짧은 역사

—그리고 하나의 미래

뇌와 컴퓨터와 집적회로의 탄생

138억 년 전 거대한 에너지 덩어리가 돌연 나타났다. 빅뱅이다.

에너지와 물질이 상호 작용하며($E=mc^2$) 우주는 급속히 팽창했다. 최초의 미세한 요동이 은하계를 만들고 46억 년 전에 지구가 탄생했다.

물리 법칙에 따라 물질이 변화하는 가운데, 자신의 구조를 정보로 만들어 DNA에 보존해 자가 복제하는 생명체가 등장한 것이 40억 년 전이다.

생명체는 돌연변이와 적자생존을 전술 삼아 불확실한 환경에서 살아남으며 단세포에서 다세포, 식물, 동물로 진화하며 다양해졌다.

동물은 외부에서 정보를 얻어 행동을 결정하기 위한 중추신경계인 뇌를 갖게 되었다. 그리고 포유류는 700만 년 전에 인류로 분화해 뇌를 진화시켰다.

생존을 위해서는 상호 협력이 필요하다. 뇌는 사회를 만들고 마음을 만들어 냈다. 사람은 자신의 의도를 알고 이를 전달하기 위한 언어와 논리적 사고력을 습득한 것이다.

수학이 탄생한 것은 3,000년 전이다.

수학은 인간의 인지 능력을 확장했다. 4대 문명에서는 계산기나 피타고라스의 정리를 이용해 세금 계산이나 토지 측량을 했다. 이윽고 기원전 5세기 고대 그리스 시대가 되자 계산보다는 수학의 내부 세계가 연구 대상이 되어 수학은 도구에서 사고로 진화했다.

7세기 아라비아에서는 대수학이 발달하고 15세기 르네상스 시대에는 기호 대수가 발명되면서 수학은 물리적 제약을 받지 않는 보편적인 입지를 획득했다. 그리고 17세기가 되자 미적분이 고안되어 무한의 세계를 탐구할 수 있게 되었다. 극한이나 연속성 개념을 엄밀하게 고찰한 결과 주관적 직감을 초월한 추상

적인 기호 체계가 만들어졌다.

20세기 들어 ‘수학을 하는 스스로의 사고에 대해 수학을 하는’ 시도마저 이루어진다. 물리적 직감이나 주관적 감각 같은 모호한 것들을 완전히 벗어나고 **뇌에서 뺀어 나온 수학은 마침내 ‘계산하는 기계’로서의 컴퓨터를 탄생시켰다.**

최초의 컴퓨터는 제2장에서도 언급했듯 연산장치 사이에 배선을 바꿔 프로그래밍하는 ‘포선 논리 방식’이었다.

그림3-1 칩 스케일링을 통해 컴퓨터는 소형화를 거치고 두 요소는 나란히 발전했다



(출처) 저자 작성

제2장에서 언급한 대로 이 방식에는 두 가지 결점이 있었다. 처리할 수 있는 프로그램의 최대 규모가 미리 준비한 하드웨어의 규모로 제한되는 ‘규모 제약 문제’와 시스템의 규모가 커지면 접속 수가 방대해지는 ‘대규모 시스템 접속 문제’이다. 이 문제를 해결한 것이 폰 노이만과 잭 킬비이다.

단순화·극소화된 연산 자원을 칩에 집적화·병렬화함으로써 컴퓨터의 성능은 비약적으로 향상됐다. 고성능 컴퓨터는 더 큰 규모의 집적회로 설계를 가능하게 했다. 무어의 법칙에 이끌려 컴퓨터와 집적회로는 함께 발전했다(그림 3-1).

집적회로의 성장과 한계

집적 회로의 성능 대비 비용은 미세화를 통해 지수함수적으로 개선할 수 있다. ‘무어의 법칙’이라고 불리는 이러한 경험에 기반한 법칙은 집적 회로의 지도 원리이자 성장 시나리오이기도 하다.

비용은 리소그래피에 의해 결정되는데, 리소그래피 기술이 미세화 한계에 가

까워지면 포토마스크를 여러 장 조합하는 등 공정을 복잡하게 만들어 미세화를 달성한다. 결과적으로 비용이 상승하면 트랜지스터 단가가 오른다. 실제로 16nm 세대(2015년)부터 트랜지스터 단가가 상승으로 접어들고 있다.

하지만 7nm 세대(2019년)부터 EUV(극자외선) 리소그래피가 도입되어 트랜지스터 가격은 다시 내려갈 것으로 예상된다. 공정이 다시 단순화되고 제조비용이 하락하기 때문이다.

따라서 최근의 문제는 비용이 아니라 성능 개선의 한계이다. 전력, 즉 발열이 상한에 도달하고 이로 인해 회로를 아무리 집적해도 성능을 더 이상 끌어올리지 못하게 되었다는 점이 시급한 과제이다.

전력당 처리 성능, 즉 전력 효율이 무어의 법칙의 명운을 쥐고 있다. **‘전력 효율 개선 없이 성능 개선 없다’**는 것이다.

전력은 스케일링(미세화)의 부작용으로 증가한다. 사실 전계효과(電界効果)로 동작하는 트랜지스터의 전계가 일정하게 유지되도록 디바이스를 스케일링하면 전력은 증가하지 않아야 한다.

하지만 실제로는 1980년부터 1990년대 중반에 걸쳐 회로를 고속 동작시키기 위해 전원전압을 낮추지 않고 디바이스를 스케일링했다. 그 결과 전력은 3년마다 4배 늘어 15년 동안 3자리 증가했다.

전력이 너무 커져서 1995년 이후로 전원전압을 낮춰왔지만 이미 디바이스 내부 전계가 너무 높아졌기 때문에 전류는 충분히 감소되지 않았으며 그 이후에도 전력은 6년 동안 2배씩 증가하고 있다.

전력 증가의 원인이 스케일링의 부작용이므로 대책을 마련하기는 쉽지 않다. 원점으로 돌아가 생각해야 한다.

전력 저감 방안은 세 가지 있다. 저전압화(V), 저용량화(C), 스위칭 저감(fa)이다.

전압을 낮추면 전력이 효과적으로 줄어들지만 한계가 있다. 문제는 누전(leak)이다.

게이트 절연막을 얇게 형성하지 않고 트랜지스터를 미세화하면 트랜지스터의 On·Off를 제어하는 게이트의 작용이 저하되어 트랜지스터가 충분히 Off되지 않

게 된다.

그 결과 전원전압을 더 낮추어도 회로가 느려지는 만큼 누전이 증가하고 지배적으로 작용해 전력은 오히려 증대된다. 오늘날 사용되고 있는 프로세서의 전력 효율이 최대가 되는 것은 전원전압이 대략 0.45볼트일 때이다.

누전을 줄이기 위해 재료, 프로세스 및 구조를 변경했다. 예를 들어 트랜지스터를 입체 구조로 만들어 게이트를 덮어 게이트의 지배력을 개선했다. 7nm 세대 FinFET은 예상 이상으로 누전 절감에 성공했다.

범용에서 전용으로, 2D에서 3D로

상온에서 CMOS 게이트를 다단 접속할 수 있는 이론적 한계는 0.036볼트이다. 저전압화 방안도 앞으로 1자리, 전력 환산 시 2자리 여지밖에 남지 않았다.

전력 효율을 개선하는 또 하나의 방안은 저용량화이다. **범용 CPU나 GPU와 비교하여 ASIC(특정 용도용 집적회로)나 SoC(하나의 통합된 시스템이 담겨 있는 칩, System-on Chip)와 같은 전용 칩은 불필요한 회로를 제거해 저용량화가 가능하며 전력 효율을 10배 이상 높일 수 있다.**

또한 데이터 이동은 계산에 비해 많은 전력을 소비한다. 특히 칩 외부로 데이터를 입출력하는 경우 3자리 정도 많은 전력을 소비한다. 폰 노이만 아키텍처에서 필요로 하는 DRAM의 입·출력이 전력의 병목 지점이 된다.

칩의 데이터 접속에서 중요한 점은 접속 경계를 선이 아닌 면으로 만드는 것이다. 칩 내부는 스케일링 비율의 제곱이 되므로 고집적을 실현할 수 있다. 한편 외부 기기와의 접속을 위한 입출력 장치는 주로 칩 주변부에 배치되므로 집적도는 스케일링 비율에 비례한다. 그 결과 데이터 통신이 내부의 성능 요구사항을 따라잡지 못한다.

이 문제를 해결하기 위해서는 칩을 적층하여 면 전체를 사용해 연결하는 것이 유효하다. **적층 레벨을 2D(평면)에서 3D(입체)로 발전시킴으로써 전력 효율을 크게 향상시킬 수 있는 것이다.**

무어의 법칙이 둔화하는 가운데 기존 기술의 연장이 아닌 신기술(파괴적 기술) 분야에도 실용화 기회가 많아지고 있다.

2. 스케일링 시나리오—지수함수의 경이

이상적인 스케일링 시나리오

집적회로 발전을 위한 기본 원리는 디바이스 미세화, 즉 스케일링하는 것이다. 집적도를 높여 칩 제조 비용을 낮추고 성능을 향상한다.

DRAM은 3년에 4배씩, 프로세서는 2년에 2배씩 집적도가 높아졌다. 이러한 경험을 토대로 한 법칙은 ‘무어의 법칙’으로 널리 알려져 있다.

칩 제조 비용은 웨이퍼 1장당 제조 비용을 1장의 웨이퍼에서 얻을 수 있는 양품(良品) 칩의 수로 나눈 값이다.

리소그래피와 프로세스 기술을 발전시켜 디바이스를 스케일링하고 동시에 웨이퍼 직경을 크게 하거나 제조 기술을 개선해 수율을 높이고 양품 칩의 수를 늘린다.

지난 50년을 되돌아보면 2년마다 디바이스는 20% 미세화되고 칩 크기는 14% 커졌다. 그 결과 집적할 수 있는 디바이스 수는 2년마다 배로($=1.14^2/0.8^2$) 증가했다.

DRAM에서는 여기에 디바이스를 3차원 구조로 만들거나 회로를 고안함으로써 3년에 4배라는 고집적화를 이루어 냈다. 다만 이러한 노력은 한계에 다다라 DRAM의 스케일링이 조만간 중단될 것이라고도 한다.

다음으로 성능이 어떻게 달라지는지에 대해 논의해 보자. **디바이스의 치수나 전압 모두 $1/\alpha$ 로 작게 스케일링**하면 트랜지스터 내부의 전계는 일정하게 유지된다. 이 ‘**일정 전계 스케일링**’에 의해 전계효과 트랜지스터는 스케일링 전후로 동일한 작동이 보장된다.

디바이스 치수가 $1/\alpha$ 가 될 때 트랜지스터를 흐르는 전류와 용량도 마찬가지로 $1/\alpha$ 가 된다. 왜냐하면 전류는 디바이스 치수와 비례해 $1/\alpha$ 가 되고 용량은

면적÷거리로 구하는데 면적이 $1/\alpha^2$ 가 되므로 용량은 $1/\alpha$ 가 된다.

전압, 전류, 용량이 각각 $1/\alpha$ 가 되면 회로의 지연 시간도 $1/\alpha$ 가 된다. 회로의 지연 시간은 용량×전압÷전류로 결정되기 때문이다.

여기서 단위 면적당 전력인 전력 밀도를 계산하면 전압×전류÷면적으로 계산되므로 $1/\alpha$ 로 스케일링해도 변하지 않는다. **집적도가 높아지면 열을 내보내기 어려워지는 듯 하지만 전력 밀도는 일정하며 발열량도 거의 비례하기 때문에 방열 문제는 발생하지 않는다. 정말 이상적인 시나리오다.**

실제 스케일링과 그 부작용

그러나 이상대로 일이 진행되지는 않았다.

마이크로프로세서의 동작 주파수는 10년 동안 약 50배 빨라졌다. 그중 13배가 스케일링으로 인한 효과이며 나머지 4배는 아키텍처에 의한 개선이다.

환산하면 동작 속도는 2년 동안 1.6 배씩 빨라진 것이다. 일정 전계 스케일링 법칙에 따르면 1.2 배이므로 상당히 속도가 빨라진 것을 알 수 있다.

실은 1995년까지 전원전압을 낮추지 않고 디바이스를 스케일링 했다. 즉 ‘일정 전계’가 아닌 ‘일정 전압’으로 스케일링 했던 것이다.

이 경우 전류는 α 배로 증가하고 용량은 $1/\alpha$ 로 작아지므로 회로 지연 시간은 $1/\alpha^2$ 로 감소하고 회로는 더욱 빠르게 동작한다. 그러나 전력 밀도는 α^3 으로 급격히 증가하게 되어 발열량도 비례해서 늘어난다.

이유는 처리 성능이 높을수록 칩이 잘 팔렸기 때문이다. 한편 칩 전력은 애초에 충분히 작았기 때문에 전력 증가는 그다지 큰 문제가 아니었다.

1980년부터 1995년까지 15년 동안 칩 전력은 1,000배 늘었다. 그 결과 단위 면적당 발열량은 전기 그릴의 30배에 달하게 되었다.

열을 방출할 수 없으면 디바이스 내부 온도가 높아져 신뢰성이 훼손된다. 전력 한계에 부딪히면 회로는 더 이상 집적할 수 없게 된다.

이렇듯 **전력 한계의 원인은 공격적인 스케일링의 부작용이었다.**

1995년 이후 전원전압은 점차 낮아졌다.

당연한 일이지만 회로를 사용하지 않을 때는 전원을 수시로 끄거나 높은 성능이 필요하지 않을 때는 전원전압을 낮추는 등 전력을 절약하려는 세심한 노력도 거듭해 왔다.

이러한 것들은 일상생활에서 행해지는 당연한 절약처럼 들릴 수 있지만 트랜지스터를 1억 개 이상 집적한 대규모 집적회로라면 낭비를 인지하는 것부터가 쉽지 않다.

전원전압의 이론적 하한치는 상온의 경우 0.036볼트이다. 이 이하가 되면 CMOS 회로의 이득이 1 미만으로 떨어져 디지털 회로를 다단으로 연결할 수 없게 된다.

하지만 실제로는 꺼져 있는 트랜지스터의 누설 전류나 디바이스의 편차, 노이즈 등이 있기 때문에 0.45볼트 이하로 낮추기는 매우 어렵다.

28nm 세대 이후는 집적은 가능하지만 동시에 사용할 수 없는 트랜지스터 즉 ‘다크 실리콘(전원을 공급하지 못해 작동하지 않는 트랜지스터)’이 급증하고 있다. 기능은 집적할 수 있어도 성능을 끌어내기 어려워지고 있는 것이다.

따라서 지금은 전력 효율을 개선할 수 있는 사람만이 성능을 개선할 수 있는 단계에 들어섰다고 할 수 있다. 그야말로 ‘**전력 효율 개선 없이 성능 개선 없다**’는 것이다.

전원전압을 낮추는 것 외에도 전력 효율을 개선하는 방법은 용량 C를 절감하는 것이다. 이를 위해 **칩을 적층해 3차원으로 집적하는 기술이 향후 집적회로의 명운을 쥐는**. 즉 **집적 레벨을 2D에서 3D로 확장하는 것이다. 왜냐하면 칩의 두께는 칩의 너비에 비해 세 자리나 작기 때문에 칩을 3차원으로 적층하면 칩 사이의 접속 거리를 훨씬 단축할 수 있고 용량 절감으로도 이어지기 때문이다.**

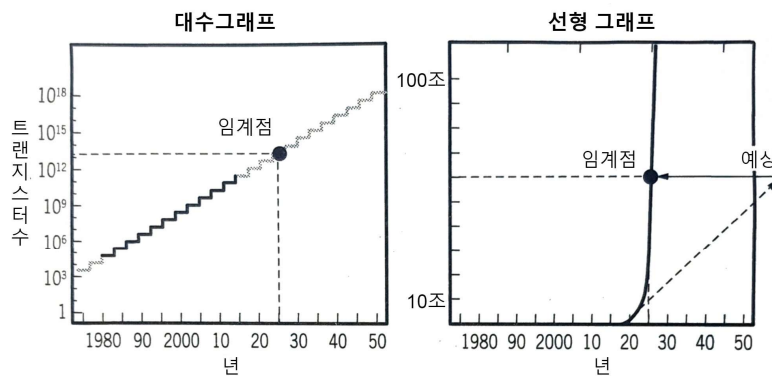
지수함수의 경이로움을 우리는 내다볼 수 없다.

연못의 잉어를 돌보는 노인이 있었다. 산소가 충분히 물속까지 녹아 들어가도록 때때로 연잎을 따 연못을 지키고 있었다. 연잎은 그다지 급격하게 늘어나지

않기 때문에 괜찮을 것이라 생각하고 1주 정도 연못을 떠났더니 연못은 완전히 연잎으로 뒤덮여 있었다.

이 이야기는 지수함수의 특징을 잘 나타낸다(코로나19 환자 수의 증가도 이와 마찬가지로이다).

그림3-2 기술이 지수함수적으로 성장해도 사람은 직선적으로 파악하기 때문에 변혁은 예상보다 빨리 찾아온다



(출처) 저자 작성

우리의 직관은 변화하는 현상을 직선 근사로 받아들인다. 태고의 옛날부터 정글 속에서 맹수(등속 운동)로부터 몸을 지키기 위해 갖게 된, DNA에 새겨진 감각이다. 현대 사회에서도 이전의 변화를 직선적으로 추정하고 미래를 예측하는 경우가 많다.

그러나 **칩이 만드는 세계는 지수함수적으로 성장한다. AI도 그중 하나이다.** AI가 갑자기 이 세상에 나타난 직후에 하늘 높이 치솟듯이 급성장하는 이유는 여기에 있다.

칩이 생성하는 데이터도 지수함수적으로 급증하고 있다. 인터넷 통신량은 연을 4배로 급격히 증가하고 있다(길더의 법칙, Gilder's law).

21세기 후반에는 전 인류의 뇌 속 뉴런의 총수에 필적하는 트랜지스터가 하나의 칩에 집적 가능할지도 모른다. 나아가 전 세계의 칩이 무선 접속된 거대한 두뇌가 지구상에 등장하는 것도 꿈만 같은 이야기는 아니다.

집적 회로의 발명으로부터 고작 100년 동안 세계는 극적으로 변하고 있다(그림 3-2).

3. 칩의 구조 개혁—누전을 줄여라

트랜지스터 구조 혁신

트랜지스터에는 세 개의 단자가 있다. 소스는 전하를 공급받는 입구, 드레인 은 전하를 내보내는 출구, 게이트는 전하의 흐름을 조정하는 수문이다. 게이트 전위를 변경함으로써 소스에서 드레인으로 전하를 흘려보내거나 멈출 수 있다. 이른바 스위치 역할을 하는 것이다.

트랜지스터 제작 방법은 반도체 기판 표면을 산화시켜 얇은 산화막을 만든 다음 그 위에 금속 게이트를 배치한다. 다음으로 반도체 기판에 첨가한 불순물과 반대 극성을 가진 불순물을 위에서부터 주입한다. 그러면 게이트 양쪽의 반도체 기판 표면에 불순물이 주입되어 소스와 드레인이 형성된다.

게이트 단면이 금속-산화물-반도체(Metal-Oxide-Semiconductor) 구조가 되므로 MOS 트랜지스터라고 불린다. 양전하인 홀이 많은 P형 반도체로 소스와 드레인을 형성한 MOS를 PMOS, 음전하인 전자가 많은 N형 반도체를 사용한 것을 NMOS라고 부른다.

NMOS의 작동 방식을 알아보자. 소스에는 전자가 담겨있다. 게이트가 소스와 동일한 전위일 때에는 소스와 드레인 사이에 P형 반도체 기판이 소스와의 사이에 전자로 된 장벽을 만들기 때문에 드레인과 소스 사이에 전압을 가해도 전자가 드레인으로 흘러가지 않는다.

그러나 게이트에 소스보다 충분히 높은 전위를 공급하면 게이트 바로 아래의 P형 반도체 기판 표면이 N형으로 반전되어 전자의 통로인 채널이 만들어지고, 소스에서 드레인으로 전자가 흘러나온다. 전류의 흐름은 전자의 이동과 반대이므로 드레인에서 소스로 전류가 흐른다.

PMOS 작동은 이와 정반대이다. 소스에 홀이 담겨있다. 게이트에 소스보다 충분히 낮은 전위를 공급하면 채널이 형성되어 소스에서 드레인으로 홀이 이동해 전류가 같은 방향으로 흐른다.

여기서 PMOS와 NMOS의 소스를 전원과 접지에 각각 연결하고, 게이트를 연결하여 입력전압을 받고 드레인을 연결해 출력전압이 나오게 하면 CMOS 인버터를 구성할 수 있다.

CMOS 인버터의 입력으로 낮은 전위(L)가 들어오면 NMOS는 Off되고 PMOS는 On되어 출력에는 전원에서 전류가 흘러나와 높은 전위(H)가 나온다. 마찬가지로 입력에 H를 넣으면 출력에서 접지로 전류가 흘러나와 출력으로 L이 나온다.

PMOS와 NMOS는 동시에 켜지지 않기 때문에 전원에서 접지로 전류가 계속 흐르지는 않는다. 출력을 H나 L로 변경하기 위해서만 전류를 사용하므로 저전력이다. PMOS와 NMOS가 이처럼 상호보완적(Complementary)인 동작을 하기 때문에 이를 CMOS라고 한다.

그런데 트랜지스터를 작게 만들면 드레인과 소스 사이에서 누전이 발생한다. 그 이유를 이해하기 위해서는 채널이 만들어지는 메커니즘에 대해 더 자세히 살펴봐야 한다.

NMOS의 작동 설명으로 돌아가 보자. 게이트에 소스보다 충분히 높은 전위를 가하면 어째서 P형 반도체 기판의 표면에 N형이 반전되어 채널이 만들어지는 것일까?

2장의 금속 전극이 마주 보는 커패시터(축전기, 콘덴서라고도 한다)를 떠올려 보자. 먼저 한 장의 금속판 A에 양전하를 주면 전하는 균일하게 분포한다.

다음으로 전하가 없는 다른 금속판 B를 평행하게 가까이 하면 정전 유도로 인해 금속판 A의 양전하에 끌려 금속판 B 안쪽에 음전하가 생기고 이와 동일한 양의 양전하가 금속판 B 바깥쪽에 분극된다. 그 결과 금속판 A에서도 양전하가 안쪽에 집중된다.

이어서 금속판 B를 접지에 연결하면 금속판 B 바깥쪽에 분극되어 있던 양전하가 접지로 흘러가게 된다. 그러나 금속판 B의 안쪽 음전하는 금속판 A의 양전하와 상호 작용하고 있어 이동할 수 없다. 그 결과 축전기 내부 전계에 전하가 저장된다.

여기서 금속판 B를 P형 반도체 기판으로 바꾼 것이 NMOS이다.

금속판 A 게이트에 소스보다 충분히 높은 전위를 주면 게이트에 양전하가 유도되며 게이트에 대항하는 P형 반도체 기판 표면에 음전하가 저장된다. 이 음전하 즉 전자가 충분히 많아지면 반도체 기판의 표면은 'N형으로 반전'되어 전자의 통로인 채널이 만들어진다.

이렇게 채널은 게이트에서의 전계 효과에 의해 제어된다.

그러나 게이트 외에도 채널에 영향을 주는 커패시터가 숨어있다. 바로 드레인이다.

드레인과 반도체 기판의 계면에서 드레인의 전자가 P형 반도체 기판으로 확산하고 P형 반도체 기판의 홀이 드레인으로 확산한다. 마치 용기 안에서 미리 설탕과 소금을 분리해 둔 판을 제거하면 설탕과 소금이 섞이는 것과 비슷하다. 차이점은 전자와 홀 사이에 정전기가 작용하기 때문에 조금 섞이면 그 이상은 확산하지 않는다는 점이다.

그 결과 드레인과 반도체 기판 계면에는 자유롭게 이동할 수 있는 전하가 결핍된 공핍층이 형성된다. 이것이 절연막이 되어 커패시터를 형성한다.

트랜지스터를 작게 만들면 소스와 드레인 사이의 거리가 짧아지고 드레인 공핍층이 소스에 가까워진다. 즉 소스 관점에서 보면 드레인도 작은 게이트인 것이다.

따라서 게이트가 닫혀 있어도 드레인에 양의 전위를 주면 소스의 전자를 가둬 둔 장벽이 조금 내려가 누전이 발생하게 된다.

미미한 누전이더라도 트랜지스터를 100억 개나 집적하면 큰 누전이 되어버린다.

누전이 생기는 원인은 게이트의 지배력 저하에 있다. 그렇다면 게이트의 지배력을 개선하기 위해서는 어떻게 해야 할까?

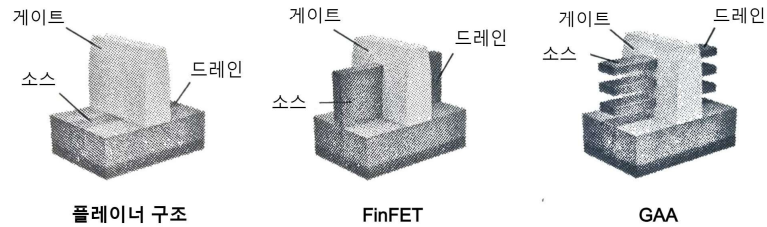
우선 재료를 바꾸는 방법이 제시되었다.

게이트 산화막을 유전율이 높은 재료로 바꾼 것이다. 이는 유효한 방법이었지만 미세화가 진행됨에 따라 효과도 약해졌다.

다음으로는 구조를 바꿀 수밖에 없다.

그래서 게이트를 두 개로 만들고 채널을 양쪽에서 끼우는 구조로 변경했다. 새로 추가되는 게이트를 채널 아래에 만들려면 제조 비용이 커지기 때문에 채

그림3-3 트랜지스터 구조 개혁



(출처) Lam Reserach

널을 반도체 기판 표면에 세워 그 양면에 게이트를 만들었다.

이것이 FinFET(핀펫) 이다(그림 3-3 중앙). 형태가 물고기의 지느러미(Fin)와 닮은 전계효과 트랜지스터(Field-Effect Transistor:FET)이기 때문에 FinFET 이라는 이름이 붙었다. 16nm 세대부터 채택하고 있다.

2nm 세대가 되면 게이트의 지배력을 더 높이는 구조가 필요해진다. 그래서 게이트가 채널을 둘러싸는 구조가 연구 및 개발 중에 있다. 이것이 GAA(Gate All Around)이다(그림 3-3 오른쪽). 게이트로 주변을 감싼 얇은 채널에 충분히 큰 전류가 On했을 때 전류가 흘러야 한다. 이를 위한 재료 특성이 연구 중에 있다.

이해를 돕기 위해 과감한 비유를 들어 보겠다. 호스의 물을 막고 싶은 경우를 상상해 보자. 기존에는 검지로 호스를 위에서 누르며 막았다(플레이너 구조). 그래도 물이 새어 나오는 문제가 발생해 다음으로는 엄지와 검지로 호스를 양쪽에서 눌렀다(FinFET). 마지막으로 다섯 손가락으로 호스를 움켜쥔 것이다(GAA).

배선 구조 개혁

회로 집적도가 늘면 칩이 소비하는 전력이 증가하고 발열량도 늘어난다. 온도가 올라가 칩이 고장 나지 않도록 전력의 허용 범위가 결정된다.

전력을 늘리지 않고 집적도를 높이기 위해서는 전원전압을 낮추는 것이 효과적이다. 전원전압을 절반으로 낮추면 CMOS 회로의 소비 전력을 1/4로 낮출

수 있으므로 집적도를 4배 높일 수 있다. 1980년대에는 전원전압이 5볼트였지만 현재는 0.5볼트이다. 이론적으로는 상온의 경우 0.036볼트까지 전원전압을 낮출 수 있다.

그러나 그전에 해결해야 할 큰 과제가 등장했다. 전원 배선이다.

전력은 전압과 전류의 곱이다. 집적도를 높여도 전력이 늘지 않도록 전압을 낮추면 그만큼 전류가 늘어난다. 예를 들어 전력이 50와트인 경우 전압이 5볼트이면 전류는 10암페어이지만 전압이 0.5볼트로 내려가면 전류는 100암페어로 늘어난다.

전자레인지나 전기 그릴이 10암페어이다. 그 10배에 달하는 큰 전류를 어떻게 1cm^2 라는 작은 칩에 공급할 것인가?

이를 위해서는 전원 배선을 크고 두껍게 만들어야 한다. 결과적으로 칩의 전원 배선은 미세화에 역행하듯이 두껍고 다층화된다. 반도체 기판 위에 형성되는 배선층은 1980년대에는 2~3층이었지만 최근에는 15층을 넘는다.

하층은 단거리 배선에, 중층은 원거리 배선에, 상층은 전원 배선에 사용된다. 하부층일수록 얇고 가늘며 상부로 갈수록 두껍고 굵다. 배선의 대부분이 전원 배선에 사용된다. 인체의 모세 혈관에서 대동맥처럼 다양한 배선이 칩 전체에 뻗어있다.

트랜지스터를 미세화할수록 전원 배선은 두껍고 굵어야 한다. 이 딜레마를 해결하기 위해 전원 배선을 반도체 기판 안에 내장시켜 전원을 칩 뒷면에서 공급하는 구조 개혁이 시작되려 하고 있다.

한계설 트렌드

미세화는 슬슬 한계에 다다랐다는 얘기가 들려온 지 오래다.

사실 1980년대부터 이러한 지적이 반복되어 왔다. 그러나 실제로는 이러한 지적을 극복하고 지금도 미세화는 계속되고 있다.

이것으로 한계설 트렌드라고 조롱하는 사람도 있다. 즉 '10년 후에는 한계'라는 주장이 빈번하게 40년 넘게 나오고 있다는 빈정거림이다. 하지만 다른 관점에서 보면 현 상황을 방치했을 때 10년 후에 어떤 한계가 닥쳐올지를 명확히

했기 때문에 우리는 그 한계를 넘어왔다고도 할 수 있다.

실제로 게이트 산화막을 유전율이 높은 재료로 바꾸는 것은 불가능하다고 여겨지는 도전이었다. 트랜지스터 1,000만 개 이상을 수율을 높여 제조할 수 있었던 것은 실리콘 기판 표면을 산화시켜 양질의 게이트 산화막을 만들 수 있었기 때문이다. 그러나 실현 불가능하다는 비판이 다수를 차지했었음에도 불구하고 고유도를 게이트 산화막은 실용화되었다.

또한 반세기 가까이 이어져 온 트랜지스터의 플레이너 구조가 FinFET이나 GAA로 과감하게 개혁되었다. GAA 구조의 PMOS와 NMOS를 수직으로 쌓는 CFET 연구도 시작되고 있다. 이제는 교과서를 10년마다 업데이트해야 할 것이다.

팻 겔싱어 인텔 CEO는 2001년에 다음과 같이 경고했다.

“현재의 CPU는 표면 1cm²당으로 환산하면 100와트 이상의 전력 밀도를 갖고 있으며 이는 원자로에 가까운 숫자입니다. 펜티엄(CPU 이름) 시절에는 전기 그릴 수준이었지만 이대로라면 10년 후에는 태양의 표면과 동일한 수준의 밀도가 될 것입니다.”

물론 위기는 피해 갔다.

필자가 도시바 연구소에서 근무할 때 선배에게 받은 가르침이 있다.

“불가능하다고 해서는 안 된다. 지금은 불가능해 보여도 미래에는 가능해질 지도 모른다. 몹시 어렵다고 해야 한다.”

이 말을 명심하고 있다.

4. AI 칩—뇌에서 배우다

수학에서 탄생한 컴퓨터

태고의 옛날에는 사람들이 손가락을 꼽으며 셈을 하고 걸음 수로 측량을 했다. 그러나 사람들은 큰 숫자를 인지할 수 없었다. 그런데 4대 문명에서는 계산기가 등장해 인간의 인지능력을 확장했다.

앞서 언급한 것처럼 고대 그리스 시대 이후에는 수학의 내부 세계가 연구 대상이 되어 수학은 도구에서 사고로 진화했다. 15세기 르네상스 시대에 기호 대수가 발명되어 현실 세계에서는 표현할 수 없는 n차원의 공간을 고찰할 수 있게 되었다. 이렇게 수학은 물리적 제약을 받지 않는 보편적인 입지를 구축한 것이었다.

17세기가 되면 미적분이 고안되어 무한한 세계를 탐구할 수 있게 되었다. 극한이나 연속성 개념을 엄밀하게 성찰한 결과 주관적인 감각을 초월한 추상적인 기호 체계가 만들어졌다. 그리고 20세기 들어 ‘수학을 하는 스스로의 사고에 대해 수학을 하는’ 시도마저 나타났다.

이렇듯, **수학은 신체에서 떨어져 나와 뇌에 머무르며 물리적 직감이나 주관적 감각 같은 모호한 것들을 완전히 벗어나고 마침내 뇌에서 넘쳐나왔다. 이것이 컴퓨터이다.**

처음의 전자식 컴퓨터는 진공관이 곧잘 고장 났다. 진공관은 전극을 가열하여 전자를 기체에 방출하고 그 전자의 흐름을 제어하는 장치이다. 가정용 백열전구와 마찬가지로 시간이 지남에 따라 전극이 얇아지고 결국 단선된다.

그래서 기체가 아닌 고체 속에서 전자를 제어하는 트랜지스터가 1948년에 발명되었다. 장치의 신뢰성은 단숨에 높아졌다.

또한 제2장의 내용을 되짚어 보자면 컴퓨터의 기능이 회로의 배선에 의해 결정되는 ‘포선 논리’에는 두 가지 과제가 있었다. 처리할 수 있는 프로그램의 최대 규모가 하드웨어의 규모로 제약을 받는 ‘규모 제약 문제’와 시스템 규모가 커지면 접속 수가 방대해지는 ‘대규모 시스템 접속 문제’이다.

그래서 폰 노이만은 여러 개의 연산 장치를 물리적으로 연결하는 대신 하나의 연산 장치가 사이클마다 다른 명령을 실행하는 ‘**프로그램 내장 방식(폰 노이만 아키텍처)**’를 발명해 규모 제약 문제를 해결했다.

한편, 잭 킬비가 1958년에 **집적회로(IC)**를 발명했다. 포토리소그래피를 이용해 한 장의 칩에 소자와 배선을 집적함으로써 ‘대규모 시스템 접속 문제’를 해결했다. 곧이어 실리콘이 IC에 가장 적합한 재료임을 발견하게 되었다.

이렇게 단순화·극소화된 연산 자원을 실리콘 칩에 집적화·병렬화함으로써 컴퓨터의 성능은 비약적으로 향상되었으며 성능이 높아진 컴퓨터는 더 대규모적

인 집적회로 설계를 가능케 했다.

이렇게 폰 노이만 아키텍처와 집적회로, 실리콘이 만나 컴퓨터와 칩은 함께 지수함수적 진화를 이룩한 것이다.

일을 하면 에너지를 소비한다. 전자회로의 일량 즉 성능은 전력 공급과 방열이라는 제약을 받는다. 에너지 또는 에너지 유속에 해당하는 전력의 효율을 높여야 칩의 성능이 향상된다.

칩의 전력 효율은 지난 20년 동안 세 자리 개선되어 뇌의 1/100 정도까지 향상되었다. 또한 칩의 집적도도 뇌 신경세포의 1/100 정도이다. 지금까지의 기세라면 10년 후에는 뇌를 따라잡을 것이다.

그러나 폰 노이만 아키텍처에서는 대량의 데이터와 명령이 프로세서와 메모리 사이를 오가므로 이 부분이 얇고 긴 목처럼 병목현상을 일으킨다 (폰 노이만 병목). 또한 실리콘 칩은 이번 세기 들어 디바이스의 치수가 100nm보다 작아지면 서 양자 효과가 나타나 누설 전류를 제어할 수 없다. 반세기 전에 탄생한 컴퓨터와 칩의 성장 한계가 드러난 것이다.

그러나 한계를 맞이하기 전에 **컴퓨터는 스스로 학습하는 능력을 갖추었다. 기계학습이다. 그리고 뇌의 신경회로망을 모방한 AI 칩이 탄생했다.**

뇌에서 배우는 AI칩

신경회로망(Neural Network:NN)을 설계하기 위한 요소 기술은 20세기에 개발되었으나 표현 가능한 공간이 너무 방대해져, 4층 이상의 심층 신경회로망을 학습시키기는 어려웠다.

그러나 21세기가 되며 오토 인코더의 심층화에 성공하고 학습에 필요한 컴퓨터 성능이 충분히 향상됨에 따라 심층 학습은 기존의 정보 처리에 비해 압도적으로 높은 처리 성능을 발휘하게 되어 급속히 상용화되었다.

회로망의 구성이나 아키텍처 연구도 진행되었다. 이미지 인식에서는 가까운 신호만 결합하는 합성곱 신경회로망(Convolution Neural Network:CNN)이 성공을 거두었다. 또한 음성이나 자연어 처리와 같은 시계열 데이터를 다루는 인

식 처리에서는 재귀 신경회로망(Recurrent Neural Network:RNN) 및 장·단기 기억(Long Short-Term Memory:LSTM)이 연구되었다. 최근에는 중요한 부분에 주목하는 어텐션 메커니즘이 도입되어 셀프 어텐션 기능을 이용함으로써 RNN의 재귀 구조를 사용하지 않는 트랜스포머 아키텍처가 주목받고 있다.

모두 우리 뇌에서 힌트를 얻어 연구가 진행되고 있다. 그중에서도 중요하다고 여겨지는 것이 **신경회로망 가지치기(프루닝, pruning)**이다.

우리 뇌의 시냅스는 태어날 때는 50조 개 정도만 존재하지만 생후 12개월까지 1,000조 개로 늘어난다. 그러나 그 이후에는 학습을 통해 시냅스가 감소한다. 신호가 통과하며 강화된 시냅스는 남지만, 신호가 오지 않는 불필요한 시냅스는 제거되어 사라지게 된다. 10살까지 시냅스는 반으로 줄어들고 그 후에는 변화가 적어진다.

즉, **유아기 초기까지는 완전 결합에 가까운 신경회로망이 형성되지만 학습함에 따라 불필요한 배선이 제거되고 필요한 배선만 남게 되는 것이다. 이렇게 낭비 없이 효율적인 뇌 회로가 완성된다.**

어린이의 뇌는 학습하기 위해 크고, 성인의 뇌는 효율적인 추론을 위해 가지치기 되는 것이다. 작은 아이를 낳아서 크게 키우고 사회에서 학습시키는 전략은 뇌가 발달한 포유류의 생존 전략일 것이다.

뇌와 실리콘 뇌

뇌와 실리콘 뇌에 관해 이야기해 보자. 수학에서 탄생한 폰 노이만 아키텍처 컴퓨터는 미리 프로그래밍이 된 상태 전이를 기반으로 견고한 정보처리를 수행한다. 이는 마치 유전으로 기능이 갖추어진 뇌의 시상체·편도체·소뇌와 닮아 있다.

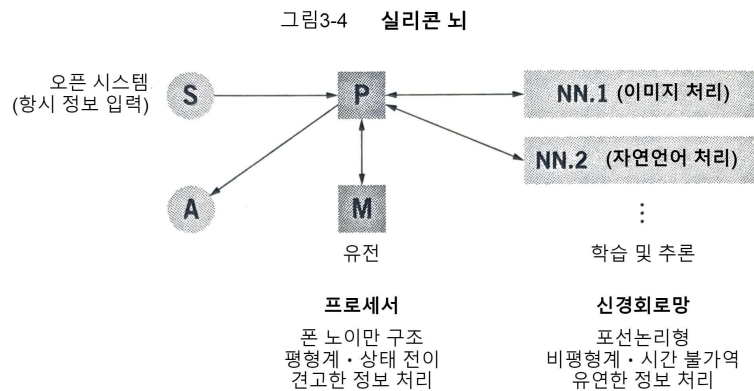
한편 뇌를 모방한 포선 논리 방식 신경회로망은 오픈 시스템에서 학습을 계속하며 프루닝을 수행하고, 시간 불가역적인 유연한 정보 처리를 에너지 효율적으로 수행한다. 이는 마치 사회에서 학습하는 대뇌 피질과 비슷하다.

이처럼 실리콘 뇌는 인간의 뇌를 참고해서 설계할 수 있다. 그렇다면 실리콘

뇌는 인간의 뇌와 비슷한 구조를 가질까? (그림 3-4)

“엄청난 다이내믹 레인지야!”

1981년 이렇게 외친 사람은 연구실 선배인 아이하라 가즈유키(合原一幸, 현 도쿄대학 특별교수)였다. 신경 축삭의 활동전위 생성과 전파를 기술한 비선형 미분방정식인 호지킨-헉슬리 방정식(Hodgkin-Huxley equations)을 계산기로 분석했더니 신경 축삭의 저항값이 크게 변했던 것이다.



프로세서가 뇌 시상체·편도체·소뇌 역할을 맡고 신경회로망이 대뇌 피질의 역할을 한다
(S:센서, A:액추에이터, M:메모리, NN: 신경회로망)

(출처) 저자 작성

유사한 특성을 갖는 인공물을 만들어 내는 것은 쉽지 않다. 뇌와 실리콘 뇌는 새와 비행기처럼 다른 원리와 구조를 갖게 될지도 모른다.

신경회로망은 배선 접속이 기능을 결정하는 ‘포선 논리’이다. 제조 후 회로를 프로그래밍할 수 있는 FPGA에 나는 많은 기대를 품고 있다.

참고문헌

[문헌]

오타야스히코(2021) 『2030 반도체지정학』 성안당
전병서(2023) 『한국 반도체 슈퍼 을 전략』 경향BP
高乘正行(2022) 『ビジネス教養としての半導体』 幻冬舎

[웹사이트]

국립국어원 표준국어대사전 <https://stdict.korean.go.kr/main/main.do>
(2023.11.16.)
SK 하이닉스 뉴스룸 <https://news.skhynix.co.kr/post/> (2023.11.23.)
삼성전자 반도체 뉴스룸 <https://news.samsungsemiconductor.com/kr/>
(2023.11.23.)

