



Flash Memory

เสนอ

อาจารย์วีรศักดิ์ เจริญรัตน์

จัดทำโดย

63102105112 นายอัศรพล พิกุลศรี

รายงานนี้เป็นส่วนหนึ่งของรายวิชาสถาปัตยกรรมคอมพิวเตอร์

(14122501)

ภาคเรียนที่ 1 ปีการศึกษา 2564

มหาวิทยาลัยราชภัฏสกลนคร

คำนำ

รายงานเล่มนี้จัดทำขึ้นเพื่อเป็นส่วนหนึ่งของรายวิชา สถาปัตยกรรมคอมพิวเตอร์ (Computer Architecture) รหัสวิชา 14122501 เพื่อให้ได้ศึกษาหาความรู้เกี่ยวกับ Flash Memory ว่ามีความเป็นมาอย่างไร และมีลักษณะอย่างไร

ผู้จัดทำหวังว่ารายงานเล่มนี้จะเป็นประโยชน์กับผู้อ่านหรือผู้ที่กำลังหาข้อมูลเรื่องนี้อยู่หากมีข้อเสนอแนะหรือข้อผิดพลาดประการใด ผู้จัดทำขอน้อมรับไว้และขออภัยมา ณ ที่นี้ด้วย

นายอัศวพล พิกุลศรี

สารบัญ

หน้าปก	ก
คำนำ	ข
สารบัญ	ค
ประวัติความเป็นมา	1
หลักการทำงาน	2
ประตูลอย MOSFET (Floating-gate MOSFET)	2
อุโมงค์ฟาวเลอร์-นอร์ดไฮม์ (Fowler–Nordheim tunneling)	4
ปั๊มชาร์จภายใน (Internal charge pumps)	4
NOR flash	5
Programming	6
Erasing	7
NAND flash	8
การเขียนและการลบ (Writing and erasing)	9
Vertical NAND	10
Construction	11
ประสิทธิภาพ	11
ค่าใช้จ่าย (Cost)	12
ข้อจำกัด	12
บดบังการลบ	12
หน่วยความจำสวมใส่ (Memory wear)	13

การรบกวนการอ่าน (Read disturb)	14
การเข้าถึงระดับต่ำ	15
NOR memories	15
NAND memories	16
มาตรฐาน (Standardization)	18
ความแตกต่างระหว่าง NOR และ NAND Flash	19
ระบบไฟล์แฟลช	20
ความจุ	21
อัตราการโอน	22
แอปพลิเคชัน	23
แฟลชอนุกรม (Serial flash)	23
การจัดเก็บเฟิร์มแวร์ (Firmware storage)	24
หน่วยความจำแฟลชแทนฮาร์ดไดรฟ์ (Flash memory as a replacement for hard drives)	25
ที่เก็บถาวรหรือการจัดเก็บระยะยาว (Archival or long-term storage)	26
การเก็บรักษาข้อมูล (Data retention)	26
การกำหนดค่า FPGA (FPGA configuration)	27
อุตสาหกรรม	27
ความสามารถในการปรับขนาดแฟลช (Flash scalability)	27
อ้างอิง	28

ประวัติความเป็นมา

ต้นกำเนิดของหน่วยความจำแฟลช ย้อนไปถึงการพัฒนาของ MOSFET แบบ Floating-gate (FGMOS) หรือเรียกว่าทรานซิสเตอร์แบบ Floating-gate

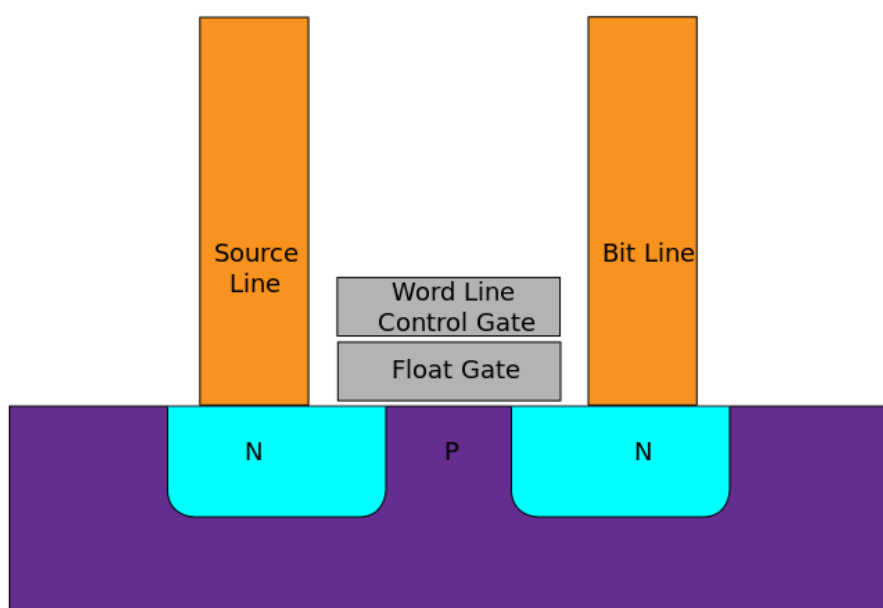
MOSFET ดั้งเดิม ทรานซิสเตอร์แบบ field-effect (ของโลหะออกไซด์-เซมิคอนดักเตอร์) หรือที่รู้จักในชื่อทรานซิสเตอร์ MOS ถูกคิดค้นโดยวิศวกรชาวอียิปต์ Mohammed M. Atalla และวิศวกรชาวเกาหลี Dawon Kahng ที่ Bell Labs ในปี 1959

ต่อมาในปี 1967 Kahng กับชาววิศวกรชาวจีน Simon Min Sze ที่ Bell Labs ได้พัฒนารูปแบบอื่นคือ Floating-gate MOSFET ซึ่งได้เสนอว่าสามารถใช้เป็นเซลล์หน่วยความจำแบบ Floating-gate เพื่อจัดเก็บรูปแบบของหน่วยความจำแบบอ่านอย่างเดียวที่ตั้งโปรแกรมได้ (Programmable read-only memory) ซึ่งเป็นทั้งแบบไม่ลบเลือนและตั้งโปรแกรมซ้ำได้

หน่วยความจำแบบ Floating-gate รุ่นแรกๆ ได้แก่ EPROM (PROM ที่ลบได้) และ EEPROM (PROM แบบลบด้วยไฟฟ้า) ในปี 1970 อย่างไรก็ตาม หน่วยความจำแบบ Floating-gate ในยุคแรกๆ ต้องการให้วิศวกรสร้างเซลล์หน่วยความจำสำหรับข้อมูลแต่ละบิต ซึ่งพิสูจน์แล้วว่าช้าและมีราคาแพง โดยจำกัดหน่วยความจำแบบ Floating-gate ให้กับแอปพลิเคชันเฉพาะในช่วงทศวรรษ 1970 เช่นเป็นอุปกรณ์ทางการทหารและโทรศัพท์มือถือรุ่นทดลองรุ่นแรกสุด

หลักการทำงาน

หน่วยความจำแฟลชจัดเก็บข้อมูลในอาร์เรย์ของเซลล์หน่วยความจำที่สร้างจากทรานซิสเตอร์แบบ Floating-gate ในอุปกรณ์เซลล์ระดับเดียว (SLC) แต่ละเซลล์จะเก็บข้อมูลเพียงบิตเดียว อุปกรณ์เซลล์หลายระดับ (MLC) รวมถึงอุปกรณ์เซลล์สามระดับ (TLC) สามารถจัดเก็บได้มากกว่าหนึ่งบิตต่อเซลล์ Floating gate อาจเป็นสื่อกระแสไฟฟ้า (โดยทั่วไปคือโพลีซิลิคอนใน หน่วยความจำแฟลช SONOS)



ประตูลอย MOSFET (Floating-gate MOSFET)

ในหน่วยความจำแฟลช แต่ละเซลล์หน่วยความจำจะมีลักษณะคล้ายกับทรานซิสเตอร์แบบ field-effect ของโลหะออกไซด์-เซมิคอนดักเตอร์มาตรฐาน (MOSFET) ยกเว้นว่าทรานซิสเตอร์มีเกทสองเกตแทนที่จะเป็นหนึ่งเกต เซลล์สามารถมองได้ว่าเป็นสวิตช์ไฟฟ้าที่กระแสไหลระหว่างขั้วทั้งสอง (แหล่งจ่ายและที่ระบาย) และถูกควบคุมโดย Floating gate (FG) และ control gate (CG)

CG นั้นคล้ายกับเกตในทรานซิสเตอร์ MOS อื่นๆ แต่ด้านล่างมี FG ที่หุ้มด้วยชั้นออกไซด์ โดยรอบ FG ถูกสอดแทรกระหว่างช่อง CG และ MOSFET เนื่องจาก FG ถูกแยกด้วยไฟฟ้าโดยชั้นฉนวน อิเล็กตรอนที่วางอยู่บนนั้นจึงติดอยู่ เมื่อ FG ถูกประจุด้วยอิเล็กตรอน ประจุนี้จะคัดกรองสนามไฟฟ้าจาก CG ดังนั้นจึงเป็นการเพิ่มแรงดันธรณีประจุ (VT1) ของเซลล์

ซึ่งหมายความว่าขณะนี้ต้องใช้แรงดันไฟฟ้าที่สูงขึ้น (VT2) กับ CG เพื่อให้ช่องนำไฟฟ้าเพื่ออ่านค่าจากทรานซิสเตอร์แรงดันไฟฟ้าปานกลางระหว่างแรงดันไฟฟ้าเกณฑ์

หากช่องสัญญาณดำเนินการที่แรงดันไฟฟ้าระดับกลางนี้ FG จะต้องไม่ถูกชาร์จ (หากมีการชาร์จ เราจะไม่ได้รับการนำไฟฟ้าเนื่องจากแรงดันไฟฟ้าปานกลางน้อยกว่า VT2) และด้วยเหตุนี้ "1" ทางตรรกะจึงถูกเก็บไว้ในเกต หากช่องสัญญาณไม่ทำงานที่แรงดันไฟฟ้าปานกลาง แสดงว่ามีการชาร์จ FG และด้วยเหตุนี้ จึงเก็บตรรกะ "0" ไว้ในเกต การมีอยู่ของตรรกะ "0" หรือ "1" นั้นรับรู้ได้โดยการพิจารณาว่ามีกระแสไหลผ่านทรานซิสเตอร์หรือไม่เมื่อแรงดันไฟฟ้าระดับกลางถูกยืนยันบน CG ในอุปกรณ์เซลล์แบบหลายระดับ ซึ่งเก็บมากกว่าหนึ่งบิตต่อเซลล์ จะตรวจนับปริมาณกระแสไฟ (แทนที่จะเป็นเพียงการมีอยู่หรือไม่มีอยู่) เพื่อกำหนดระดับประจุบน FG ได้แม่นยำยิ่งขึ้น

Floating gate(ประตูลอย) MOSFETs ตั้งชื่ออย่างนั้นเพราะมีชั้นอุโมงค์ออกไซด์ที่เป็นฉนวนไฟฟ้าระหว่าง Floating gate กับซิลิกอน ดังนั้นประตูจึง "Floats(ลอย)" เหนือซิลิกอน ออกไซด์ทำให้อิเล็กตรอนถูกกักขังไว้ที่ประตูลอย การเสื่อมสภาพหรือการสึกหรอ (และความทนทานที่จำกัดของหน่วยความจำแฟลชแบบลอยตัว) เกิดขึ้นเนื่องจากสนามแม่เหล็กที่สูงมาก (10 ล้านโวลต์ต่อเซนติเมตร) ที่เกิดขึ้นจากออกไซด์

ความหนาแน่นของไฟฟ้าแรงสูงดังกล่าวสามารถทำลายพันธะอะตอมเมื่อเวลาผ่านไป ในออกไซด์ที่ค่อนข้างบาง ค่อยๆ ลดระดับคุณสมบัติการเป็นฉนวนไฟฟ้าและปล่อยให้อิเล็กตรอนถูกกักขังและผ่านได้อย่างอิสระ (เร็ว) จากประตูลอยสู่ออกไซด์ เพิ่มโอกาสที่ข้อมูลจะสูญหาย เนื่องจากอิเล็กตรอน (ปริมาณที่ใช้เพื่อแสดงระดับประจุที่แตกต่างกัน ซึ่งแต่ละอิเล็กตรอนจะกำหนดให้กับบิตที่แตกต่างกันใน MLC Flash) โดยปกติแล้วจะอยู่ที่ประตูลอย นี่คือเหตุผลที่การเก็บรักษาข้อมูลลดลงและความเสี่ยงของการสูญเสียข้อมูลจะเพิ่มขึ้นตามการลดลงที่เพิ่มขึ้น

อุโมงค์ฟาวเลอร์-นอร์ดไฮม์ (Fowler–Nordheim tunneling)

กระบวนการย้ายอิเล็กตรอนจากประตูควบคุมไปยังประตูลอยเรียกว่าอุโมงค์ Fowler–Nordheim และเปลี่ยนแปลงลักษณะเฉพาะของเซลล์โดยพื้นฐานโดยการเพิ่มแรงดันธรณีประตูของ MOSFET ในทางกลับกันสิ่งนี้จะเปลี่ยนกระแสแหล่งจ่ายที่ไหลผ่านทรานซิสเตอร์สำหรับแรงดันเกตที่กำหนด

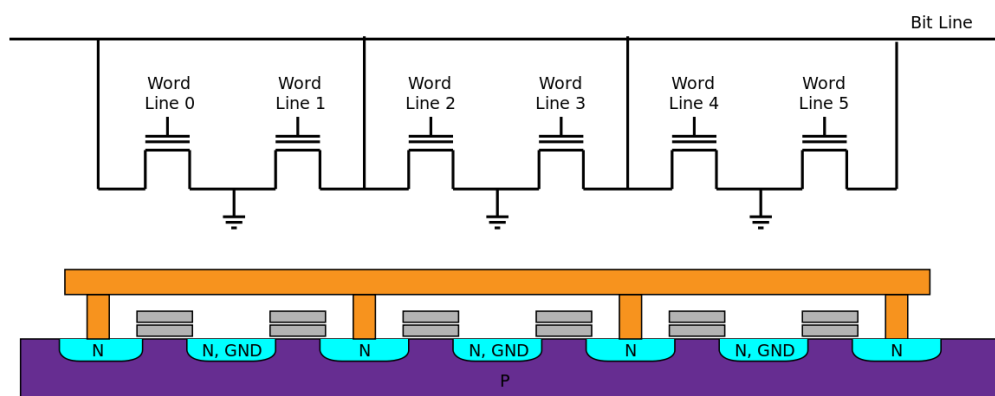
ซึ่งท้ายที่สุดแล้วจะใช้เพื่อเข้ารหัสค่าไบนารี เอฟเฟกต์อุโมงค์ Fowler–Nordheim สามารถย้อนกลับได้ ดังนั้นจึงสามารถเพิ่มหรือนำอิเล็กตรอนออกจากประตูลอยได้ ซึ่งกระบวนการที่เรียกกันทั่วไปว่าการเขียนและการลบ

ปั๊มชาร์จภายใน (Internal charge pumps)

ชิปแฟลชแทบทั้งหมดในปัจจุบันต้องการแรงดันไฟฟ้าเพียงแหล่งเดียวและผลิตไฟฟ้าแรงสูงที่ต้องการโดยใช้ปั๊มประจุบนชิป พลังงานที่ใช้โดยชิปแฟลช NAND ขนาด 1.8 V สูญเสียไปในปั๊มชาร์จเอง เนื่องจากบูสต์คอนเวอร์เตอร์มีประสิทธิภาพมากกว่าปั๊มชาร์จโดยเนื้อแท้ นักวิจัยที่พัฒนา SSD พลังงานต่ำได้เสนอให้กลับไปใช้แรงดันไฟฟ้า V_{cc}/V_{pp} คู่ที่ใช้กับชิปแฟลชรุ่นก่อนๆ ทั้งหมด

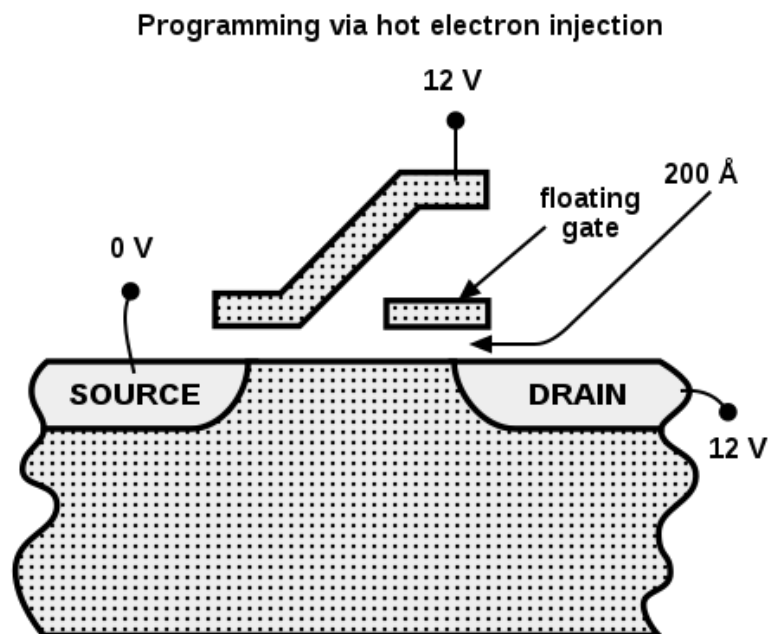
ซึ่งจะทำให้แรงดันไฟฟ้า V_{pp} สูงสำหรับชิปแฟลชทั้งหมดใน SSD ด้วยตัวเดียว ตัวแปลงบูสต์ภายนอกที่ใช้ร่วมกัน

NOR flash



ในแฟลช NOR แต่ละเซลล์มีปลายด้านหนึ่งเชื่อมต่อโดยตรงกับกราวด์ และปลายอีกด้านหนึ่งเชื่อมต่อโดยตรงกับเส้นบิต การจัดเรียงนี้เรียกว่า "แฟลช NOR" เพราะมันทำหน้าที่เหมือนเกต NOR: เมื่อบรรทัดคำใดบรรทัดหนึ่ง (เชื่อมต่อกับ CG ของเซลล์) ถูกทำให้สูง ทรานซิสเตอร์หน่วยเก็บข้อมูลที่เกี่ยวข้องจะดึงบรรทัดบิตเอาต์พุตให้ต่ำ แฟลช NOR ยังคงเป็นเทคโนโลยีทางเลือกสำหรับการใช้งานแบบฝังตัวที่ต้องการอุปกรณ์หน่วยความจำแบบไม่ลบเลือนแบบแยกส่วน คุณลักษณะเวลาแฝงในการอ่านต่ำของอุปกรณ์ NOR ช่วยให้สามารถใช้โค้ดได้โดยตรงและการจัดเก็บข้อมูลในผลิตภัณฑ์หน่วยความจำเดียว

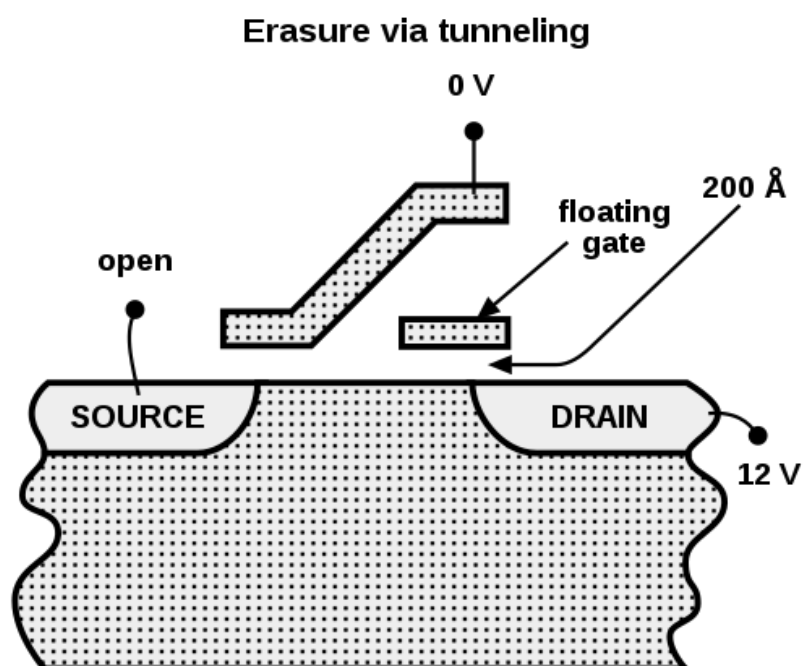
Programming



เซลล์แฟลช NOR ระดับเดียวในสถานะเริ่มต้นจะเทียบเท่ากับค่าไบนารี "1" ในทางตรรกะ เนื่องจากกระแสจะไหลผ่านช่องสัญญาณภายใต้การใช้แรงดันไฟฟ้าที่เหมาะสมกับเกตควบคุม เพื่อให้แรงดันบิตไลน์ถูกดึงลง สามารถตั้งโปรแกรมเซลล์แฟลช NOR หรือตั้งค่าเป็นค่าไบนารี "0" โดยทำตามขั้นตอนต่อไปนี้:

- แรงดันสูง (โดยทั่วไป >5 V) ใช้กับ CG
- ตอนนี้อิเล็กตรอนจึงสามารถไหลจากแหล่งกำเนิดไปยังท่อระบายน้ำได้ (สมมติว่าเป็นทรานซิสเตอร์ NMOS)
- กระแสไฟที่ไหลออกจากแหล่งกำเนิดสูงเพียงพอที่จะทำให้อิเล็กตรอนพลังงานสูงบางตัวกระโดดผ่านชั้นฉนวนไปยัง FG ผ่านกระบวนการที่เรียกว่าการฉีดอิเล็กตรอนร้อน

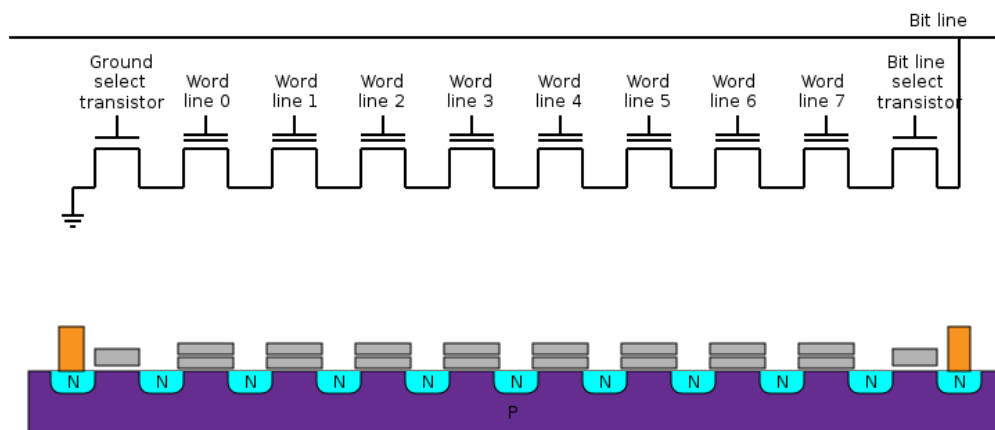
Erasing



ในการลบเซลล์แฟลช NOR (รีเซ็ตเป็นสถานะ "1") จะใช้แรงดันไฟฟ้าขนาดใหญ่ของขั้วตรงข้ามระหว่าง CG และขั้วต้นทาง โดยดึงอิเล็กตรอนออกจาก FG ผ่านอุโมงค์ควอนตัม ซิป หน่วยความจำแฟลช NOR สมัยใหม่แบ่งออกเป็นส่วนการลบ (มักเรียกว่าบล็อกหรือเซกเตอร์)

การลบสามารถทำได้บนพื้นฐานบล็อกเท่านั้น เซลล์ทั้งหมดในส่วนการลบจะต้องถูกลบพร้อมกัน อย่างไรก็ตาม การเขียนโปรแกรมของเซลล์ NOR โดยทั่วไปสามารถทำได้ครั้งละหนึ่งไบต์หรือคำ

NAND flash



แฟลช NAND ยังใช้ทรานซิสเตอร์แบบ Floating-gate แต่เชื่อมต่อในลักษณะที่คล้ายกับเกท NAND: ทรานซิสเตอร์หลายตัวเชื่อมต่อแบบอนุกรม และเส้นบิตจะถูกดึงให้ต่ำก็ต่อเมื่อบรรทัดคำทั้งหมดถูกดึงขึ้นสูง (เหนือทรานซิสเตอร์ VT) จากนั้นกลุ่มเหล่านี้จะเชื่อมต่อผ่านทรานซิสเตอร์เพิ่มเติมบางตัวกับอาร์เรย์ไลน์บิตแบบ NOR ในลักษณะเดียวกับที่ทรานซิสเตอร์ตัวเดียวเชื่อมโยงในแฟลช NOR

เมื่อเปรียบเทียบกับแฟลช NOR การแทนที่ทรานซิสเตอร์ตัวเดียวด้วยกลุ่มที่เชื่อมโยงแบบอนุกรมจะเพิ่มระดับของการกำหนดแอดเดรสเพิ่มเติม ในขณะที่แฟลช NOR อาจระบุหน่วยความจำที่ละหน้า ตามด้วยคำ แฟลช NAND อาจระบุด้วยหน้า คำและบิต การกำหนดแอดเดรสระดับบิตเหมาะกับแอปพลิเคชันบิตซีเรียล (เช่น การจำลองฮาร์ดดิสก์) ซึ่งเข้าถึงได้ครั้งละหนึ่งบิตเท่านั้น ในทางกลับกัน แอปพลิเคชันแบบ Execute-in-place ต้องการการเข้าถึงทุกบิตในหนึ่งคำพร้อมกัน สิ่งนี้ต้องการที่อยู่ระดับคำ ไม่ว่าในกรณีใด โหมดการกำหนดแอดเดรสแบบบิตและแบบคำสามารถทำได้โดยใช้แฟลช NOR หรือ NAND

ในการอ่านข้อมูล ก่อนอื่นให้เลือกกลุ่มที่ต้องการ (ในลักษณะเดียวกับที่เลือกทรานซิสเตอร์ตัวเดียวจากอาร์เรย์ NOR) ถัดไป บรรทัดคำส่วนใหญ่จะถูกดึงขึ้นเหนือ VT ของบิตที่ตั้งโปรแกรมไว้ ในขณะที่บรรทัดหนึ่งถูกดึงขึ้นไปเหนือ VT ของบิตที่ถูกลบ กลุ่มอนุกรมจะดำเนินการ (และดึงบรรทัดบิตต่ำ) หากไม่ได้ตั้งโปรแกรมบิตที่เลือกไว้

แม้จะมีทรานซิสเตอร์เพิ่มเติม แต่การลดสายกราวด์และสายบิตช่วยให้มีเลย์เอาต์ที่หนาแน่นขึ้นและความจุต่อชิปที่มากขึ้น (จริง ๆ แล้วสายกราวด์และสายบิตกว้างกว่าเส้นในไดอะแกรมมาก) นอกจากนี้ แฟลช NAND มักจะได้รับอนุญาตให้มีข้อผิดพลาดจำนวนหนึ่ง (แฟลช NOR เช่นเดียวกับที่ใช้สำหรับ ROM ของ BIOS คาดว่าจะเป็น ปราชญ์จากข้อผิดพลาด) ผู้ผลิตพยายามเพิ่มปริมาณพื้นที่เก็บข้อมูลที่ใช้งานได้สูงสุดโดยลดขนาดของทรานซิสเตอร์ลง เซลล์แฟลช NAND สามารถอ่านได้โดยการวิเคราะห์การตอบสนองต่อแรงดันไฟฟ้าต่างๆ

การเขียนและการลบ (Writing and erasing)

แฟลช NAND ใช้การแทรกอุโมงค์สำหรับการเขียนและการปล่อยช่องสัญญาณสำหรับการลบ หน่วยความจำแฟลช NAND เป็นแกนหลักของอุปกรณ์จัดเก็บข้อมูล USB แบบถอดได้ที่เรียกว่าแฟลชไดรฟ์ USB รวมถึงรูปแบบการ์ดหน่วยความจำและโซลิดสเตตไดรฟ์ส่วนใหญ่ที่มีในปัจจุบัน

โครงสร้างลำดับชั้นของ NAND Flash เริ่มต้นที่ระดับเซลล์ซึ่งสร้างสตริง จากนั้นหน้า บล็อก ระบาย และท้ายที่สุด สตริงคือชุดของเซลล์ NAND ที่เชื่อมต่อซึ่งแหล่งที่มาของเซลล์หนึ่งเชื่อมต่อกับท่อระบายน้ำของเซลล์ถัดไป ขึ้นอยู่กับเทคโนโลยี NAND โดยปกติสตริงจะประกอบด้วยเซลล์ NAND 32 ถึง 128 เซลล์ สตริงถูกจัดระเบียบเป็นหน้าซึ่งจัดเป็นบล็อก

แต่ละสตริงเชื่อมต่อกับบรรทัดแยกที่เรียกว่า bitline (BL) เซลล์ทั้งหมดที่มีตำแหน่งเดียวกันในสตริงจะเชื่อมต่อผ่านประตูควบคุมด้วย wordline (WL) เครื่องบิน มีบล็อกจำนวนหนึ่งที่เชื่อมต่อผ่าน BL เดียวกัน แฟลชได้ประกอบด้วยระนาบอย่างน้อยหนึ่งระนาบ และวงจรต่อพ่วงที่จำเป็นสำหรับการดำเนินการอ่าน/เขียน/ลบทั้งหมด

สถาปัตยกรรมของ NAND Flash หมายความว่าข้อมูลสามารถอ่านและตั้งโปรแกรมในหน้าได้ โดยทั่วไปแล้วจะมีขนาดระหว่าง 4 KiB ถึง 16 KiB แต่สามารถลบได้เฉพาะที่ระดับของบล็อกทั้งหมดซึ่งประกอบด้วยหลายหน้าและขนาด MB เมื่อบล็อกถูกลบ เซลล์ทั้งหมดจะถูกตั้งค่าตามตรรกะเป็น 1 ข้อมูลสามารถตั้งโปรแกรมได้เพียงครั้งเดียวผ่านไปยังหน้าในบล็อกที่ถูกลบ

เซลล์ใดๆ ที่ตั้งค่าเป็น 0 โดยการเขียนโปรแกรมสามารถรีเซ็ตเป็น 1 ได้เท่านั้นโดยการลบ บล็อกทั้งหมด ซึ่งหมายความว่าก่อนที่จะสามารถตั้งโปรแกรมข้อมูลใหม่ลงในหน้าที่มีข้อมูลอยู่แล้ว เนื้อหาปัจจุบันของหน้าและข้อมูลใหม่จะต้องถูกคัดลอกไปยังหน้าใหม่ที่ถูกลบ หากมีหน้าที่เหมาะสม สามารถเขียนข้อมูลได้ทันที หากไม่มีหน้าที่ลบ ต้องลบบล็อกก่อนคัดลอกข้อมูลไปยังหน้าในบล็อกนั้น หน้าเก่าจะถูกทำเครื่องหมายว่าไม่ถูกต้องและสามารถลบและนำกลับมาใช้ใหม่ได้

Vertical NAND

หน่วยความจำ NAND แนวตั้ง (V-NAND) หรือ 3D NAND จะสแต็คเซลล์หน่วยความจำใน แนวตั้งและใช้สถาปัตยกรรมแฟลชดักประจุ เลเยอร์แนวตั้งช่วยให้มีความหนาแน่นบิตของพื้นที่ที่ใหญ่ ขึ้นโดยไม่ต้องใช้เซลล์แต่ละเซลล์ที่เล็กกว่า

นอกจากนี้ยังจำหน่ายภายใต้เครื่องหมายการค้า BiCS Flash ซึ่งเป็นเครื่องหมายการค้าของ Kioxia Corporation (อดีต Toshiba Memory Corporation) 3D NAND ได้รับการประกาศครั้งแรกโดยโตชิบาในปี 2550 V-NAND ถูกผลิตขึ้นเชิงพาณิชย์ครั้งแรกโดย Samsung Electronics ในปี 2556

โครงสร้าง

V-NAND ใช้รูปทรงแฟลชกับดักประจุ (ซึ่งเปิดตัวในเชิงพาณิชย์ในปี 2545 โดย AMD และ Fujitsu) ซึ่งเก็บประจุไว้บนฟิล์มซิลิคอนไนไตรด์ที่ฝังอยู่ ฟิล์มดังกล่าวมีความทนทานต่อจุดบกพร่อง และสามารถทำให้หนาขึ้นเพื่อเก็บอิเล็กตรอนจำนวนมากขึ้นได้ V-NAND ห่อหุ้มเซลล์ดักจับประจุ ระนาบให้อยู่ในรูปทรงกระบอก

ในปี 2020 หน่วยความจำ 3D NAND Flash โดย Micron และ Intel จะใช้เกตแบบลอย แทน อย่างไรก็ตาม หน่วยความจำ 3D NAND ที่เลเยอร์ 128 และสูงกว่านั้นใช้โครงสร้างกับดักประจุ แบบเดิม เนื่องจากการล่มสลายของความร่วมมือระหว่าง Micron และ Intel แพ่นชาร์จ 3D NAND Flash นั้นบางกว่าประตูลอย 3D NAND ใน Float Gate 3D NAND เซลล์หน่วยความจำจะถูกแยก ออกจากกันโดยสิ้นเชิง ในขณะที่ 3D NAND ของกับดักประจุ กลุ่มเซลล์หน่วยความจำในแนวตั้งใช้ วัสดุซิลิคอนไนไตรด์เดียวกัน

เซลล์หน่วยความจำแต่ละเซลล์ประกอบด้วยชั้นโพลีซิลิคอนที่มีระนาบหนึ่งชั้นซึ่งมีรูที่เต็มไปด้วยกระบอกสูบแนวตั้งที่มีศูนย์กลางหลายอัน พื้นผิวโพลีซิลิคอนของรูทำหน้าที่เป็นอิเล็กโทรดเกต กระบอกสูบซิลิกอนไดออกไซด์ที่อยู่บนสุดทำหน้าที่เป็นไดอิเล็กตริกเกตล้อมรอบกระบอกสูบซิลิกอน ไนไตรด์ที่เก็บประจุ ในทางกลับกันก็ปิดกั้นซิลิกอนไดออกไซด์เป็นไดอิเล็กตริกแบบอุโมงค์ที่ล้อมรอบ แกนกลางของตัวนำโพลีซิลิคอนซึ่งทำหน้าที่เป็นช่องทางนำไฟฟ้า

เซลล์หน่วยความจำในชั้นแนวตั้งที่ต่างกันไม่รบกวนซึ่งกันและกัน เนื่องจากประจุไม่สามารถเคลื่อนที่ในแนวตั้งผ่านสื่อเก็บซิลิคอนไนไตรด์ได้ และสนามไฟฟ้าที่เกี่ยวข้องกับเกตจะถูกจำกัดอย่างใกล้ชิดภายในแต่ละชั้น คอลเล็กชันแนวตั้งจะเหมือนกันทางไฟฟ้ากับกลุ่มที่ลิงก์อนุกรมซึ่งมีการกำหนดค่าหน่วยความจำแฟลช NAND แบบธรรมดา

Construction

การเติบโตของกลุ่มเซลล์ V-NAND เริ่มต้นด้วยชั้นโพลีซิลิคอนที่นำไฟฟ้า (doped) สลับกัน และชั้นซิลิกอนไดออกไซด์ที่เป็นฉนวน ขั้นตอนต่อไปคือการสร้างรูทรงกระบอกผ่านชั้นเหล่านี้ ในทางปฏิบัติ ชิป 128 Gbit V-NAND ที่มีเซลล์หน่วยความจำ 24 ชั้น ต้องการรูดังกล่าวประมาณ 2.9 พันล้านรู ถัดไป พื้นผิวด้านในของรูได้รับการเคลือบหลายชั้น ซิลิกอนไดออกไซด์ตัวแรก ซิลิกอนไนไตรด์ จากนั้นชั้นที่สองของซิลิกอนไดออกไซด์ ในที่สุด หลุมก็เต็มไปด้วยโพลีซิลิคอนตัวนำ (doped)

ประสิทธิภาพ

ในปี 2013 สถาปัตยกรรมแฟลช V-NAND ช่วยให้สามารถอ่านและเขียนข้อมูลได้เร็วกว่า NAND ทัวไปถึงสองเท่า และใช้งานได้ยาวนานขึ้นถึง 10 เท่า ในขณะที่ใช้พลังงานน้อยลง 50 เปอร์เซ็นต์ พวกเขาเสนอความหนาแน่นบิตทางกายภาพที่เปรียบเทียบได้โดยใช้การพิมพ์หิน 10 นาโนเมตร แต่อาจสามารถเพิ่มความหนาแน่นของบิตได้มากถึงสองลำดับความสำคัญ เนื่องจาก V-NAND มีการใช้งานมากถึงหลายร้อยเลเยอร์ ณ ปี 2020 ชิป V-NAND ที่มี 160 เลเยอร์ อยู่ระหว่างการพัฒนาโดยซัมซุง

ค่าใช้จ่าย (Cost)

ราคาของเวเฟอร์ของ 3D NAND นั้นเทียบได้กับระนาบ NAND Flash แบบลดขนาด (32 นาโนเมตรหรือน้อยกว่า) อย่างไรก็ตาม ด้วยการหยุดสเกล NAND ระนาบที่ 16 นาโนเมตร การลดต้นทุนต่อบิตสามารถดำเนินต่อไปโดย 3D NAND โดยเริ่มจาก 16 เลเยอร์

อย่างไรก็ตามเนื่องจากผนังไม่แนวตั้งของรูถูกสลักผ่านชั้นต่างๆ แม้แต่ความเป็ยงเบนเล็กน้อยก็นำไปสู่ต้นทุนบิตชั้นต่ำ เช่น กฎการออกแบบชั้นต่ำที่เทียบเท่า (หรือความหนาแน่นสูงสุด) สำหรับจำนวนเลเยอร์ที่กำหนด จำนวนเลเยอร์ต้นทุนบิตชั้นต่ำนี้ลดลงสำหรับเส้นผ่านศูนย์กลางรูที่เล็กกว่า

ข้อจำกัด

บล็อกการลบ

ข้อจำกัดอย่างหนึ่งของหน่วยความจำแฟลชคือ แม้ว่าจะสามารถอ่านหรือตั้งโปรแกรมทีละบิตหรือคำในรูปแบบการเข้าถึงโดยสุ่ม แต่ก็สามารถลบได้ครั้งละหนึ่งบล็อกเท่านั้น

โดยทั่วไปจะตั้งค่าบิตทั้งหมดในบล็อกเป็น 1 โดยเริ่มจากบล็อกที่เพิ่งลบใหม่ ตำแหน่งใดๆ ภายในบล็อกนั้นสามารถตั้งโปรแกรมได้ อย่างไรก็ตาม เมื่อบิตถูกตั้งค่าเป็น 0 โดยการลบบล็อกทั้งหมดเท่านั้น สามารถเปลี่ยนกลับเป็น 1 ได้ กล่าวอีกนัยหนึ่ง หน่วยความจำแฟลช (โดยเฉพาะแฟลช NOR) ให้การดำเนินการอ่านและเขียนโปรแกรมโดยสุ่มเข้าถึงแต่ไม่มีการสุ่มตามอำเภอใจ – เข้าถึงการดำเนินการเขียนใหม่หรือลบ อย่างไรก็ตาม ตำแหน่งสามารถเขียนใหม่ได้ราบใดที่ค่าใหม่เป็น 0 บิตเป็น superset ของค่าที่เขียนทับ ตัวอย่างเช่น ค่า nibble อาจถูกลบไปที่ 1111 จากนั้นเขียนเป็น 1110 การเขียนที่ต่อเนื่องไปยัง nibble นั้นสามารถเปลี่ยนเป็น 1,010 จากนั้น 0010 และสุดท้ายคือ 0000

โดยพื้นฐานแล้ว การลบจะตั้งค่าบิตทั้งหมดเป็น 1 และการเขียนโปรแกรมสามารถล้างบิตเป็น 0 เท่านั้นระบบไฟล์บางระบบที่ออกแบบมาสำหรับอุปกรณ์แฟลชใช้ประโยชน์จากความสามารถในการเขียนซ้ำนี้ เช่น Yaffs1 เพื่อแสดงข้อมูลเมตาของเซกเตอร์ ระบบไฟล์แฟลชอื่นๆ เช่น YAFFS2 ไม่เคยใช้ความสามารถ "เขียนซ้ำ" นี้เลย เนื่องจากระบบทำงานพิเศษหลายอย่างเพื่อให้เป็นไปตามกฎ

"เขียนครั้งเดียว" แม้ว่าโครงสร้างข้อมูลในหน่วยความจำแฟลชจะไม่สามารถอัปเดตในลักษณะทั่วไปได้อย่างสมบูรณ์ แต่สิ่งนี้ทำให้สมาชิกสามารถ "ลบ" โดยทำเครื่องหมายว่าไม่ถูกต้อง

เทคนิคนี้อาจจำเป็นต้องแก้ไขสำหรับอุปกรณ์เซลล์หลายระดับ โดยที่เซลล์หน่วยความจำหนึ่งเซลล์มีมากกว่าหนึ่งบิต

อุปกรณ์แฟลชทั่วไป เช่น แฟลชไดรฟ์ USB และการ์ดหน่วยความจำมีเฉพาะอินเทอร์เฟซระดับบล็อก หรือเลเยอร์การแปลงแฟลช (FTL) ซึ่งเขียนไปยังเซลล์อื่นในแต่ละครั้งเพื่อให้ระดับการสึกหรอของอุปกรณ์ สิ่งนี้จะป้องกันการเขียนที่เพิ่มขึ้นภายในบล็อก อย่างไรก็ตาม มันช่วยให้อุปกรณ์ไม่เสื่อมสภาพก่อนเวลาอันควรจากรูปแบบการเขียนที่เข้มข้น

หน่วยความจำสวมใส่ (Memory wear)

ข้อจำกัดอีกประการหนึ่งคือหน่วยความจำแฟลชมีจำนวนโปรแกรมที่จำกัด – รอบการลบ (โดยทั่วไปเขียนเป็นรอบ P/E) ผลิตภัณฑ์แฟลชที่มีจำหน่ายทั่วไปส่วนใหญ่รับประกันว่าสามารถทนต่อรอบ P/E ได้ประมาณ 100,000 รอบ ก่อนที่การสึกหรอจะทำให้ความสมบูรณ์ของการจัดเก็บลดลง [80] Micron Technology และ Sun Microsystems ได้ประกาศชิปหน่วยความจำแฟลช SLC NAND ที่มีอัตรา 1,000,000 P/E รอบในวันที่ 17 ธันวาคม 2551

จำนวนรอบที่รับประกันอาจใช้เฉพาะกับบล็อกศูนย์ (เช่นเดียวกับอุปกรณ์ TSOP NAND) หรือกับบล็อกทั้งหมด (เช่นใน NOR) ผลกระทบนี้จะลดลงในเฟิร์มแวร์ชิปบางตัวหรือไดรเวอร์ระบบไฟล์โดยการนับบล็อกการเขียนและการแมปใหม่แบบไดนามิกเพื่อกระจายการดำเนินการเขียนระหว่างเซกเตอร์ เทคนิคนี้เรียกว่าการปรับระดับการสึกหรอ อีกวิธีหนึ่งคือดำเนินการตรวจสอบการเขียนและทำการแมปใหม่ไปยังเซกเตอร์สำรองในกรณีที่เกิดข้อผิดพลาดในการเขียน ซึ่งเป็นเทคนิคที่เรียกว่าการจัดการบล็อกที่ไม่ดี (BBM)

สำหรับอุปกรณ์พกพาสำหรับผู้บริโภค เทคนิคการจัดการเหล่านี้มักจะช่วยยืดอายุหน่วยความจำแฟลชให้ยาวนานกว่าอายุของอุปกรณ์เอง และข้อมูลบางส่วนอาจสูญหายได้ในแอปพลิเคชันเหล่านี้

อย่างไรก็ตาม สำหรับการจัดเก็บข้อมูลที่มีความน่าเชื่อถือสูง ไม่แนะนำให้ใช้หน่วยความจำแฟลชที่จะต้องผ่านรอบการเขียนโปรแกรมจำนวนมาก ข้อจำกัดนี้ไม่มีความหมายสำหรับแอปพลิเคชันที่ 'อ่านอย่างเดียว' เช่น อินโคลเอ็นด์และเราเตอร์ ซึ่งถูกตั้งโปรแกรมเพียงครั้งเดียวหรือมากที่สุดสองสามครั้งในช่วงอายุการใช้งาน

การรบกวนการอ่าน (Read disturb)

วิธีที่ใช้ในการอ่านหน่วยความจำแฟลช NAND อาจทำให้เซลล์ใกล้เคียงในบล็อกหน่วยความจำเดียวกันเปลี่ยนแปลงเมื่อเวลาผ่านไป (กลายเป็นโปรแกรม) สิ่งนี้เรียกว่าการรบกวนการอ่าน จำนวนเกณฑ์การอ่านโดยทั่วไปอยู่ในการอ่านหลายแสนครั้งระหว่างการดำเนินการลบที่แทรกแซง หากอ่านอย่างต่อเนื่องจากเซลล์หนึ่ง เซลล์นั้นจะไม่ลืมหเลว แต่จะอ่านเซลล์รอบข้างหนึ่งเซลล์ในการอ่านครั้งถัดไป

เพื่อหลีกเลี่ยงปัญหาการรบกวนการอ่าน โดยทั่วไปตัวควบคุมแฟลชจะนับจำนวนการอ่านทั้งหมดไปยังบล็อกหนึ่งๆ นับตั้งแต่การลบครั้งล่าสุด เมื่อการนับเกินขีดจำกัดเป้าหมาย บล็อกที่ได้รับผลกระทบจะถูกคัดลอกไปยังบล็อกใหม่ ลบออก แล้วปล่อยไปยังบล็อกพูล บล็อกเดิมดีเหมือนใหม่หลังจากลบ อย่างไรก็ตาม หากตัวควบคุมแฟลชไม่เข้าไปแทรกแซงตามเวลา จะเกิดข้อผิดพลาดในการอ่านข้อมูลโดยที่ข้อมูลอาจสูญหายได้ หากข้อผิดพลาดมีมากเกินไปที่จะแก้ไขด้วยการรบกวนการอ่านข้อมูลโดยที่ข้อมูลอาจสูญหายได้ หากข้อผิดพลาดมีมากเกินไปที่จะแก้ไขด้วยการรบกวนการอ่านข้อมูลโดยที่ข้อมูลอาจสูญหายได้ หากข้อผิดพลาดมีมากเกินไปที่จะแก้ไขด้วยการรบกวนการอ่านข้อมูลโดยที่ข้อมูลอาจสูญหายได้

เอฟเฟกต์เอ็กซ์เรย์ (X-ray effects)

แฟลชไอซีส่วนใหญ่มาในแพ็คเกจ ball grid array (BGA) และแม้แต่วงจรที่ไม่ได้ติดตั้งบน PCB ถัดจากแพ็คเกจ BGA อื่นๆ หลังจากการประกอบ PCB บอร์ดที่มีแพ็คเกจ BGA มักจะถูกเอ็กซ์เรย์เพื่อดูว่าลูกบอลกำลังเชื่อมต่อกับแพดที่เหมาะสมหรือไม่ หรือหาก BGA จำเป็นต้องทำใหม่ รังสีเอกซ์เหล่านี้สามารถลบบิตที่ตั้งโปรแกรมไว้ในชิปแฟลช (แปลงบิต "0" ที่ตั้งโปรแกรมไว้เป็นบิต "1" ที่ถูกลบ) บิตที่ถูกลบ (บิต "1") จะไม่ได้รับผลกระทบจากรังสีเอกซ์

ผู้ผลิตบางรายกำลังผลิตอุปกรณ์หน่วยความจำ USB แบบ SD[90] ที่กันรังสีเอกซ์

การเข้าถึงระดับต่ำ

อินเทอร์เฟซระดับต่ำกับชิปหน่วยความจำแฟลชแตกต่างจากหน่วยความจำประเภทอื่นๆ เช่น DRAM, ROM และ EEPROM ซึ่งรองรับการเปลี่ยนแปลงบิต (ทั้งศูนย์ถึงหนึ่งและหนึ่งถึงศูนย์) และการเข้าถึงแบบสุ่มผ่านแอดเดรสบัสที่เข้าถึงได้จากภายนอก

หน่วยความจำ NOR มีบัสที่อยู่ภายนอกสำหรับการอ่านและการเขียนโปรแกรม สำหรับหน่วยความจำ NOR การอ่านและการตั้งโปรแกรมเป็นการเข้าถึงแบบสุ่ม และการปลดล๊อคและการลบจะเป็นแบบบล็อก สำหรับหน่วยความจำ NAND การอ่านและการตั้งโปรแกรมเป็นแบบเพจ และการปลดล๊อคและการลบจะเป็นแบบบล็อก

NOR memories

การอ่านจากแฟลช NOR นั้นคล้ายกับการอ่านจากหน่วยความจำเข้าถึงโดยสุ่ม โดยมีการระบุที่อยู่และบัสข้อมูลไว้อย่างถูกต้อง ด้วยเหตุนี้ ไมโครโปรเซสเซอร์ส่วนใหญ่จึงสามารถใช้หน่วยความจำแฟลช NOR เป็นหน่วยความจำแบบดำเนินการในสถานที่ (XIP) ได้ ซึ่งหมายความว่าโปรแกรมที่จัดเก็บไว้ในแฟลช NOR สามารถทำงานได้โดยตรงจากแฟลช NOR โดยไม่จำเป็นต้องคัดลอกลงใน RAM ก่อน แฟลช NOR อาจถูกตั้งโปรแกรมในลักษณะการเข้าถึงแบบสุ่มคล้ายกับการอ่าน การเขียนโปรแกรมเปลี่ยนบิตจากตรรกะหนึ่งเป็นศูนย์ บิตที่เป็นศูนย์อยู่แล้วจะไม่เปลี่ยนแปลง การลบจะต้องเกิดขึ้นทีละบล็อก และรีเซ็ตบิตทั้งหมดในบล็อกที่ถูกลบกลับไปเป็นบล็อกเดียว ขนาดบล็อกทั่วไปคือ 64, 128 หรือ 256 KiB

การจัดการบล็อกที่ไม่ดีเป็นคุณสมบัติที่ค่อนข้างใหม่ในชิป NOR ในอุปกรณ์ NOR รุ่นเก่าที่ไม่รองรับการจัดการบล็อกที่ไม่ดี ซอฟต์แวร์หรือไดรเวอร์อุปกรณ์ที่ควบคุมชิปหน่วยความจำจะต้องแก้ไขสำหรับบล็อกที่เสื่อมสภาพ ไม่เช่นนั้นอุปกรณ์จะหยุดทำงานอย่างน่าเชื่อถือ

คำสั่งเฉพาะที่ใช้ในการล๊อค ปลดล๊อค โปรแกรม หรือลบหน่วยความจำ NOR นั้นแตกต่างกันไปสำหรับผู้ผลิตแต่ละราย เพื่อหลีกเลี่ยงความต้องการซอฟต์แวร์ไดรเวอร์เฉพาะสำหรับอุปกรณ์ทุกชิ้นที่ผลิต คำสั่ง Common Flash Memory Interface (CFI) พิเศษช่วยให้อุปกรณ์สามารถระบุตัวเองและพารามิเตอร์การทำงานที่สำคัญได้

นอกจากการใช้เป็น ROM เข้าถึงโดยสุ่มแล้ว แฟลช NOR ยังสามารถใช้เป็นอุปกรณ์จัดเก็บข้อมูลได้ด้วยการใช้ประโยชน์จากการเขียนโปรแกรมเข้าถึงโดยสุ่ม อุปกรณ์บางอย่างมีฟังก์ชันอ่านขณะเขียนเพื่อให้โค้ดทำงานต่อไปได้แม้ในขณะที่โปรแกรมหรือการดำเนินการลบเกิดขึ้นในเบื้องหลังสำหรับการเขียนข้อมูลตามลำดับ โดยปกติชิป NOR แฟลชจะมีความเร็วในการเขียนที่ช้า เมื่อเทียบกับแฟลช NAND

แฟลช NOR ทั่วไปไม่ต้องการรหัสแก้ไขข้อผิดพลาด

NAND memories

สถาปัตยกรรมแฟลช NAND เปิดตัวโดยโตชิบาในปี 1989[93] ความทรงจำเหล่านี้เข้าถึงได้เหมือนกับอุปกรณ์บล็อก เช่น ฮาร์ดดิสก์ แต่ละบล็อกประกอบด้วยหลายหน้า หน้าโดยทั่วไปจะมีขนาด 512 [94] 2,048 หรือ 4,096 ไบต์ ที่เกี่ยวข้องกับแต่ละหน้าคือสองสามไบต์ (โดยทั่วไปคือ 1/32 ของขนาดข้อมูล) ที่สามารถใช้สำหรับการจัดเก็บการตรวจสอบรหัสแก้ไขข้อผิดพลาด (ECC)

ขนาดบล็อกทั่วไป ได้แก่ :

- 32 หน้า 512+16 ไบต์สำหรับขนาดบล็อก (มีผล) 16 KiB
- 64 หน้า 2,048+64 ไบต์ต่อบล็อกขนาด 128 KiB[95]
- 64 หน้า 4,096+128 ไบต์ต่อบล็อกขนาด 256 KiB[96]
- 128 หน้า 4,096+128 ไบต์ต่อบล็อกขนาด 512 KiB

ในขณะที่อ่านและเขียนโปรแกรมบนพื้นฐานหน้า การลบสามารถทำได้บนพื้นฐานบล็อกเท่านั้น

อุปกรณ์ NAND ยังต้องการการจัดการบล็อกที่ไม่ดีโดยซอฟต์แวร์ไดรเวอร์อุปกรณ์หรือโดยชิปควบคุมแยกต่างหาก ตัวอย่างเช่น การ์ด SD มีวงจรควบคุมเพื่อดำเนินการจัดการบล็อกที่ไม่ดีและปรับระดับการสึกหรอ เมื่อซอฟต์แวร์ระดับสูงเข้าถึงบล็อกเชิงตรรกะ

บล็อกนั้นจะถูกจับคู่กับบล็อกจริงโดยไดรเวอร์อุปกรณ์หรือตัวควบคุม อาจมีการแบ่งบล็อกจำนวนหนึ่งบนชิปแฟลชเพื่อจัดเก็บตารางการแมปเพื่อจัดการกับบล็อกที่เสียหาย หรือระบบอาจตรวจสอบแต่ละบล็อกเมื่อเปิดเครื่องเพื่อสร้างแผนที่บล็อกที่ไม่ดีใน RAM ความจุของหน่วยความจำโดยรวมจะค่อยๆ ลดขนาดลงเมื่อมีบล็อกจำนวนมากขึ้นถูกทำเครื่องหมายว่าไม่ดี

NAND อาศัย ECC เพื่อชดเชยบิตที่อาจล้มเหลวโดยธรรมชาติระหว่างการทำงานของอุปกรณ์ปกติ ECC ทั่วไปจะแก้ไขข้อผิดพลาดหนึ่งบิตในแต่ละ 2048 บิต (256 ไบต์) โดยใช้ ECC 22 บิต หรือข้อผิดพลาดหนึ่งบิตในแต่ละ 4096 บิต (512 ไบต์) โดยใช้ ECC 24 บิต[98] หาก ECC ไม่สามารถแก้ไขข้อผิดพลาดระหว่างการอ่านได้ ระบบอาจยังตรวจพบข้อผิดพลาด เมื่อทำการลบหรือการทำงานของโปรแกรม อุปกรณ์สามารถตรวจจับบล็อกที่ไม่สามารถตั้งโปรแกรมหรือลบและทำเครื่องหมายว่าบล็อกนั้นไม่ถูกต้อง ข้อมูลจะถูกเขียนไปยังบล็อกอื่นที่ดีและมีการอัปเดตแผนที่บล็อกที่ไม่ดี

รหัส Hamming เป็น ECC ที่ใช้บ่อยที่สุดสำหรับแฟลช SLC NAND รหัส Reed-Solomon และรหัส BCH (รหัส Bose-Chaudhuri-Hocquenghem) มักใช้ ECC สำหรับแฟลช MLC NAND ชิปแฟลช MLC NAND บางตัวสร้างรหัสแก้ไขข้อผิดพลาด BCH ที่เหมาะสมภายใน

อุปกรณ์ NAND ส่วนใหญ่จัดส่งมาจากโรงงานโดยมีบล็อกที่ไม่ดี โดยทั่วไปแล้วสิ่งเหล่านี้จะถูกทำเครื่องหมายตามกลยุทธ์การทำเครื่องหมายบล็อกที่ไม่ถูกต้องที่ระบุ โดยการอนุญาตบล็อกที่ไม่ดี ผู้ผลิตจะได้ผลตอบแทนที่สูงกว่าที่จะเป็นไปได้หากบล็อกทั้งหมดต้องได้รับการตรวจสอบว่าดี ซึ่งช่วยลดต้นทุนแฟลช NAND ได้อย่างมากและลดความจุในการจัดเก็บของชิ้นส่วนลงเพียงเล็กน้อยเท่านั้น

เมื่อเรียกใช้ซอฟต์แวร์จากหน่วยความจำ NAND มักใช้กลยุทธ์หน่วยความจำเสมือน: เนื้อหาหน่วยความจำจะต้องถูกเพจหรือคัดลอกไปยัง RAM ที่แมปหน่วยความจำก่อนและดำเนินการที่นั่น (นำไปสู่การรวมกันระหว่าง NAND + RAM) หน่วยการจัดการหน่วยความจำ (MMU) ในระบบมีประโยชน์ แต่ก็สามารถทำได้ด้วยการซ้อนทับ ด้วยเหตุนี้ บางระบบจะใช้หน่วยความจำ NOR และ NAND ร่วมกัน โดยที่หน่วยความจำ NOR ที่เล็กกว่าจะถูกใช้เพื่อเป็น ROM ของซอฟต์แวร์ และหน่วยความจำ NAND ที่ใหญ่กว่าจะถูกแบ่งพาร์ติชันด้วยระบบไฟล์เพื่อใช้เป็นพื้นที่จัดเก็บข้อมูลแบบไม่ลบเลือน

NAND สละข้อได้เปรียบในการเข้าถึงแบบสุ่มและดำเนินการแทน NOR NAND เหมาะที่สุดสำหรับระบบที่ต้องการการจัดเก็บข้อมูลที่มีความจุสูง มีความหนาแน่นสูงขึ้น ความจุที่มากขึ้น และต้นทุนที่ต่ำกว่า มีการลบที่เร็วขึ้น การเขียนตามลำดับ และการอ่านตามลำดับ

มาตรฐาน (Standardization)

กลุ่มที่เรียกว่า Open NAND Flash Interface Working Group (ONFI) ได้พัฒนา อินเทอร์เฟซระดับต่ำที่ได้มาตรฐานสำหรับชิปแฟลช NAND ซึ่งช่วยให้สามารถทำงานร่วมกันได้ ระหว่างอุปกรณ์ NAND ที่สอดคล้องจากผู้ขายรายต่างๆ ข้อมูลจำเพาะ ONFI เวอร์ชัน 1.0 เผยแพร่ เมื่อวันที่ 28 ธันวาคม พ.ศ. 2549 โดยระบุ:

- อินเทอร์เฟซทางกายภาพมาตรฐาน (pinout) สำหรับแฟลช NAND ในแพ็คเกจ TSOP-48, WSOP-48, LGA-52 และ BGA-63
- ชุดคำสั่งมาตรฐานสำหรับการอ่าน เขียน และลบชิปแฟลช NAND
- กลไกสำหรับการระบุตัวเอง (เทียบได้กับคุณสมบัติการตรวจจับการมีอยู่ของซีเรียลของโมดูล หน่วยความจำ SDRAM)

กลุ่ม ONFI ได้รับการสนับสนุนจากผู้ผลิตแฟลช NAND รายใหญ่ ซึ่งรวมถึง Hynix, Intel, Micron Technology และ Numonyx ตลอดจนผู้ผลิตรายใหญ่ของอุปกรณ์ที่ใช้ชิปแฟลช NAND

ผู้ผลิตอุปกรณ์แฟลชรายใหญ่สองรายคือ Toshiba และ Samsung เลือกใช้อินเทอร์เฟซของการออกแบบของตนเองที่เรียกว่า Toggle Mode (และปัจจุบันคือ Toggle V2.0) อินเทอร์เฟซนี้ไม่สามารถใช้งานร่วมกับข้อกำหนดของ ONFI ได้ ผลที่ได้คือผลิตภัณฑ์ที่ออกแบบมาสำหรับอุปกรณ์ของผู้ขายรายหนึ่งอาจไม่สามารถใช้อุปกรณ์ของผู้ขายรายอื่นได้

กลุ่มผู้จำหน่าย รวมทั้ง Intel, Dell และ Microsoft ได้ก่อตั้งคณะทำงาน Non-Volatile Memory Host Controller Interface (NVMHCI) เป้าหมายของกลุ่มนี้คือการจัดหาอินเทอร์เฟซการเขียนโปรแกรมซอฟต์แวร์และฮาร์ดแวร์มาตรฐานสำหรับระบบย่อยหน่วยความจำแบบไม่ลบเลือน ซึ่งรวมถึงอุปกรณ์ "แฟลชแคช" ที่เชื่อมต่อกับบัส PCI Express

ความแตกต่างระหว่าง NOR และ NAND Flash

แฟลช NOR และ NAND แตกต่างกันในสองวิธีที่สำคัญ:

- การเชื่อมต่อของเซลล์หน่วยความจำแต่ละเซลล์นั้นแตกต่างกัน
- อินเทอร์เฟซสำหรับอ่านและเขียนหน่วยความจำแตกต่างกัน NOR อนุญาตการเข้าถึงแบบสุ่ม ในขณะที่ NAND อนุญาตเฉพาะการเข้าถึงหน้าเท่านั้น

แฟลช NOR และ NAND ได้ชื่อมาจากโครงสร้างของการเชื่อมต่อระหว่างเซลล์หน่วยความจำ [ต้องการอ้างอิง] ในแฟลช NOR เซลล์จะเชื่อมต่อแบบขนานกับบรรทัดบิต ทำให้เซลล์สามารถอ่านและตั้งโปรแกรมแยกกันได้

การเชื่อมต่อแบบขนานของเซลล์คล้ายกับการเชื่อมต่อแบบขนานของทรานซิสเตอร์ในเกต CMOS NOR ในแฟลช NAND เซลล์จะเชื่อมต่อแบบอนุกรม คล้ายกับเกต CMOS NAND การเชื่อมต่อแบบอนุกรมใช้พื้นที่น้อยกว่าการเชื่อมต่อแบบขนาน ซึ่งช่วยลดต้นทุนของแฟลช NAND โดยตัวมันเองไม่ได้ป้องกันไม่ให้เซลล์ NAND ถูกอ่านและตั้งโปรแกรมทีละรายการ

เซลล์แฟลช NOR แต่ละเซลล์มีขนาดใหญ่กว่าเซลล์แฟลช NAND – 10 F2 เทียบกับ 4 F2 – แม้ว่าจะใช้การผลิตอุปกรณ์เซมิคอนดักเตอร์เหมือนกันทุกประการ ดังนั้นทรานซิสเตอร์ หน้าสัมผัส ฯลฯ แต่ละตัวจึงมีขนาดเท่ากันทุกประการ เนื่องจากเซลล์แฟลช NOR ต้องใช้โลหะแยกต่างหากติดต่อแต่ละเซลล์

เนื่องจากการเชื่อมต่อแบบอนุกรมและการลบผู้ติดต่อ wordline กริดขนาดใหญ่ของเซลล์หน่วยความจำแฟลช NAND อาจใช้พื้นที่เพียง 60% ของพื้นที่เซลล์ NOR ที่เทียบเท่ากัน[105] (สมมติว่ามีความละเอียดของกระบวนการ CMOS เท่ากัน เช่น 130 นาโนเมตร, 90 นาโนเมตร หรือ 65 นาโนเมตร) นักออกแบบของแฟลช NAND ตระหนักดีว่าพื้นที่ของชิป NAND และด้วยเหตุนี้ต้นทุนจึงสามารถลดลงได้อีกโดยการถอดที่อยู่ภายนอกและวงจรบัสข้อมูลออก

แต่อุปกรณ์ภายนอกสามารถสื่อสารกับแฟลช NAND ผ่านคำสั่งที่เข้าถึงตามลำดับและการลงทะเบียนข้อมูล ซึ่งจะดึงและส่งออกข้อมูลที่จำเป็นภายใน ตัวเลือกการออกแบบนี้ทำให้การเข้าถึงหน่วยความจำแฟลช NAND แบบสุ่มเป็นไปได้ แต่เป้าหมายของแฟลช NAND คือการเปลี่ยนฮาร์ดดิสก์แบบกลไก ไม่ใช่เพื่อแทนที่ ROM

Attribute	NAND	NOR
Main application	File storage	Code execution
Storage capacity	High	Low
Cost per bit	Low	
Active power	Low	
Standby power		Low
Write speed	Fast	
Read speed		Fast
Execute in place (XIP)	No	Yes
Reliability		High

ระบบไฟล์แฟลช

เนื่องจากคุณสมบัติเฉพาะของหน่วยความจำแฟลช จึงเหมาะที่สุดที่จะใช้กับคอนโทรลเลอร์เพื่อปรับระดับการสึกหรอและการแก้ไขข้อผิดพลาด หรือระบบไฟล์แฟลชที่ออกแบบมาโดยเฉพาะ ซึ่งจะกระจายการเขียนบนสื่อและจัดการกับเวลาการลบที่ยาวนานของบล็อกแฟลช NOR แนวคิดพื้นฐานเบื้องหลังระบบไฟล์แฟลชมีดังต่อไปนี้: เมื่อมีการอัปเดตที่เก็บแฟลช ระบบไฟล์จะเขียนสำเนาใหม่ของข้อมูลที่เปลี่ยนแปลงไปยังบล็อกใหม่ ทำการแมปตัวชี้ไฟล์ใหม่ จากนั้นจึงลบบล็อกเก่าในภายหลัง มีเวลา

ในทางปฏิบัติ ระบบไฟล์แฟลชจะใช้เฉพาะกับอุปกรณ์เทคโนโลยีหน่วยความจำ (MTD) ซึ่งเป็นหน่วยความจำแฟลชแบบฝังที่ไม่มีตัวควบคุม การดหน่วยความจำแฟลชแบบถอดได้, SSD, ชิพ eMMC/eUFS และแฟลชไดรฟ์ USB มีตัวควบคุมในตัวเพื่อทำการปรับระดับการสึกหรอและการแก้ไขข้อผิดพลาด ดังนั้นการใช้ระบบไฟล์แฟลชเฉพาะจึงไม่มีประโยชน์ใดๆ

ความจุ

ชิปหลายตัวมักจะถูกจัดเรียงหรือเรียงซ้อนกันเพื่อให้ได้ความจุที่สูงขึ้น สำหรับใช้ในอุปกรณ์อิเล็กทรอนิกส์สำหรับผู้บริโภค เช่น เครื่องเล่นมัลติมีเดียหรือ GPS การปรับขนาดความจุ (เพิ่มขึ้น) ของชิปแฟลชใช้ตามกฎหมายของมัวร์ เนื่องจากผลิตขึ้นด้วยเทคนิคและอุปกรณ์วงจรรวมหลายแบบที่เหมือนกัน นับตั้งแต่มีการนำ 3D NAND มาใช้ การปรับขนาดไม่จำเป็นต้องเกี่ยวข้องกับกฎของมัวร์อีกต่อไป เนื่องจากไม่มีการใช้ทรานซิสเตอร์ (เซลล์) ที่มีขนาดเล็กกว่าอีกต่อไป

โดยทั่วไปแล้วอุปกรณ์จัดเก็บข้อมูลแฟลชสำหรับผู้บริโภคจะได้รับการโฆษณาด้วยขนาดที่ใช้งานได้ซึ่งแสดงเป็นจำนวนเต็มขนาดเล็กสอง (2, 4, 8 เป็นต้น) และการกำหนดเมกะไบต์ (MB) หรือ กิกะไบต์ (GB) เช่น 512 MB, 8 GB. ซึ่งรวมถึง SSD ที่วางตลาดเป็นฮาร์ดไดรฟ์ทดแทน ตามฮาร์ดไดรฟ์แบบเดิมซึ่งใช้เลขนำหน้าทศนิยม[139] ดังนั้น SSD ที่มีเครื่องหมาย "64 GB" จะต้องมีความจุอย่างน้อย 64×10003 ไบต์ (64 GB) ผู้ใช้ส่วนใหญ่จะมีความจุน้อยกว่านี้เล็กน้อยสำหรับไฟล์ของตน เนื่องจากพื้นที่ที่ใช้โดยข้อมูลเมตาของระบบไฟล์

ชิปหน่วยความจำแฟลชภายในมีขนาดเป็นทวีคูณไบนารีที่เข้มงวด แต่ความจุรวมที่แท้จริงของชิปไม่สามารถใช้ได้ทีอินเทอร์เฟซของไดรฟ์ ซึ่งมีขนาดใหญ่กว่าความจุที่โฆษณาไว้มาก เพื่อให้สามารถกระจายการเขียน (การปรับระดับการสึกหรอ) เพื่อการประหยัด สำหรับรหัสการแก้ไขข้อผิดพลาด และสำหรับข้อมูลเมตาอื่นๆ ที่จำเป็นสำหรับเฟิร์มแวร์ภายในของอุปกรณ์

ในปี 2548 โตชิบาและแซนด์สกีได้พัฒนาชิปแฟลช NAND ที่สามารถจัดเก็บข้อมูลขนาด 1 GB โดยใช้เทคโนโลยีเซลล์หลายระดับ (MLC) ซึ่งสามารถจัดเก็บข้อมูลได้สองบิตต่อเซลล์ ในเดือนกันยายน 2548 Samsung Electronics ประกาศว่าได้พัฒนาชิป 2 GB ตัวแรกของโลก

ในเดือนมีนาคม พ.ศ. 2549 ซัมซุงได้ประกาศเปิดตัวแฟลชไดรฟ์ที่มีความจุ 4 GB ซึ่งโดยทั่วไปแล้วจะมีลำดับความสำคัญเท่ากับฮาร์ดไดรฟ์แล็ปท็อปที่มีขนาดเล็กกว่า และในเดือนกันยายน พ.ศ. 2549 ซัมซุงได้ประกาศชิปขนาด 8 GB ที่ผลิตโดยใช้กระบวนการผลิตขนาด 40 นาโนเมตร[141] ในเดือนมกราคม 2008 SanDisk ได้ประกาศความพร้อมใช้งานของการ์ด MicroSDHC ขนาด 16 GB และ SDHC Plus ขนาด 32 GB แฟลชไดรฟ์รุ่นล่าสุด (ณ ปี 2555) มีความจุมากกว่ามาก โดยจุได้ 64, 128 และ 256 GB การพัฒนาร่วมกันที่ Intel และ Micron จะช่วยให้สามารถผลิตแฟลชสติก NAND ขนาด 32 เลเยอร์ 3.5 เทราไบต์ (ต้องการการชาร์จแรง) ได้ 32 เลเยอร์ และ SSD ขนาดมาตรฐาน 10 TB อุปกรณ์ประกอบด้วย 5 แพ็คเกจขนาด 16 × 48 GB TLC dies โดยใช้การออกแบบเซลล์เกทแบบลอย

ชิปแฟลชยังคงผลิตต่อไปโดยมีความจุต่ำกว่าหรือประมาณ 1 MB (เช่น สำหรับ BIOS-ROM และแอปพลิเคชันแบบฝัง)

ในเดือนกรกฎาคม พ.ศ. 2559 ซัมซุงได้ประกาศเปิดตัว Samsung 850 EVO ขนาด 4 TB ซึ่งใช้ TLC 3D V-NAND ขนาด 256 Gbit 48 เลเยอร์[146] ในเดือนสิงหาคม 2559 Samsung ได้ประกาศเปิดตัว SAS SSD ขนาด 32 TB 2.5 นิ้ว โดยใช้ TLC 3D V-NAND 512 Gbit 64-layer นอกจากนี้ Samsung คาดว่าจะเปิดตัว SSD ที่มีพื้นที่จัดเก็บสูงสุด 100 TB ภายในปี 2020

อัตราการโอน

อุปกรณ์หน่วยความจำแฟลชมักจะอ่านได้เร็วกว่าการเขียนมากประสิทธิภาพยังขึ้นอยู่กับคุณภาพของตัวควบคุมการจัดเก็บข้อมูลซึ่งมีความสำคัญมากขึ้นเมื่ออุปกรณ์บางส่วนเต็ม แม้ว่าการเปลี่ยนแปลงในการผลิตเพียงอย่างเดียวคือการหดตัว การไม่มีตัวควบคุมที่เหมาะสมอาจส่งผลให้ความเร็วลดลง

แอปพลิเคชัน

แฟลชอนุกรม (Serial flash)

แฟลชซีเรียลเป็นหน่วยความจำแฟลชขนาดเล็กที่ใช้พลังงานต่ำซึ่งให้การเข้าถึงข้อมูลแบบอนุกรมเท่านั้น แทนที่จะระบุไบต์เดียว ผู้ใช้จะอ่านหรือเขียนกลุ่มไบต์ที่ต่อเนื่องกันขนาดใหญ่ในพื้นที่ที่อยู่แบบอนุกรม Serial Peripheral Interface Bus (SPI) เป็นโปรโตคอลทั่วไปสำหรับการเข้าถึงอุปกรณ์

เมื่อรวมเข้ากับระบบฝังตัว แฟลชซีเรียลต้องใช้สายไฟบน PCB น้อยกว่าหน่วยความจำแฟลชแบบขนาน เนื่องจากจะส่งและรับข้อมูลที่ละบิต ซึ่งอาจช่วยลดพื้นที่บอร์ด การใช้พลังงาน และต้นทุนรวมของระบบ

มีสาเหตุหลายประการที่ว่าทำไมอุปกรณ์อนุกรมที่มีพินภายนอกน้อยกว่าอุปกรณ์ขนานจึงสามารถลดต้นทุนโดยรวมได้อย่างมาก:

- ASIC จำนวนมากจำกัดเฉพาะแผ่น ซึ่งหมายความว่าขนาดของแม่พิมพ์ถูกจำกัดด้วยจำนวนของแผ่นลวดเชื่อม มากกว่าความซับซ้อนและจำนวนประตูที่ใช้สำหรับตรรกะของอุปกรณ์ การจำกัดแผ่นยึดเกาะทำให้วงจรรวมมีขนาดเล็กลง นี่จะเพิ่มจำนวนของแม่พิมพ์ที่อาจประดิษฐ์บนแผ่นเวเฟอร์ และด้วยเหตุนี้จึงลดต้นทุนต่อตาย
- การลดจำนวนพินภายนอกยังช่วยลดต้นทุนการประกอบและบรรจุภัณฑ์อีกด้วย อุปกรณ์อนุกรมอาจบรรจุในแพ็คเกจที่เล็กกว่าและง่ายกว่าอุปกรณ์แบบขนาน
- แพ็คเกจจำนวนพินที่เล็กกว่าและน้อยกว่าใช้พื้นที่ PCB น้อยลง
- อุปกรณ์จำนวนพินที่ต่ำกว่าทำให้การกำหนดเส้นทาง PCB ง่ายขึ้น

แฟลช SPI หลักๆ มีอยู่สองประเภท ประเภทแรกมีลักษณะเป็นหน้าขนาดเล็กและบัฟเฟอร์ หน้า SRAM ภายในอย่างน้อยหนึ่งหน้าทำให้สามารถอ่านหน้าที่สมบูรณ์ไปยังบัฟเฟอร์ แก้ไขบางส่วน แล้วเขียนกลับ (เช่น Atmel AT45 DataFlash หรือ Micron Technology Page Erase NOR Flash).

ประเภทที่สองมีเซลล์ที่ใหญ่กว่า โดยที่เซลล์ที่เล็กที่สุดที่มักพบในแฟลช SPI ประเภทนี้คือ 4 kB แต่อาจมีขนาดใหญ่ถึง 64 kB เนื่องจากแฟลช SPI ประเภทนี้ไม่มีบัฟเฟอร์ SRAM ภายใน จึงต้องอ่านและแก้ไขหน้าทั้งหมดก่อนที่จะเขียนกลับ ทำให้จัดการได้ช้า อย่างไรก็ตาม ประเภทที่สองมีราคาถูกกว่าประเภทแรก ดังนั้น จึงเป็นตัวเลือกที่ดีเมื่อแอปพลิเคชันเป็นการสร้างโค้ดแฮดแวร์ ทั้งสองประเภทไม่สามารถแลกเปลี่ยนได้ง่าย เนื่องจากไม่มีพินเอาต์เดียวกัน และชุดคำสั่งไม่เข้ากัน

FPGA ส่วนใหญ่ใช้เซลล์การกำหนดค่า SRAM และต้องใช้อุปกรณ์กำหนดค่าภายนอก ซึ่งมักจะเป็นชิปแฟลชอนุกรม เพื่อโหลดการกำหนดค่าบิตสตรีมเข้าสู่รอบการจ่ายไฟ

การจัดเก็บเฟิร์มแวร์ (Firmware storage)

ด้วยความเร็วที่เพิ่มขึ้นของซีพียูสมัยใหม่ อุปกรณ์แฟลชแบบขนานมักจะช้ากว่าบัส หน่วยความจำของคอมพิวเตอร์ที่เชื่อมต่ออยู่มาก ในทางกลับกัน SRAM ที่ทันสมัยใช้เวลาในการเข้าถึงต่ำกว่า 10 ns ในขณะที่ DDR2 SDRAM ใช้เวลาในการเข้าถึงต่ำกว่า 20 ns ด้วยเหตุนี้จึงมักเป็นที่ต้องการเพื่อให้โค้ดที่จัดเก็บไว้ในแฟลชลงในแรม นั่นคือรหัสจะถูกคัดลอกจากแฟลชไปยัง RAM ก่อนดำเนินการเพื่อให้ CPU สามารถเข้าถึงได้ด้วยความเร็วเต็มที่ เฟิร์มแวร์ของอุปกรณ์อาจถูกจัดเก็บไว้ในอุปกรณ์แฟลชซีเรียล จากนั้นจึงคัดลอกไปยัง SDRAM หรือ SRAM เมื่อเปิดเครื่อง

การใช้อุปกรณ์แฟลชอนุกรมภายนอกแทนการใช้แฟลชบนชิปช่วยลดความจำเป็นในกระบวนการประนีประนอม (กระบวนการผลิตที่ดีสำหรับลอจิกความเร็วสูงมักไม่ดีสำหรับแฟลชและในทางกลับกัน)

เมื่อตัดสินใจอ่านเฟิร์มแวร์ในบล็อกขนาดใหญ่หนึ่งบล็อก เป็นเรื่องปกติที่จะเพิ่มการบีบอัดเพื่อให้สามารถใช้ชิปแฟลชขนาดเล็กลงได้ แอปพลิเคชันทั่วไปสำหรับแฟลชซีเรียลรวมถึงการจัดเก็บเฟิร์มแวร์สำหรับฮาร์ดไดรฟ์ ตัวควบคุมอีเทอร์เน็ต โมเด็ม DSL อุปกรณ์เครือข่ายไร้สาย ฯลฯ

หน่วยความจำแฟลชแทนฮาร์ดไดรฟ์ (Flash memory as a replacement for hard drives)

แอปพลิเคชันล่าสุดสำหรับหน่วยความจำแฟลชใช้แทนฮาร์ดดิสก์ หน่วยความจำแฟลชไม่มีข้อจำกัดทางกลไกและเวลาแฝงของฮาร์ดไดรฟ์ ดังนั้นโซลิดสเตตไดรฟ์ (SSD) จึงน่าสนใจเมื่อพิจารณาถึงความเร็ว เสียง การใช้พลังงาน และความน่าเชื่อถือ

แฟลชไดรฟ์กำลังได้รับแรงฉุดในฐานะอุปกรณ์จัดเก็บข้อมูลสำรองของอุปกรณ์พกพา นอกจากนี้ยังใช้แทนฮาร์ดไดรฟ์ในคอมพิวเตอร์เดสก์ท็อปประสิทธิภาพสูงและเซิร์ฟเวอร์บางตัวที่มีสถาปัตยกรรม RAID และ SAN

ยังมีบางแง่มุมของ SSD แบบแฟลชที่ทำให้ไม่สวย ราคาต่อกิกะไบต์ของหน่วยความจำแฟลชยังคงสูงกว่าฮาร์ดดิสก์อย่างมาก

หน่วยความจำแฟลชยังมีรอบ P/E (โปรแกรม/การลบ) ในจำนวนที่จำกัด แต่ดูเหมือนว่าจะอยู่ภายใต้การควบคุมในขณะนี้ เนื่องจากการรับประกันสำหรับ SSD ที่ใช้แฟลชนั้นใกล้จะถึงแล้วสำหรับฮาร์ดไดรฟ์ปัจจุบัน

นอกจากนี้ ไฟล์ที่ถูกลบบน SSD สามารถคงอยู่ได้เป็นระยะเวลาไม่จำกัดก่อนที่จะถูกเขียนทับด้วยข้อมูลใหม่ เทคนิคการลบหรือทำลายหรือซอฟต์แวร์ที่ทำงานได้ดีกับฮาร์ดดิสก์ไดรฟ์แบบแม่เหล็กไม่มีผลกระทบต่อ SSD ซึ่งส่งผลต่อความปลอดภัยและการตรวจสอบทางนิติเวช อย่างไรก็ตาม เนื่องจากคำสั่ง TRIM ที่เรียกใช้โดยโซลิดสเตตไดรฟ์ส่วนใหญ่ ซึ่งทำเครื่องหมายที่อยู่บล็อกแบบลอจิกที่ถูกครอบครองโดยไฟล์ที่ถูกลบว่าไม่ได้ใช้เพื่อเปิดใช้งานการรวบรวมขยะ ซอฟต์แวร์กู้คืนข้อมูลจึงไม่สามารถกู้คืนไฟล์ที่ถูกลบจากไฟล์ดังกล่าวได้

สำหรับฐานข้อมูลเชิงสัมพันธ์หรือระบบอื่น ๆ ที่ต้องใช้ธุรกรรม ACID แม้แต่พื้นที่เก็บข้อมูลแฟลชจำนวนเล็กน้อยก็สามารถเพิ่มความเร็วให้กับอาร์เรย์ของดิสก์ไดรฟ์ได้อย่างมาก

ในเดือนพฤษภาคม 2549 Samsung Electronics ได้ประกาศพิธีที่ใช้หน่วยความจำแฟลชสองเครื่อง นั่นคือ Q1-SSD และ Q30-SSD ที่คาดว่าจะวางจำหน่ายในเดือนมิถุนายน 2549 ซึ่งทั้งสองเครื่องใช้ SSD ขนาด 32 GB และอย่างน้อยในตอนแรกเริ่มมีเฉพาะในเกาหลีใต้เท่านั้น การเปิดตัว Q1-SSD และ Q30-SSD นั้นล่าช้าและในที่สุดก็ถูกจัดส่งในปลายเดือนสิงหาคม 2549

พีซีที่ใช้หน่วยความจำแฟลชเครื่องแรกที่วางจำหน่ายคือ Sony Vaio UX90 ซึ่งประกาศให้ส่งจองล่วงหน้าเมื่อวันที่ 27 มิถุนายน พ.ศ. 2549 และเริ่มจัดส่งในญี่ปุ่นเมื่อวันที่ 3 กรกฎาคม พ.ศ. 2549 ด้วยฮาร์ดไดรฟ์หน่วยความจำแฟลชขนาด 16GB ปลายเดือนกันยายน 2549 Sony ได้อัพเกรดหน่วยความจำแฟลชใน Vaio UX90 เป็น 32Gb

ไดรฟ์โซลิดสเตตเป็นตัวเลือกสำหรับ MacBook Air เครื่องแรกที่เปิดตัวในปี 2008 และตั้งแต่ปี 2010 เป็นต้นไป ทุกรุ่นจะมาพร้อมกับ SSD เริ่มตั้งแต่ปลายปี 2011 ซึ่งเป็นส่วนหนึ่งของโครงการริเริ่ม Ultrabook ของ Intel แล็บท็อปบางเฉียบจำนวนมากขึ้นมาพร้อมกับมาตรฐาน SSD

นอกจากนี้ยังมีเทคนิคไฮบริด เช่น ไฮบริดไดรฟ์และ ReadyBoost ที่พยายามรวมข้อดีของเทคโนโลยีทั้งสองเข้าด้วยกัน โดยใช้แฟลชเป็นแคชความเร็วสูงแบบไม่ลบเลือนสำหรับไฟล์บนดิสก์ที่มักถูกอ้างอิง แต่ไม่ค่อยได้รับการแก้ไข เช่น แอปพลิเคชันและ ไฟล์ปฏิบัติการระบบปฏิบัติการ

ที่เก็บถาวรหรือการจัดเก็บระยะยาว (Archival or long-term storage)

ทรานซิสเตอร์แบบ Floating-gate ในอุปกรณ์เก็บข้อมูลแฟลชจะเก็บประจุซึ่งแสดงถึงข้อมูล ประจุนี้จะค่อยๆ รั่วไหลไปตามกาลเวลา ทำให้เกิดข้อผิดพลาดทางตรรกะสะสม หรือที่เรียกว่า "บิตเน่า" หรือ "จางลงเล็กน้อย"

การเก็บรักษาข้อมูล (Data retention)

ยังไม่ชัดเจนว่าข้อมูลในหน่วยความจำแฟลชจะคงอยู่ได้นานเพียงใดภายใต้สภาวะการเก็บถาวร (เช่น อุณหภูมิที่ไม่เป็นอันตรายและความชื้นที่มีการเข้าถึงไม่บ่อยนักโดยมีหรือไม่มีการเขียนซ้ำเพื่อป้องกัน) เอกสารข้อมูลของไมโครคอนโทรลเลอร์ "ATmega" ที่ใช้แฟลชของ Atmel มักรับประกันเวลาเก็บรักษา 20 ปีที่ 85 °C (185 °F) และ 100 ปีที่ 25 °C (77 °F)

ช่วงการเก็บข้อมูลจะแตกต่างกันไปตามประเภทและรุ่นของที่เก็บข้อมูลแฟลช เมื่อจ่ายไฟและไม่ได้ใช้งาน เฟิร์มแวร์ของแฟลชต่อเราจะมีเฟรชประจุของทรานซิสเตอร์ที่เก็บข้อมูลเป็นประจำ

ความสามารถในการเก็บข้อมูลจะแตกต่างกันไปตามอุปกรณ์จัดเก็บข้อมูลแฟลช เนื่องจากความแตกต่างของเฟิร์มแวร์ ความซับซ้อนของข้อมูล และอัลกอริธึมการแก้ไขข้อผิดพลาด

บทความจาก CMU ในปี 2558 ระบุว่า "อุปกรณ์แฟลชของวันนี้ซึ่งไม่ต้องการการรีเฟรชแฟลช มีอายุการเก็บรักษาโดยทั่วไปคือ 1 ปีที่อุณหภูมิห้อง" และเวลาเก็บรักษานั้นจะลดลงอย่างทวีคูณเมื่ออุณหภูมิเพิ่มขึ้น ปรากฏการณ์นี้สามารถจำลองได้จากสมการอาร์เรเนียส

การกำหนดค่า FPGA (FPGA configuration)

FPGA บางตัวใช้เซลล์กำหนดค่าแฟลชที่ใช้โดยตรงเป็นสวิตช์ (ตั้งโปรแกรมได้) เพื่อเชื่อมต่อองค์ประกอบภายในเข้าด้วยกัน โดยใช้ทรานซิสเตอร์แบบ Floating-gate ชนิดเดียวกับเซลล์จัดเก็บข้อมูลแฟลชในอุปกรณ์จัดเก็บข้อมูล

อุตสาหกรรม

แหล่งข่าวรายหนึ่งระบุว่าในปี 2551 อุตสาหกรรมหน่วยความจำแฟลชมีมูลค่าการผลิตและการขายประมาณ 9.1 พันล้านดอลลาร์สหรัฐ แหล่งข้อมูลอื่นระบุว่าตลาดหน่วยความจำแฟลชมีมูลค่ามากกว่า 20 พันล้านดอลลาร์ในปี 2549 คิดเป็นสัดส่วนมากกว่าแปดเปอร์เซ็นต์ของตลาดเซมิคอนดักเตอร์โดยรวม และมากกว่า 34 เปอร์เซ็นต์ของตลาดหน่วยความจำเซมิคอนดักเตอร์ทั้งหมด ในปี 2555 ตลาดมีมูลค่าประมาณ 26.8 พันล้านดอลลาร์ อาจใช้เวลาถึง 10 สัปดาห์ในการผลิตชิปหน่วยความจำแฟลช

ความสามารถในการปรับขนาดแฟลช (Flash scalability)

เนื่องจากโครงสร้างที่ค่อนข้างเรียบง่ายและมีความต้องการความจุสูง หน่วยความจำแฟลช NAND จึงเป็นเทคโนโลยีที่มีการปรับขนาดอย่างเข้มงวดที่สุดในบรรดาอุปกรณ์อิเล็กทรอนิกส์ การแข่งขันที่รุนแรงระหว่างผู้ผลิตเพียงไม่กี่รายอันดับต้นๆ นั้น เป็นการเพิ่มความก้าวร้าวในการย่อขนาดกฎการออกแบบ MOSFET แบบ Floating-gate หรือโหนดเทคโนโลยีกระบวนการ[86] แม้ว่าแนวโน้มการย่อขนาดที่คาดหวังจะเป็นปัจจัยสองทุกๆ สามปีต่อฉบับดั้งเดิมของกฎของมัวร์ แต่เมื่อไม่นานมานี้สิ่งนี้ได้รับการเร่งให้เร็วขึ้นในกรณีของ NAND flash เป็นสองเท่าทุกๆ สองปี

อ้างอิง

Anton Shilov. (5 ธันวาคม 2560). Samsung Starts Production of 512 GB UFS NAND Flash Memory: 64-Layer V-NAND, 860 MB/s Reads. สืบค้นเมื่อวันที่ 11 ตุลาคม 2564. จากเว็บไซต์: <https://www.anandtech.com/show/12120/samsung-starts-production-of-512-gb-ufs-chips>

Avinash Aravindan. (23 กรกฎาคม 2561). Flash 101: NAND Flash vs NOR Flash. สืบค้นเมื่อวันที่ 9 ตุลาคม 2564. จากเว็บไซต์ <https://www.embedded.com/flash-101-nand-flash-vs-nor-flash/>

economist. (11 มีนาคม 2549). Not just a flash in the pan. สืบค้นเมื่อวันที่ 8 ตุลาคม 2564. จากเว็บไซต์: <https://www.economist.com/technology-quarterly/2006/03/11/not-just-a-flash-in-the-pan>

Matt Basinger. (15 กันยายน 2551). Flash . สืบค้นเมื่อวันที่ 8 ตุลาคม 2564 จากเว็บไซต์: https://web.archive.org/web/20091031121330/http://www.psocdeveloper.com/uploads/tx_piapappnote/an2209_03.pdf

Open NAND Flash Interface Specification. (28 ธันวาคม 2551). Open NAND Flash Interface. สืบค้นเมื่อวันที่ 12 ตุลาคม 2564. จากเว็บไซต์: https://web.archive.org/web/20110727145313/http://onfi.org/wp-content/uploads/2009/02/onfi_1_0_gold.pdf

Wikipedia. (21 กันยายน 2564). Distinction between NOR and NAND flash. สืบค้นเมื่อวันที่ 8 ตุลาคม 2564. จากเว็บไซต์: https://en.wikipedia.org/wiki/Flash_memory#Distinction_between_NOR_and_NAND_flash

Wikipedia. (21 กันยายน 2564). Flash memory. สืบค้นเมื่อวันที่ 8 ตุลาคม 2564. จากเว็บไซต์:
https://en.wikipedia.org/wiki/Flash_memory

Wikipedia. (21 กันยายน 2564). History. สืบค้นเมื่อวันที่ 12 ตุลาคม 2564. จากเว็บไซต์:
https://en.wikipedia.org/wiki/Flash_memory#History

Wikipedia. (21 กันยายน 2564). Limitations. สืบค้นเมื่อวันที่ 12 ตุลาคม 2564. จากเว็บไซต์:
https://en.wikipedia.org/wiki/Flash_memory#Limitations

Wikipedia. (21 กันยายน 2564). Low-level access. สืบค้นเมื่อวันที่ 12 ตุลาคม 2564. จาก
เว็บไซต์: https://en.wikipedia.org/wiki/Flash_memory#Low-level_access

Wikipedia. (21 กันยายน 2564). Principles of operation สืบค้นเมื่อวันที่ 12 ตุลาคม 2564. จาก
เว็บไซต์: https://en.wikipedia.org/wiki/Flash_memory#Principles_of_operation

Wikipedia. (16 มกราคม 2564). Flash file system. สืบค้นเมื่อวันที่ 9 ตุลาคม 2564. จากเว็บไซต์:
https://en.wikipedia.org/wiki/Flash_file_system