# 차세대 지능형 소자 구현을 위한 모노리식 3D 집적화 기술 이슈

## Issues on Monolithic 3D Integration Techniques for Realizing Next Generation Intelligent Devices

**문제현** (J. Moon, jmoon@etri.re.kr) 플렉시블전자소자연구실 책임연구원
**남수지** (S. Nam, sjnam15@etri.re.kr) 플렉시블전자소자연구실 선임연구원
**주철웅** (C.W. Joo, cwjoo@etri.re.kr) 플렉시블전자소자연구실 선임연구원
**성치훈** (C. Sung, sungch90@etri.re.kr) 플렉시블전자소자연구실 UST학생연구원
**김희옥** (H.O. Kim, istoo85@etri.re.kr) 실감디스플레이연구실 기술실무원
**조성행** (S.H. Cho, helloshcho@etri.re.kr) 플렉시블전자소자연구실 책임연구원
**박찬우** (C.W. Park, chanwoo@etri.re.kr) 플렉시블전자소자연구실 책임연구원/실장

## ABSTRACT

Since the technical realization of self-aligned planar complementary metal-oxide-semiconductor field-effect transistors in 1960s, semiconductor manufacturing has aggressively pursued scaling that fruitfully resulted in tremendous advancement in device performances and realization of features sizes smaller than 10 nm. Due to many intrinsic material and technical obstacles, continuing the scaling progress of semiconductor devices has become increasingly arduous. As an effort to circumvent the areal limit, stacking devices in a three-dimensional fashion has been suggested. This approach is commonly called monolithic three-dimensional (M3D) integration. In this work, we examined technical issues that need to be addressed and overcome to fully realize energy efficiency, short latency and cost competency. Full-fledged M3D technologies are expected to contribute to various new fields of artificial intelligence, autonomous gadgets and unknowns, which are to be discovered.

**KEYWORDS** 반도체, 모노리식 3D 집적화, 스케일링, 상보형금속산화막반도체, 후공정, 공정호환

## Ⅰ. 서론

1960년대 후반 자발정렬 Si 기반 MOSFET의 기술적 구현은 반도체 소자와 집적회로의 기술적 가능성을 열어주었다. 관련하여 1965년에 Fairchild Semiconductor와 인텔의 공동창업자 무어(G.

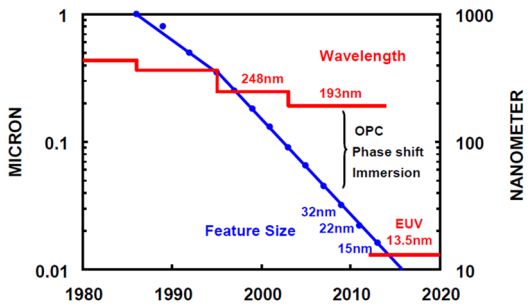
* DOI: https://doi.org/10.22648/ETRI.2021.J.360302
* 이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단-차세대지능형반도체기술개발(소자)사업 지원을 받아 수행된 연구임 [연구과제명: 저온 공정 산화물 반도체 기반 초저전력 단일3차원 집적 로직 소자 및 아키텍쳐 개발(NRF-2020M3F3A2A01085791)].

본 저작물은 공공누리 제4유형
출처표시+상업적이용금지+변경금지 조건에 따라 이용할 수 있습니다.
©2021 한국전자통신연구원

Moore)는 반도체 소자의 평면 밀도가 격년주기로 2배 증가할 것이라고 예측하였다[1]. 무어의 예측은 다분히 현상적인 관찰에 근거하고 있다. 무어의 법칙을 좀 더 정량적으로 기술하기 위하여 Meindle은 $N=F^{-2}D^2P$라는 식을 제시하였다. 여기서 $N$은 칩당 트랜지스터의 개수이다. $F$와 $D$는 각각 소자에서 가장 작은 요소의 치수와 칩의 면적이다. $P$는 트랜지스터의 면적당 충진율이다. 반도체 공정 관점에서는 $F$에 기술적 역량이 주로 집중되어 있다고 하겠다. 이 식에 의하면 $F$가 1/10으로 작아지는 경우, $N$은 100배로 커짐을 알 수 있다. 증착기술과 패터닝 공정의 정교화로 소자밀도는 무어의 예측과 추세적으로 상당히 유사하게 높아졌다. 지난 50년간 반도체 사업은 꾸준한 소형화 스케일링을 진행하며 소자성능 고도화, 전력소모 저감, 집적회로 안정화 등을 추구하면서 발전해왔다. 2000년대 초반까지 이른바 무어의 법칙이 비교적 잘 적중했던 기간을 통상적으로 기하학적 스케일링(Geometrical Scaling) 시대라 부른다. 이 시기에는 주로 평면상에서 트랜지스터의 구성요소의 치수를 작게 만드는 공정들이 비약적으로 발전하였다. 기하학적 스케일링은 흔히 MOSFET의 채널이 길이($L$)의 관점에 고찰된다. 통상적으로 소자의 길이가 1/n으로 작아지면 전압과 전류도 각각 1/n배 감소하므로 소비전력은 $1/n^2$으로 줄일 수 있다. 또한 크기가 작아졌으므로 소자의 평면 밀도 또한 증가시킬 수 있다. 드레인 전류($I_D$)는 게이트 절연막의 유전상수($k$)와 채널폭($W$)에는 비례하고, 게이트 절역막 두께($t$)와 채널길이($L$)에는 반비례한다($I_D \sim k \times W/L \times 1/t$). 따라서 채널의 $L$을 줄이면 $I_D$가 증가한다. 이를 조절하기 위하여 $I_D$에 관련된 치수와 소재를 조절하는 연구가 활발하게 진행되었다. 스케일을 통한 소자성능 향상, 소비전력의 이점은 거의 포화단계에 도달하였다. 이에 대한 대응으로 기하학적 스케일 접근법보다는 기능적 요소를 갖는 다이(Die)를 적층하는 접근법이 심각하게 재고되기 시작하였다. 다이적층은 상하부를 연결하는 TSV(Through Si Via) 기술에 크게 힘입었다. 본 고에서는 무어의 법칙 유지가 기술적 한계에 도달한 시점에 중요하게 부각되고 있는 모노리식 3차원(M3D: Monolithic 3 Dimensional) 집적화에 대하여 살펴보고자 한다.

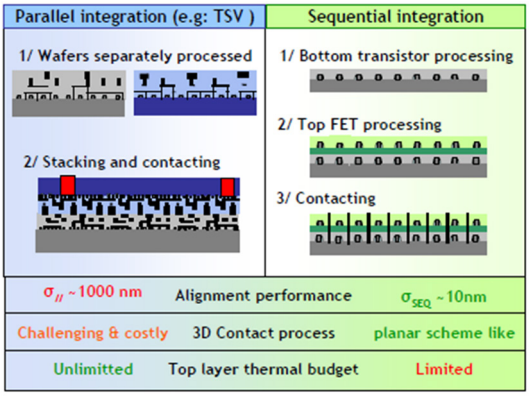
## Ⅱ. 반도체 발전 및 M3D의 등장

그림 1은 연도에 따른 패턴크기와 포토리소그래피에 사용된 광원 파장(λ)을 보여준다. 2000년대 초기까지는 λ=248nm KrF 광원을 활용하여 약 100nm 크기 수준의 패터닝이 구현되었으며, 이후에 λ=193nm ArF 광원을 활용하여 최소~20nm 수준까지 안정되게 패터닝되었다. 2004년에는 렌즈와 기판표면 사이에 굴절률이 큰 액체를 위치시켜 패턴 분해능을 증가시킨 액침 리소그래피(Immersion Lithography)도 소개되었다. 2015년 이후에 네덜란드의 ASML(Advanced Semiconductor Material Lithography)에서 성공적으로 극자외선(Extreme Ultra Violet (EUV), λ=13.5nm) 장비를 개발하여




출처 Reprinted with permission from IEEE[2]

그림 1 CMOS 패턴과 리소그래피광원의 변천



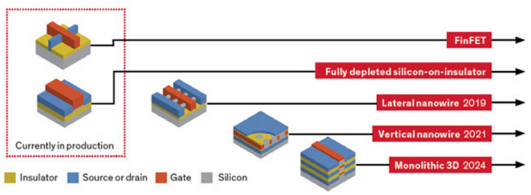

출처 Reprinted with permission from IEEE[3]

**그림 2 M3D 구현 방법. 병행집적화(왼편)와 연속집적화(오른편) 방식**

10nm 이하의 패터닝이 가능하게 되었다.

2018년 이후로 반도체 집적회로가 직면한 평면상의 스케일에 대처하기 위한 노력은 크게 두 가지로 대별할 수 있다(그림 2). 첫째는 새로운 패터닝 기법을 개발하여 소자의 크기를 아주 작게 만드는 More Moore 접근 방식이다. 둘째는 평면 차원을 탈피하여 수직방향으로 소자구성 층위를 3차원으로 적층하는 More than Moore 방식이다. 첫째의 대표적인 경우로는 EUV 공정을 들 수 있다. 이 방식에서는 기본적으로 패턴 소형화가 핵심이며, 이에 관련된 광원, 포토레지스트, 마스크, 펠리클 개발이 주요한 토픽이다. 둘째의 경우는 흔히 3차원(3D) 집적화로 일컫는다. 이 접근 방식에서는 상부층과 하부층이 집적화된다. 상부층과 하부층은 기능이 동일하거나 상이하다. M3D 구현에서는 상·하부층을 전기적으로 연결해 주는 저저항 Via 기술과 상하부층 간의 공정 호환성이 매우 중요하다. 3D 집적화를 구현하기 위한 기술적인 접근법은 대별하여 병행 집적화 및 연속 집적화 방식 두 가지가 있다. 병행 집적화 방식에서는 이종의 기판에 소자/회로 제작이 진행되며 개별기판 공정 완료 후 합착이 이루어진다. 이 방식에서는 공정이 별도로 진행되므로 상·하부공정의 불합치성으로 인한 제조 제약이 완화되지만 전기적 연결인 Via(관통전극)의 정렬 및 형성이 소형패턴 및 고집적도 회로에서는 상당히 어렵다. 병행집적화의 예로는 TSV가 적용된 다층의 DRAM 모듈을 들 수 있다. TSV의 치수는 지름 10mm, 높이 200mm 수준이다. TSV가 적용되는 경우 치수의 소형화에 한계가 있다. 소형화의 한계를 극복하기 위하여 SOI(Si On Insulator) 기판을 사용하는 경우도 있다. 연속집적화 공정에서는 하부 MOSFET 기반의 CMOS 공정을 진행한 후 상부에 절연막을 구비하고 기능적으로 상이한 소자/회로층을 제작한다. 이론적으로 연속 집적화 공정에서는 상·하부가 동일한 소자밀도를 가질 수 있다. 450℃ 이상에서는 도판트 확산, 소자특성변화, 메탈배선 치수변경 등으로 CMOS 부의 성능 저하가 나타나므로 상부층의 공정 온도는 상기 온도 이하로 유지함이 매우 중요하다.

최근의 새로운 반도체칩 시장에서 칩 소비자와 칩 제조사의 위치 혹은 입장이 뒤바뀌고 있다. 과거에는 칩 제조사에서 성능 및 집적도를 발표했다. 최근에는 최종 소비자 제품 제작사가 요구되는 사양을 제시하는 상황으로 반도체 사업이 전개되고 있다. 즉, 소비품 제조사가 자사 제품들의 기술적 성능을 만족하는 칩의 성능을 정의하게 되었다. 그림 3에는 이를 만족하기 위한 몇 가지 핵심 기술들




출처 Reprinted with permission from IEEE[4]

**그림 3 반도체 제작 핵심 요소 기술**

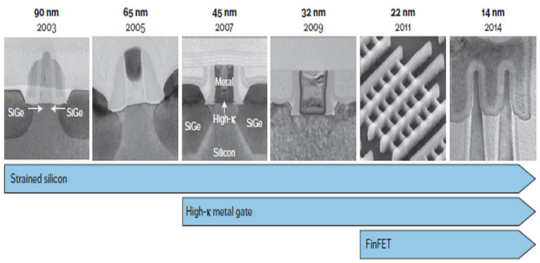
이 제시되어 있다. FinFET 같은 기술은 꾸준히 사용될 것으로 예상되며 나노선 집적화와 M3D가 또한 연구개발의 단계에서 상용화수준으로 발전하게 될 것으로 예상된다.

## III. 반도체 로드맵 개괄

반도체 및 전자 소자의 로드맵을 발표하는 IEEE 산하 IRDS(International Roadmap for Devices and Systems)에서는 2020 보고에서 미래의 빅데이터, 클라우드, IoT 모빌리티 등에 효과적으로 대응하기 위하여 소자 측면에서 요구되는 PPAC 기준을 제안하였다. PPAC는 각각 Performance, Power, Area, Cost의 첫 글자를 모은 약어이다. 최근에는 딥러닝과 인공감각과 같이 고속으로 대용량의 자료를 처리해야 하는 응용분야가 많아짐에 따라 PPAC 요구사항은 점점 중요하게 대두되고 있다. 이의 구체적인 내용은 다음과 같다.

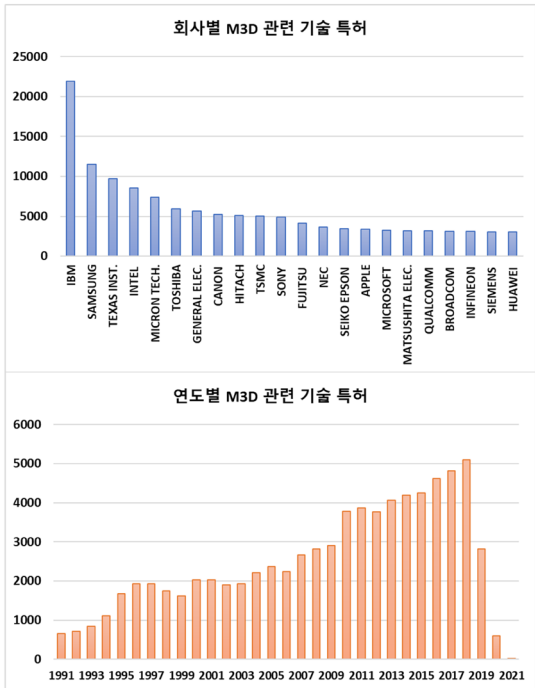
- Performance: 현재의 전압수준에서 처리 속도 15% 증대
- Power: 동일성능 기준 30% 저감
- Area: 동일성능 기준 칩 면적 30% 감소
- Cost: 15% 수준의 생산단가 감소

이를 구현하기 위한 중요한 기술로 극자외선 패터닝 기술, 게이트 올 어라운드(GAA: Gate All Around) 구조와 3D 집적화가 제시/시연되었다. 앞서 언급한 평면상의 기하학적 스케일링 기술적 한계들이 여러 가지 요소에서 등장하기 시작하였다. 대표적으로 채널의 길이가 10nm 이하로 매우 짧아지면서 숏채널 효과가 주된 소자 작동 장애로 등장하기 시작하였다. 이와 관련한 스위칭 소자의 On/Off 및 문턱전압 이하 스윙(Subthreshold Swing) 등의 특성을 조절하기가 어려워졌다. 기술적 한계를 극복하기 위한 대응으로 소재 및 구조 측면에서 극복기술이 제시되었다. 대표적으로 $HfO_2$나 $ZrO_2$ 같은 고유전율 게이트 절연막, 비평면구조의 FinFET, 채널에 응력을 인가하는 Strained Channel 등이 제시되었다. FinFET에서 등장한 전류조절의 어려움을 극복하기 위하여 채널의 모든 면을 게이트 층으로 두르는 GAA 구조도 제안되었다. 편의




출처 Reprinted with permission from IEEE[5]

그림 4 반도체 제작 핵심 요소 기술




출처 Wips 분석 결과

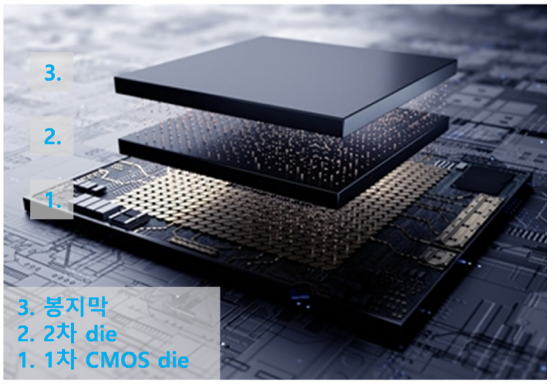
그림 5 회사 및 연도별 M3D 특허현황

상 이 시기를 등가 또는 유효 스케일링(Equivalent 혹은 Effective scaling) 시대라 부른다. 2025년 이후론 EUV 노광 공정과 함께 GAA가 주류 기술로 자리 잡을 것으로 IRDS는 예측하고 있다. 그림 4에 유효스케일링에 관련된 소자구조와 주요 치수를 도시하였다.

그림 5는 "Monolithic 3D integration"을 핵심 검색어로 삼아 실시한 특허 검색 결과이다. 미국 IBM이 약 17%의 비율로 1위를 차지하고 있으며, 이어서 삼성이 9%를 유지하고 있다. 인텔, TSMC, 애플 및 마이크로소프트는 각각 6.7%, 4%, 2.7%, 2.5%의 비율을 차지하고 있다. 국가별로 보면 미국과 일본이 큰 비중을 차지하고 있으며, 중국의 화웨이는 20위에 자리 잡고 있다. 연도별 특허 수는 꾸준히 증가하고 있다. 2010년에 특히 특허 수가 급격히 증가하고 있다. 이 시기는 FinFET이 본격적으로 상업용 반도체 소자에 등장하기 시작한 시기이다. 평면상 스케일링이 한계에 도달했을 때 M3D 특허가 급증하는 이유는 어렵지 않게 이해된다. 기술적으로는 트랜지스터 분야가 가장 많으며 요소 기술로는 열 방출, 배선, 패터닝 분야가 많다.

## Ⅳ. M3D의 기술적 이슈

이 장에서는 반도체 분야 한계 극복 기술로 제안되고 있는 M3D의 기술적 이슈를 살펴보고자 한다. 3D 적층기술은 이미 수직 NAND(V-NAND) 메모리 기술에서 도입되어 상용화되었다. 동일한 소자가 반복해서 쌓아 올린 형태의 메모리에서 활용되고 있는 삼차원 적층 기술과는 달리 M3D의 경우에는 상이한 기능을 수행하는 구성 요소들이 적층된다는 점에서 V-NAND와 차이가 있다. V-NAND에서는 동일한 기능을 수행하는 컨트럴




출처 Samsung Newsroom,"Samsung's new 3D integration technology X-Cube(eXtended-Cube)," 2020, Reprinted with permission from Samsung.

그림 6 삼성전자의 3D 적층기술의 X-Cube 칩 모식도[6]

게이트가 기둥 형태로 켜켜이 적층되어 있다. 그림 6은 M3D의 적층기술의 예시를 담은 모식도이다. 그림에서 볼 수 있듯이 하부에는 모노리식 공정으로 소자 어레이가 구비되어 있으며, 상부의 절연막 상에는 하부와 기능이 상이한 어레이가 구비되어 있다. 최종적으로 봉지막이 구성된다.

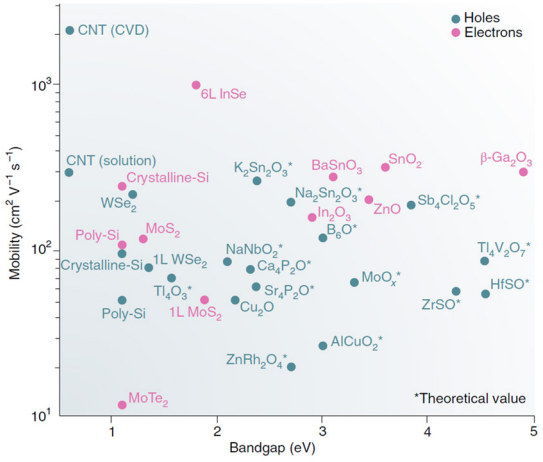
### 1. 성능측면

무어의 법칙은 단위 면적당 트랜지스터 밀도의 증가율에 대하여 기술한다. 그렇지만 밀도 자체는 회로와 소자 성능에 대해서 직접적으로 말해 주지는 않는다. 소모 전력과 구동 가능 주파수의 경우 흔히 Dennard 스케일링이 적용된다. 소모전력($P$)과 작동주파수($f$), 정전용량($C$) 및 구동전압($V$) 사이에는 $P \sim fCV^2$의 관계가 있다. 따라서 소자의 크기가 작아지면 게이트 절연막의 $C$도 작아지므로 동일한 $P$에서는 높은 주파수에서 작동이 가능하다. 또 다른 가능성으로는 동일한 주파수에서는 $V$를 낮출 수 있다. 따라서 소자가 작아지면 주파수, 구동전압 및 소모전력에 대한 이득이 있을 수 있

다. 그렇지만 작아진 크기로 인하여 발생하는 누설 전류와 숏채널 형성은 부정적인 효과이다.

## 2. M3D 상부 소재 선택

전술한 바와 같이 병행 집적화는 상·하부를 전기적으로 연결하는 Via 정렬의 어려움 및 이로 인한 한계 때문에 연속집적화 방식이 선호된다. 하부가 CMOS 기판인 경우 성능 열화 방지 및 BEOL(Back End of Line) 공정 호환성이 상부 소재 선택에 중요한 기준이 된다. 특히 450℃ 이하의 공정 온도 가이드라인 준수가 중요하다. 그림 7은 상부 트랜지스터 반도체 채널에 활용될 수 있는 물질들의 평균적인 이동도와 밴드갭이 도시되어 있다. 이동도 측면에서는 화학기상증착법으로 제작된 탄소나노튜브(CNT)가 ~1,000cm$^2$/Vs로 가장 높으며 밴드갭 측면에서는 베타 갈륨 산화물이 4.6eV 수준으로 가장 넓다. 소재 선정에 덧붙여 공정적인 측면에서는 패턴성이 매우 중요하다고 하겠다. 현재까지는 진공 증착 및 포토리소그래피보다 더 우수한 균일성, 치수, 형상 및 상·하부 패턴 정렬성을 제공해 주는 기술적 방법은 제시되지 않고 있다. 이 관점에서 보자면 용액 공정이나 전사법은 M3D에 적합하지 않다고 판단된다. 상기를 염두에 두면 다결정 Si 박막과 금속산화물 계열 물질로 소재의 선택폭이 좁아짐을 알 수 있다. Si은 n형과 p형을 모두 구현하기에 용이하지만 금속산화물의 경우 p형은 아직 안정성과 성능이 미진한 상태이다. 안정한 금속 산화물 p형 반도체를 구현하기 위해서는 유효질량이 낮고 양이온의 s 전자 궤도가 원자가 띠(Valence Band) 형성에 기여하는 물질을 발굴 내지는 설계하여야 한다. 금속 산화물의 경우 스퍼터링 혹은 원자층 증착방법 등으로 박막 형성이 용이하며, 포토리소그래피로 치수와 형상이 정



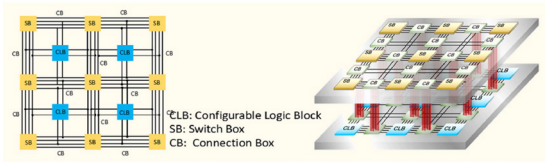

출처 Reprinted with permission from Nature[7]

**그림 7 BEOL 공정과 호환성이 있는 M3D 집적화 채널 물질 후보군**

확한 패턴 형성이 가능하다. 또한, Si에 비해 저온 공정이 가능하며, 매우 낮은 off 전류를 가지고 있어서 M3D의 상부 소재로 각광받고 있다.

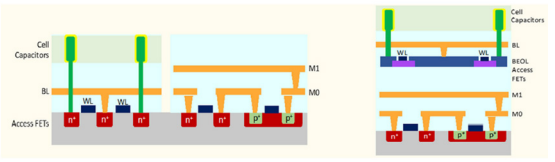
## 3. 실리콘 계열 M3D의 실시 예

M3D 집적화 공정은 평면상의 소자 제작과 비교하여 최종 결과물이 성능 향상 또는 에너지 효율 증대가 확실할 때 유의미하다. 실제 상용화 측면에서는 진공 증착 및 포토리소그래피 사용이 중요한 공정 기준이 될 수 있다. 가장 대표적인 Si 계열의 M3D는 V-NAND를 들 수 있다. 이 구조는 고집적 메모리 용량을 성공적으로 구현하였다. 그렇지만 이는 기능적인 측면에서는 서로 다른 기능을 갖는 소자의 집적화라기보다는 다층 소자막의 수직형 스트링 제작 기술 능력 측면이 (예를 들면, 높은 aspect ratio를 갖는 hole etching) 더 크다고 하겠다. 이종 기능의 집적화 실시에 대한 예는 문헌상으로 보고된 예가 아직 드물다. 이로부터 M3D 집적화 분야 신규 소재 및 신규 3D 집적화 아키텍처




출처 Reprinted with permission from IEEE[8]

그림 8 M3D 집적화를 이용한 FPGA 예시도




출처 Reprinted with permission from IEEE[8]

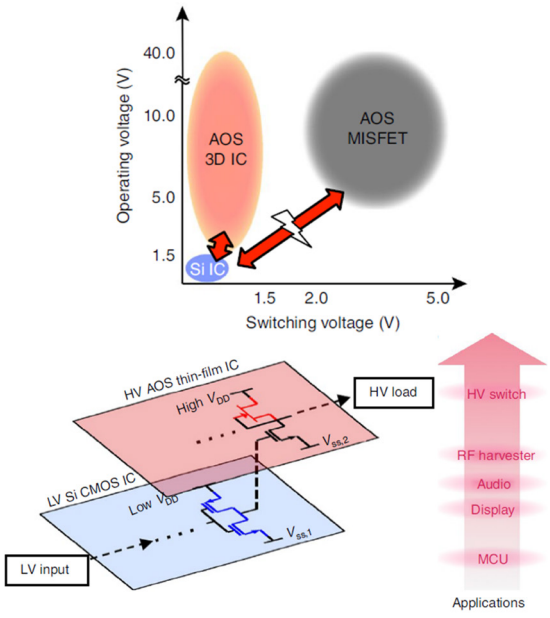
그림 9 M3D 집적화를 이용한 DRAM 예시도

의 원천 연구 개발이 필요함을 알 수 있다. 성능 향상 및 에너지 효율 등의 요구사항에 앞선 M3D 집적화 공정 구현의 중요한 이슈는 저온 공정으로 구현 가능한 상부층 소자의 하부 Si 소자 대비 성능, Complementary 소자 구현의 용이성, 그리고 패턴 밀도 등을 들 수 있다. 이 절에서는 문헌을 통하여 보고된 예를 살펴보고자 한다.

FPGA는 사용자가 필요에 따라 논리회로 구조를 변경할 수 있는 논리회로가 포함된 소자이다. FPGA에 M3D를 접목시키면 성능을 개선할 수 있다(그림 8). 통상의 FPGA에서는 CLB(Configurable Logic Block)가 평면상에 배열되며 CLB 사이의 데이터는 SB(Switch Block)으로 전송 혹은 라우팅된다. 평면구조상에서는 SB 사이의 라우팅 시간을 확보하기 위한 지체(Delay) 때문에 속도 저하(Latency Loss)가 발생한다. 또한 소자 간의 연결구역인 CB(Connection Box)의 물리적 이격으로 소비전력이 증가하게 된다. 평면상의 단점으로 인하여 집적도에 한계가 있다. 이를 해결하기 위한 방안으로 CLB를 하부에 위치시키고 상부에 SB와 CB를 위치시키는 구조가 제안되었다. 문헌에 의하면 이 구조는 로직 밀도 10배 증가, 속도 저하 4배 감소 및 소비 전력을 5배 정도 줄일 수 있다.

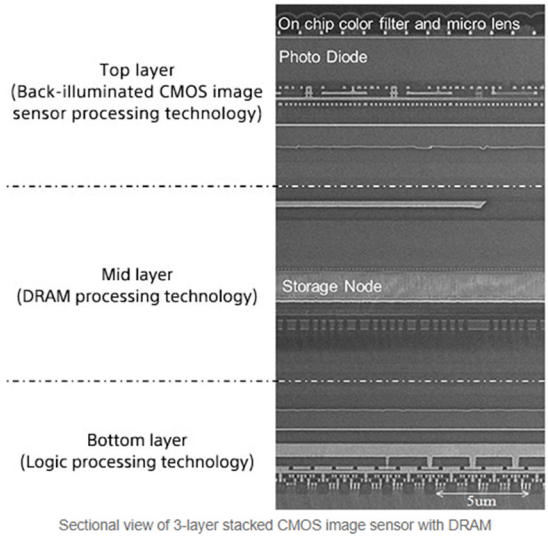
축적 용량 소자의 종횡비가 지수적으로 증가하는 이유 때문에 DRAM의 스케일링은 상당한 기술 장벽에 직면하고 있다. 이를 극복하기 위하여 Fin-FET에서 나아가 GAA FET가 사용되고 있다. M3D 접근법은 난이도가 높은 소자제작 기술을 우회할 수 있는 기술적 방안을 제시한다. M3D DRAM에서는 상부에 축적 용량 소자와 어레이 트랜지스터가 위치하며 하부에 로직 트랜지스터가 위치한다. 이러한 구조에서는 통상적인 BEOL과 공정 호환성이 양호한 채널 소재를 찾는 일이 기술적 과제로 남아 있다.

M3D 접근법은 고전압 파워 IC에도 유의미하게 적용될 수 있다. 수십 nm 이하 크기의 Si CMOS 기반 소자는 약 1V 정도의 구동 전압으로 동작되나, 전력관리시스템, 교류-직류 정류 회로, 정전기 보호회로 등은 작게는 2V에서 크게는 100V에서 동작된다. 따라서 System-on-Chip(SoC) 설계 관점에서 저전압 실리콘 기반 Microcontroller Unit(MCU)과 고전압 Power Management 회로를 Monolithic 3D 집적 방법을 이용하여 구현할 경우, 저전압에서 구동되는 소자와 고전압에서 구동되는 소자 간의 연동 소자가 필요하게 된다. 미시간 대학교 전기 컴퓨터 공학과 R. L. Peterson 교수 그룹은 2019년에 ZnSnO 기반의 산화물 반도체를 이용하여 통상적인 MISFET(Metal-Insulator-Semiconductor Field-Effect Transistor)과 Schottky-gated 박막트랜지스터 형태의 MESFET(Metal-Semiconductor Field-Effect Transistor), 그리고 Schottky Diode를 100nm-node FinFET이 구현된 웨이퍼 위에 직접 제작함으로써 MCU 구성 요소인 Si CMOS 구동용 저전압 조건으로 고전압 로직 인버터와 고전압 정




출처 Reprinted with permission from Nature[9]

**그림 10 산화물 반도체와 Si CMOS 집적화를 이용한 고전압 파워 연동의 예시도**




출처 Reprinted with permission from Sony[10]

**그림 11 소니의 DRAM 구비 CMOS 이미지센서의 단면구조 사진**

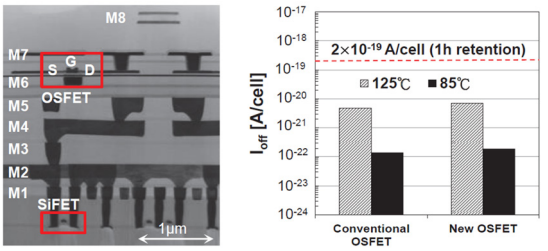
류기를 구동할 수 있음을 Nature Electronics에 보고하였다. 이때 최대 공정 온도는 520℃ −2시간 조건이 적용되었다. 향후 BEOL(Back End of Line)에 적용할 수 있기 위해서는 400℃ 미만의 열처리공정이 추가로 개발해야 함을 예고하고 있다.

## 4. 이종 집적화 3D의 실시 예

이종 M3D 집적화는 주로 이미지 센서 분야에서 폭넓게 활용되었다. 이 응용에서는 하부에 Si 기반 MOSFET 어레이를 사용하고 상부에는 반도체 기반 광학 소자 이미지 센서를 집적화시킨다. 대표적인 예로 스마트폰에 장착된 디지털 카메라를 들 수 있다. 최근에는 고해상도 디스플레이가 요구되는 AR/VR(증강현실/가상현실) 분야에 동일한 접근 방식이 활용되고 있다. 대표적인 예로 마이크로 LED와 고해상도 OLED 기구를 들 수 있다. 현재 상용화된 예로는 소니의 3층 구조 CMOS 이미지 센서가 있다(그림 11). 이 접근법은 콤팩트한 구조와 수직 배선 연결의 이점을 살려 RC 지연을 크게 줄일 수 있는 장점이 있다.

IoT 모바일 기기는 휴대용이라는 특성 때문에 전력소모절감이 매우 중요한 이슈이다. 소자어레이에서 전력낭비는 대부분 대기상태에서 발생한다. 대부분의 산화물박막은 단극성이므로 매우 낮은 누설전류는 off 상태에서 확보가 가능하다는 점에 착안하여 전력소모를 줄이기 위한 Noff(Normally Off) CPU 구현용 요소기술로 65nm Si MOSFET과 60nm 산화 반도체가 집적된 사례가 보고되었다(그림 12).

이 보고에 의하면 c축으로 결정화된 고이동도 IGZO가 핵심기술이며, 이의 적용으로 ~$10^{-21}$A 수준의 매우 낮은 Off 상태의 전류달성이 가능했다. 상기의 예를 구현하기 위해서는 Si MOSFET급의

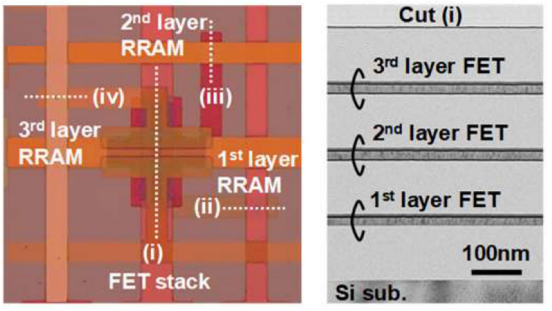



출처 Reprinted with permission from IEEE[11]

그림 12 65nm SiFET과 60nm OSFET 적층구조와 누설전류 수준

안정된 산화물 패터닝 기술, 상·하부층 Via 기술 및 고이동도 산화물 확보가 필수적이다. 상부 결정화를 위한 공정 진행 시 Si MOSFET의 성능을 저하시키지 않은 공정 개발이 매우 중요하다고 하겠다. 또한 매우 낮은 전류 수준을 측정하기 위해선 측정기술/장비 및 평가 표준화도 필요하다고 사료된다.

이종 결합의 또 다른 예로는 인공신경망이 있다. RAM을 기반으로 한 이원 신경망(Binary Neural Network)은 통상적으로 안정성과 노이즈 측면에서 유리하지만 다수의 뉴런 필요 및 평면상의 구조로 인한 신호 지연과 소비전력의 단점이 있었다. 이를 해결하기 위한 기술적 수단으로 모노리식 공정으로 $HfO_2$ 기반 저항 RAM(RRAM: Resistivie RAM)과




출처 Reprinted with permission from IEEE[12]

그림 13 $HfO_2$ RRAM과 IGZO FET 적층구조를 갖는 3D BNN

IGZO FET 삼차원 구조로 제작하는 방법이 제안되었다. 이 구조에서는 RRAM/IGZO FET의 3층 구조가 사용되었으며, BNN의 작동이 검증되었다(그림 13).

이종 집적화는 오감 센서 및 디스플레이 분야는 자율운행/안전기기 작동과 모니터에 폭넓은 활용이 가능하므로 향후 전망이 매우 밝다고 하겠다.

## V. 결론 및 전망

무어의 법칙으로 대변되는 평면상 집적화는 극자외선 패터닝 공정으로 당분간 유지될 수 있을 것으로 예상되지만 평면을 탈피한 수직 적층 개념의 M3D는 집적밀도 및 스케일링 이외의 중요한 장점을 제공한다. 첫째, 동일 수의 소자를 좁은 다이 면적상에 제작할 수 있다. 또한 평면상의 각종 배선의 길이를 줄일 수 있어 전력 소모 및 신호 지연을 현저히 줄일 수 있다. 잠재적인 단점은 열 방출의 어려움이다. 구조 자체가 3차원 부피형이므로 M3D는 열소산에 불리한 측면이 있다. 이종의 기능의 조합은 발굴에 따라 가능한 경우가 굉장히 많기에 M3D는 잠재적으로 정체되어 있는 반도체 분야의 기능적 확장을 가능케 한다. 대표적으로 인공지능, 빅데이터, 뉴런 반도체, 자율작동기기 분야 등에 효율적인 대응이 가능케 할 수 있다. 하지만 M3D 기술이 전자응용분야에 상용화 수준으로 접근하기 위해서는 여러 가지 기술적 장벽이 존재하고 있다. 전술한 바와 같이 집적화 관점에서는 순차적 집적화가 유리하다고 판단된다. 이 접근법에서는 상부 구조 제작 시 하부 Si MOSFET 성능을 저하시키지 않은 온도가 선정되어야 한다. 고온 열처리 불가 제약 상황에서는 상부의 고품위의 액티브 재료 확보가 중요한 이슈라 할 수 있겠다. 중요한 후보로는 고이동도 산화물 반도체가 제시될 수

있다. 상·하부 연결 측면에서는 상·하부의 전기 연결이 중요하다. 작동 측면에서는 성능저하를 야기하는 기생용량을 제거할 수 있는 설계가 필요하다. 소재 측면에서는 고이동도(>50cm$^2$/Vs) 산화물 소재, 고종횡비 저저항 Via 물질, 고이동도 안정한 p-형 산화물 반도체 소재 등의 개발이 중요하다고 사료된다. M3D는 유효 스케일링 후속 단계 More than Moore를 기술적으로 구현해 줄 수 있는 캐치 프레이즈로 등장하고 있다. 응용을 염두에 둔 연구개발 단계에서는 기존의 존재하는 응용처를 대신하는 접근 방식이 아닌 반드시 M3D를 사용해야만 하는 응용처를 발굴하는 접근법이 중요하다고 사료된다. 유망 M3D 신규 응용분야로는 뇌 모사 소자 개발을 조심스럽게 제안한다. 하부에는 신호 증폭 및 프로세싱을 조절하는 CMOS 어레이가 위치하고 상부에는 산화물 기반 시냅스/뉴런 모사의 멤리스터 어레이가 위치하는 M3D 구조가 제안될 수 있다. 신규 기술은 초기에는 현안 한계를 극복하기 위하여 등장하는 경우가 대부분이다. 기술의 성숙도가 높아짐에 따라 신규 기술은 전혀 뜻하지 않은 방향으로 새로운 "무엇인가"로 구현되는 경우가 아주 빈번하게 목도되어 왔다. M3D도 이와 같은 과정을 거치리라 기대한다.

**용어해설**

**스케일링** 반도체 소자를 구성하는 요소들을 미세화함과 동시에 소자의 성능을 높이는 기술적 접근법

**FET(Field Effect Transistor)** 반도체막에 전기장을 인가하여 반도체막을 통하여 흐르는 전류의 양을 외부에서 조절할 수 있는 소자

**Monolithic 3D technology** 하부에 MOSFET 기판 소자 어레이가 위치하고 상부에 상이한 기능을 갖는 반도체층을 3차원으로 적층하는 전자소자 제작 기술

## 약어 정리

| | |
|---|---|
| AR | Augmented Reality |
| CMOS | Complementary Metal–Oxide–Semiconductor |
| DRAM | Dynamic Random Access Memory |
| EUV | Extreme Ultra Violet |
| FPGA | Field Programmable Gate Array |
| MOSFET | Metal–Oxide–Semiconductor Field–Effect Transistor |
| OLED | Organic Light Emitting Diode |
| TSV | Through Silicon Via |
| VR | Virtual Reality |

## 참고문헌


[1] G.E. Moore, "Cramming more components onto integrated circuits," Electronics, vol. 38, no. 8, Apr. 1965, pp. 114-117.

[2] K.J. Kuhn, "CMOS transistor scaling past 32nm and implications on variation," in Proc. IEEE/SEMI Adv. Semiconductor Manuf. Conf. (ASMC), San Francisco, CA, USA, July 2010, pp. 241-246, doi: 10.1109/ASMC.2010.5551461.

[3] P. Batude et al., "Advances, challenges and opportunities in 3D CMOS sequential integration," in Proc. Int. Electron Devices Meeting, Washington, DC, USA, Dec. 2011, pp. 7.3.1-7.3.4, doi: 10.1109/IEDM.2011.6131506.

[4] R. Courtland, "Transistors could stop shrinking in 2021," IEEE Spectrum, vol. 53, no. 9, Sept. 2016, pp. 9-11, doi: 10.1109/MSPEC.2016.7551335.

[5] M.T. Bohr and I.A. Young, "CMOS scaling trends and beyond," IEEE Micro, vol. 37, no. 6, Nov./Dec. 2017, pp. 20-29, doi: 10.1109/MM.2017.4241347.

[6] Samsung Newsroom, "Samsung's new 3D integration technology X-Cube(eXtended-Cube)," 20 2 0, Samsung, https://news.samsung.com/global/samsung-announces-availability-of-its-silicon-proven-3d-ic-technology-for-high-performance-applications.

[7] S. Salahuddin, K. Ni, and S. Datta, "The era of hyper-scaling in electronics," Nature Electron., vol. 1, 2018, pp. 442-450.

[8] S. Datta et al., "Back-end-of-line compatible transistors for monolithic 3-D integration," IEEE Micro, vol. 39, no. 6, Nov./Dec. 2019, pp. 8-15, doi: 10.1109/MM.2019.2942978.

[9] Y. Son et al., "Monolithic integration of high-voltage thin-film electronics on low-voltage integrated circuits using a solution process," Nat. Electron., vol. 2, 2019, pp. 540-548.

[10] Sony Semiconductor Solutions Corporation, "3-layer stacked CMOS image sensor with DRAM for smartphones," Sony Corporation, 2017, https://www.sony.net/SonyInfo/News/Press/201702/17-013E/

[11] S.H. Wu et al., "Performance boost of crystalline In-Ga-Zn-O material and transistor with extremely low leakage for IoT normally-off CPU application," in Proc. Symp. VLSI Circuits, Kyoto, Japan, June 2017, pp. T166-T167, doi: 10.23919/VLSIC.2017.8008580.

[12] J. Wu et al., "A monolithic 3D integration of RRAM array with oxide semiconductor FET for in-memory computing in quantized neural network AI applications," in Proc. IEEE Symp. VLSI Technol., Honolulu, HI, USA, June 2020, pp. 1-2, doi: 10.1109/VLSITechnology18217.2020.9265062.