

Flash memory의 기술 동향

최기환 (삼성전자)

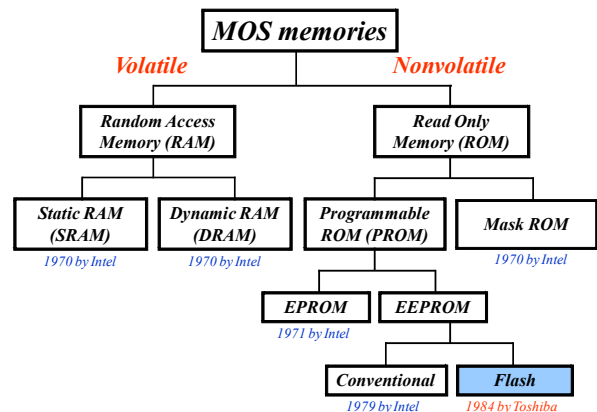
I. 서론

플래시 메모리(Flash Memory)는 전기적으로 데이터를 지우고 다시 기록할 수 있는 기억 장치를 말하는데, 잘 알려진 대로 전원이 공급되지 않을 경우에 데이터가 소멸되는 DRAM 및 SRAM 등과 같은 휘발성 메모리(Volatile Memory)와는 달리 전원이 공급되지 않더라도 기존에 저장된 데이터는 그대로 존재하는 불 휘발성 메모리 (Nonvolatile Memory)이다. 1980년대 초, 마쯔오카 교수에 의해 처음 만들어진 불 휘발성 메모리의 원형은 EPROM (Electrically Programmable ROM), EEPROM (Electrically Erasable & Programmable ROM)을 거쳐 현재 사용되고 있는 플래시 메모리의 모습으로 진화하였다.

현재 MP3 Player, Note Book, Smart Phone, Tablet PC 등 휴대용 전자 기기(Mobile Electronic Device)의 상용화 및 Green Technology에 대한 관심으로 인해 일상에서 흔히 볼 수 있는 USB (Universal Serial Bus) Drive를 포함 하여 eMMC(embedded Multi Media Card) 및SSD(Solid State Disk)에 이르기까지 다양한 응용 처에 있어서 플래시 메모리의 수요는 가히 폭발적으로 증가하고 있다. 이 같은 플래시 메모리 시장규모의 증가는 불 휘발성이라는 메모리로서의 매력적인 장점 이외에 기타 반도체 메모리에 비해 플래시 메모리 자체가 갖는 미세 공정의 편이성(Minimum Design Pitch 셀 array전체에 균일하게 사용됨) 및 이로 인한 높은 고 집적, 고 용량 기억장치 수단 확보 가능성에 기인한다고 해도 크게 틀리지 않을 것이다. 그러나, 플래시 메모리의 이러한 추세를 향후 지속적으로 유지하기 위해 가장 먼저 선행되어야 할 필수조건이 Technology Scaling Down인데, 여기에는 크게 두 가지 장애가 존재하고 있으며, 그 중 첫 번째는 Lithography

장비의 한계이다. 현재 알려져 있는 Lithography 기술(cf) Double Patterning 기법)로 구현할 수 있는 가장 작은 선 폭은 20nm근방으로 알려져 있으며, Quad Patterning 기법을 사용할 경우 ~15nm 까지도 가능하다고 보고되고 있지만^[1], 공정 cost의 증가 및 20nm급 제품이 2010년에 양산화되기 시작한 것을 고려하면 10nm급 제품 개발은 다소 지연될 것으로 예상되고 있다.

또 한가지 극복해야 할 커다란 문제는 셀 사이즈가 작아질 수록 점점 더 심각해지는 플래시 메모리 고유의 신뢰성(성능) 특성 열화이다. 일반적으로 셀 사이즈가 작아지면 “셀 산포 불 균일”, “인접 셀 간섭현상” 등 여러 가지 원인들에 의해서 플래시 메모리 자체의 성능, 즉 읽기(Read) 및 쓰기(Program) 속도, 개서 횟수 (Program/Erase Count: Endurance), 그리고 불 휘발성 메모리에 있어서 필수적으로 요구되는 데이터 저장시간 (Data Retention)등의 특성이 열화 되어 기존 세대 제품의 성능/신뢰성 특성을 유지시키는데 큰 어려움이 있다. 본 논문에서는 플래시 메모리 제품의 Scaling Down에 따라



〈그림 1〉 Tree of MOS Memory

발생하게 되는 문제점 및 이에 대한 극복 노력들을 살펴보고자 한다.

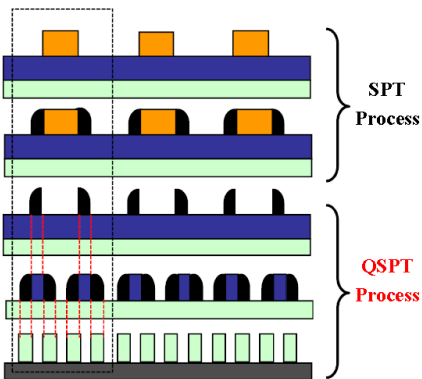
II. NAND 플래시 메모리

1. 기본 동작 : Program, Erase, Read

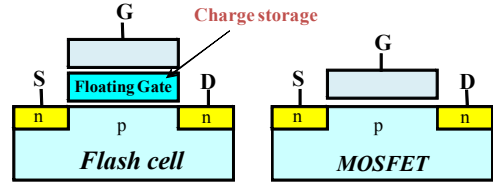
플래시 메모리는 기타 다른 메모리와 같이 '셀'이라 불리는 기본 저장 단위의 배열로 이루어져 있으며, 플래시 메모리 셀에는 비트 정보 저장을 위한 Floating Gate(FG)라는 Charge Storage Node가 존재한다. 셀의 구조는 두 개의 게이트를 가지고 있는 것을 제외하면, 표준 MOSFET과 동일하다. '0'과 '1'의 비트 정보는 각 셀의 Tunnel Oxide양단에 고전계를 인가하여 FG로 전자를 주입시키거나 (Program동작) 혹은 FG에서 전자를 빼냄으로써(Erase동작) 각각 저장된다. FG가 산화물 층에 의해 절연되어 있기 때문에 그 곳에 위치한 전자는 갇히게 되어 전원이 공급되지 않더라도 저장된 정보는 사라지지 않게 된다. 전자가 FG에 있을 때, Control Gate(CG)에서 나오는 전기장에 영향을 주어 셀의 문턱 전압(Threshold Voltage, V_{th})이 변경된다. 이와 같이 CG에 특정 전압을 가하여 그 셀의 정보를 읽을 때, FG에 있는 전자의 수에 따라 V_{th} 가 다르기 때문에 전류가 흐르거나 흐르지 않는다. 이러한 전류의 흐름과 차단이 각각 '1'과 '0'로 판독되게 된다.

일반적으로 플래시 메모리는 낸드형(NAND-type)과 노어형(NOR-type)이 있는데 이는 플래시 메모리 기본 소자의 구성형태가 마치 Logic Gate의 NAND 및 NOR gate의 Pull Down을 형성하고 있는 부분의 모양과 유사한 데에서 유래한다.

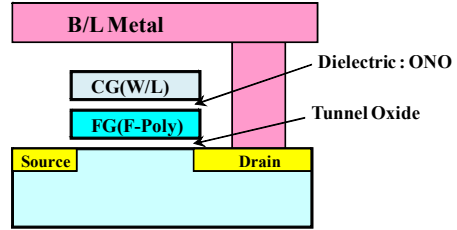
메모리 셀 unit의 contact 수가 적은 NAND 플래시 메모리는 고집적화가 용이하여 고 용량 데이터 저장용으로 사용되는 반면에, Byte단위의 고속 sensing이 가능한 NOR 플래시 메모리는 빈번하게 개사하는 저 용량 데이터 저장용으로 주



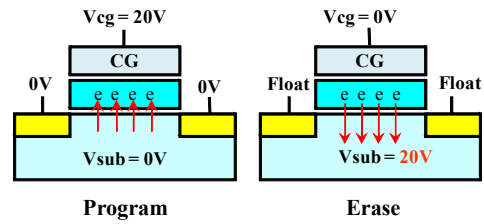
〈그림 2〉 Schematic diagram of QSPT (Quad Spacer Patterning Technology) ^[1]



(a) 플래시 셀 vs. MOSFET



(b) 플래시 셀 구조



(c) Program & Erase 동작

〈그림 3〉 플래시 셀 구조 및 기본 동작

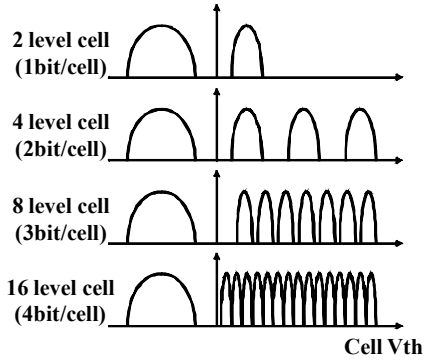
	NAND	NOR
Cell Array & Size	$2F \times 2.5F$	$2F \times 5F$
Cross-section		
Features	Small Cell Size, High Density Low Power & Good Endurance → Mass Storage	Large Cell Current, Fast Random Access → Code Storage

〈그림 4〉 플래시 메모리 : 낸드형 및 노어형

로 사용하며 본 논문에서는 주로 NAND 플래시 메모리에 대해서 살펴보고자 한다.

2. Multi Level Cell(MLC) NAND Flash

고 용량 NAND 플래시 수요를 만족시키기 위해 실제 셀 사이즈를 줄이는 방법 이외에 셀 하나에 여러 비트를 저장하는 MLC (Multi Level Cell) 기술을 이용하는 방법이 있으며, 현재 20nm급 3bit/cell 제품까지 상용화 되어 있다. 〈그림 5〉에서 볼 수 있듯이 셀 당 저장 비트수가 늘어 날수록 셀 state 산포 간 간격이 줄어들게 되어 (혹은 겹치게 되어) 저장된 데이터 Read시 에러의 확률이 높아지게 된다. 기본적으로 MLC제품



〈그림 5〉 MLC NAND Flash Cell state

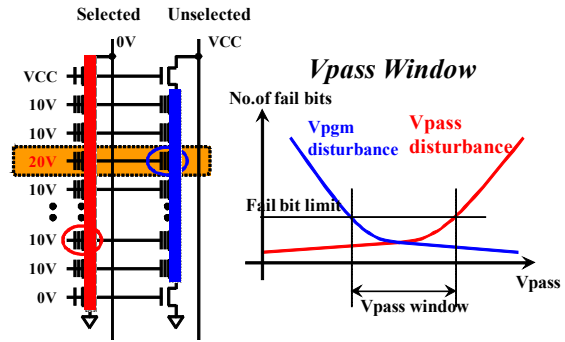
의 경우 각 셀 State의 산포 폭을 좁게 유지해야 하므로 프로그램 하는데 시간이 많이 걸리며, Read시에도 여러 번 읽어야 하므로 Read 성능도 떨어지게 된다. 3bit/cell 제품의 신뢰성(Endurance) 및 성능 (tPROG: Program Time)은 일반적으로 각각 ~0.5Kcycle, 2.5ms~3ms 정도인데, ~3Kcycle, 1~1.5ms 정도의 값을 갖는 2bit/cell 제품에 비하여 상대적으로 열세이지만, 그럼에도 불구하고 셀 size 축소 없이 동일한 공정을 사용하여 30%~40% (cf) 2bit vs. 3bit)의 생산성 향상을 가져올 수 있기 때문에 MLC 제품에 대한 관심은 지속적으로 유지되고 있다.

3. Disturbances

다른 메모리와 마찬가지로 NAND 플래시에서 가장 중요한 것은 기존에 저장된 ‘1’또는 ‘0’데이터가 차후 Read시 에러 없이 읽어 낼 수 있어야 하는데, 셀 사이즈가 작아질수록 저장된 데이터는 왜곡되어 에러가 발생할 가능성이 점점 더 증가하게 된다. 최근 통신분야에서 주로 사용되던 ECC기술이 NAND 플래시 응용에 있어서 필수 불가결한 요소가 되고 있는 것도 다 그 때문이라고 할 수 있다. 이렇게 NAND 플래시의 BER(Bit Error Rate)를 증가시키는 요인으로서는 데이터 기록시 존재하는 각종 Disturbance (Vpgm disturbance, Vpass disturbance), 저장된 Charge의 Loss, 및 인접 셀 간섭에 의한 산포 넓어짐 등이 있다.

가. Vpgm & Vpass Disturbance

NAND 플래시 메모리의 Program동작은 잘 알려진 바와 같이 F-N tunneling으로 이루어 지게 되는데, 선택 WL에는 Vpgm, 비 선택 WL에는 Vpass가 인가되고, 선택 및 비 선택 String의 BL에는 각각 0V 및 VDD를 인가함으로써 프로그램 및 프로그램 방지 (Program inhibit) 동작이 이루어 지게 된다. 이와 같은 프로그램 동작조건 하에서 크게2가지 Disturbance가 존재하는데 그 하나는 “선택된 String의 비 선택

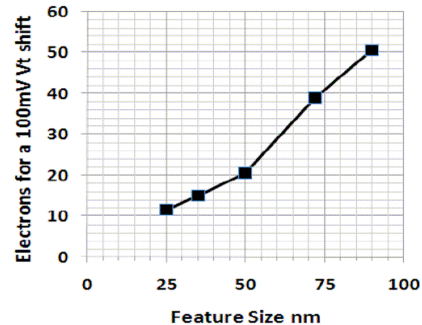


〈그림 6〉 Vpass window in NAND Flash

택 WL에 연결되어 있는 셀들”이 받는 *Vpass disturbance* 이고, 다른 하나는 “비 선택 String의 선택 WL에 연결되어 있는 Cell들”이 받는 *Vpgm disturbance* 이다. Vpgm이 높아질수록 원활한 프로그램 방지 동작을 위해 더 높은 Channel 전압이 요구되고, 그 결과 Vpass전압이 높아지게 되는데, 이때 높아진 Vpass전압에 의해 선택 String의 Cell에서 F-N Tunneling이 나타나게 되어 원치 않은 Program현상이 발생하게 된다. NAND 플래시의 Program동작은 해당 셀 특성에 의해 결정된 Vpgm전압에 대해 Program inhibit동작을 위한 적절한 Vpass전압이 존재하는가에 달려있는데, 이를 “*Vpass Window*”라고 부른다. 〈그림 6〉에서 볼 수 있는 바와 같이 Vpass전압이 너무 낮을 경우는 Vpgm disturbance에 의해서, 그리고 너무 높을 경우는 Vpass disturbance에 의해서 Fail bit이 증가하게 된다.

나. Charge loss

플래시 메모리를 일정시간 동안 방치하게 되면, 기존에 저장된 데이터는 여러 가지 원인에 의하여 변할 수 있는데, 그 정도는 주변 온도 및 Endurance횟수 등에 따라서 달라지게 된다. 이러한 경향은 셀 사이즈가 감소 할수록 점점 심하게 나타나게 되는데, 이는 셀 사이즈 감소에 따른 FG Capacitance 변화에 의해 해당 FG에 존재하는 전자의 수가 동시에 감소하기 때문이다. 〈그림 7〉에 의하면 25nm MLC의 경우 셀 Vth



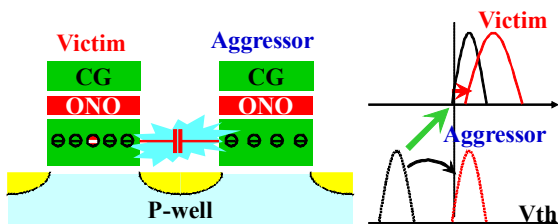
〈그림 7〉 Electrons required for a 100mV Vt shift vs. cell feature size. [2]

100mV 를 위해서 필요한 전자수가 10여개 정도밖에 되지 않는데, 동일한 개수의 전자가 Tunnel Oxide나 ONO 막을 통하여 사라질 경우를 가정해보면 셀 사이즈가 작을수록 Charge loss에 의한 셀 V_{th} 변화폭이 커지는 것은 당연하다고 볼 수 있다.

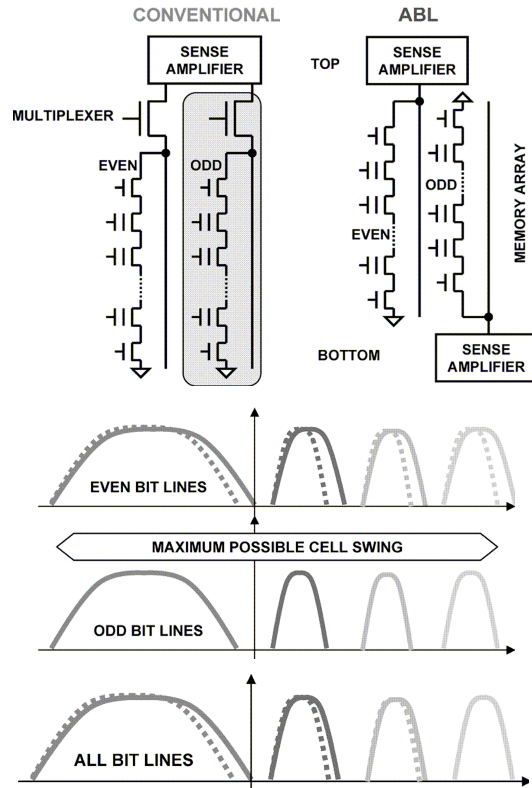
다. 인접 셀 간섭 현상

플래시 메모리의 대용량화를 위해 셀 사이즈를 줄이게 되면 가장 먼저 발생하는 문제는 인접 셀 간섭현상인데 이는 셀 간 거리가 가까워질수록 셀간의 FG사이에 Capacitive coupling이 증가하게 되어 해당 셀 V_{th} 가 인접 셀의 V_{th} 변화량에 따라 왜곡되는 현상이다 (called Yupin effect^[3]). 이러한 현상을 극복하기 위해 다양한 방법들이 사용되고 있는데, 셀 Program 구동방식을 변경하거나 혹은 공정적으로 셀간 Capacitive coupling을 감소시키는 노력들이 보고 되고 있다.

기존 플래시 메모리Chip architecture는 일종의 Sense amp. 및 write driver역할을 하는 Page buffer가 두 개의 비트라인(even & odd)에 의해 공유되는 구조인데, Program시 even 비트라인 셀을 먼저 program한 후 odd 비트라인 셀을 program하게 되는데, 그림에서 보이는 것처럼 even 비트라인 셀들은 odd 비트라인 셀들이 program되고 나면 산포의 왜곡이 생기게 된다. 이러한 현상은 기본적으로 인접 셀들이 Program되는 시점이 각각 다르기 때문에 발생하게 되는데, 이러한 시간적 차이를 없애거나 줄일 수 있다면 간섭현상을 효과적으로 억제할 수 있다. 그 예로 각 비트라인 당 Page buffer가 존재하는 All-bit Line Architecture는 비트라인 방향의 셀들을 동시에 program 할 수 있기 때문에 비트라인 방향의 간섭현상을 억제하는데 효과적이며^[4], 해당 셀의 program시 Target V_{th} 보다 낮은 레벨로 1차 program한 뒤 인접 셀에 의한 coupling후 최종 V_{th} level로 program하는 Reprogram 방식은 좋은 예라 할 수 있다^[5]. 설계적인 방법 이외에 공정적으로 FG간 Capacitive coupling 크기 자체를 감소시키는 방법도 보고 되고 있는데, WL사이에 유전물질 대신 Air gap을 형성시킬 경우 인접 셀 간섭현상을 약 25% 줄일 수 있다.^[2,6]



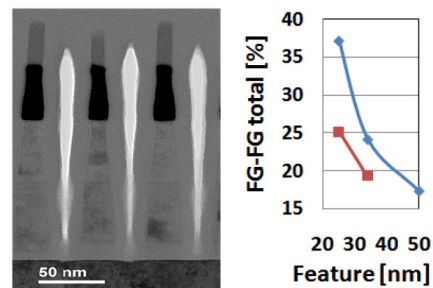
〈그림 8〉 인접 셀 간섭현상(Yupun effect)



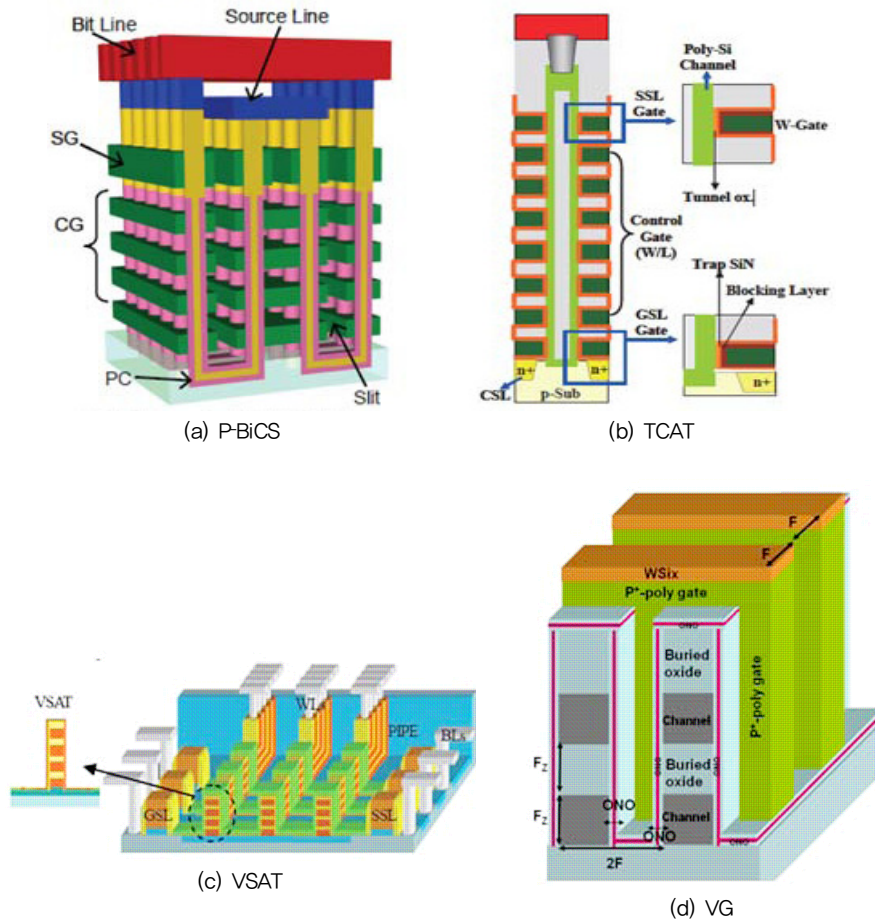
〈그림 9〉 All-bit line architecture^[4]

III. 3D NAND Flash memory

최근 < 10nm 급 NAND 플래시 제품에 필요한 Lithography 기술 개발이 지연됨에 따라 이에 대한 대응책으로 3D NAND Flash 에 관한 연구가 활발히 진행되고 있다^[7-11]. 기본적으로 3D NAND 플래시 Cell Array는 각 Cell Layer를 Stack 형식으로 쌓아 올려 형성시키기 때문에, 앞에서 언급한 2D NAND Flash Scaling Down 시 큰 장애가 되고 있는 Lithography 문제 및 인접 셀 간섭 등의 현상을 많은 부분 완화시킬 수 있다^[12]. 3D Flash Memory Architecture의 선구



〈그림 10〉 Cross-section of the cell in the WL direction showing the WL airgap and reduction in total FG-FG coupling with airgap (red square) and without (blue diamond).^[5]



〈그림 11〉 3D NAND Flash architectures^[13]

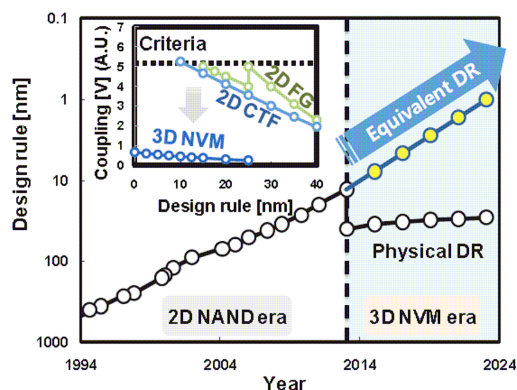
자적 역할을 한 Bit-cost scalable (BiCS) ^[7] 를 비롯하여 BiCS의 공정적 문제(tunnel oxide damage)를 개선한 P-BiCS ^[8], 그리고 Metal Gate 공정이 가능한 TCAT ^[9], twisted-channel을 사용한 VSAT ^[10], 셀 전류가 수직방향이 아닌 수평방향으로 흐르게 되는 Vertical gate (VG)^[11] 등 많은 Architecture 들이 제시되고 있다.

그러나, 3D 플래시 제품의 고층화에 따른 공정 난이도 증가 및 이에 따른 Cost 증가, 그리고 대부분의 3D NAND 플래

시에서 채용하고 있는 Nitride-based storage cell(SONOS 혹은 MANOS)이 가지고 있는 전하 재배열 (Charge re-distribution)에 의한 셀 산포 조절의 어려움 및 electron back tunneling에 의한 산포 넓어짐 등 여러 가지 극복해야 할 문제들이 존재하고 있다.

IV. 결론

현재 NAND Flash Memory는 고 집적화 및 고 용량화의 요구에 직면하고 있으며, 이에 부응하기 위해 Technology scaling down 및 이에 따른 제품의 성능 및 신뢰성 열화 등을 해결하기 위해 다각적인 노력을 하고 있다. 그 노력의 일환으로 BER 개선을 위한 ECC 기술 채택, Chip Architecture 변화 및 NAND 셀 구동 방식 변경, 그리고 MLC기술 개발 등으로 Lithography 기술 정체 등의 한계에 대응하고 있다. 다른 한편으로는 3D NAND Flash 제품 개발에 박차를 가하여 2D NAND Flash보다 훨씬 저렴하면서도 대용량의 NAND Flash 제품이 곧 상용화 될 예정이다. 그 외에도 Flash Memory 자체를 대체할 수 있는 PC (Phase Change, 상변화



〈그림 12〉 The technical transition of NAND toward the 3D NVM era [12]



메모리) RAM 및 다결정상태에 따른 저항차이를 이용한 ReRAM 등의 차세대 메모리에 관한 연구도 활발하게 진행되고 있어 지속적으로 다가올 Mobile Electronic Device 시대에 대한 준비를 하고 있다.

참고문헌

- [1] J. Hwang, et al, IEDM, pp.9.1.1-9.1.4, 2011.
- [2] K. Prall, et al, IEDM, pp.102-103, 2010.
- [3] J. Chen, et al, US Patent 5,867,429, Feb. 2, 1999.
- [4] R. Cernea, et al, IJSSC, Vol.44, No.1, Jan., 2009.
- [5] D. Lee, et al, ISSCC, pp.430-432, 2012.
- [6] C. Lee, et al, IEDM, 2010, pp. 98-101, 2010.
- [7] H. Tanaka, et al, VLSI Symp., pp.14-15, 2007.
- [8] R. Katsumata, et al, VLSI Symp., pp.136-137, 2009.
- [9] J. Jang, et al, VLSI Symp., pp.192-193, 2009.
- [10] J. Kim, et al, VLSI Symp., pp.186-187, 2009.
- [11] W. Kim, et al, VLSI Symp., pp.188-189, 2009.
- [12] J. Choi, et al, VLSIT, pp.178-179, 2011.
- [13] Y. Hsiao, et al, IMW, pp.1-4, 2010.



최기환

1992년 2월 한양대학교 전자공학과(BS).
 1994년 2월 포항공과대학교 전자전기공학과(MS).
 2005년 5월 University of Southern California, 전기공학과 (Ph.D).
 1994년 3월~2000년 4월 삼성전자 플래시 메모리 설계.
 2000년 5월~2005년 4월 Univ. of Southern California, RA.
 2005년 5월~현재 삼성전자 플래시 메모리 설계.
 <관심분야> Flash memory design