บทที่ 3 หน่วยประมวลผลกลางและหน่วยความจำหลัก

CT1701 ระบบคอมพิวเตอร์

Microprocessor

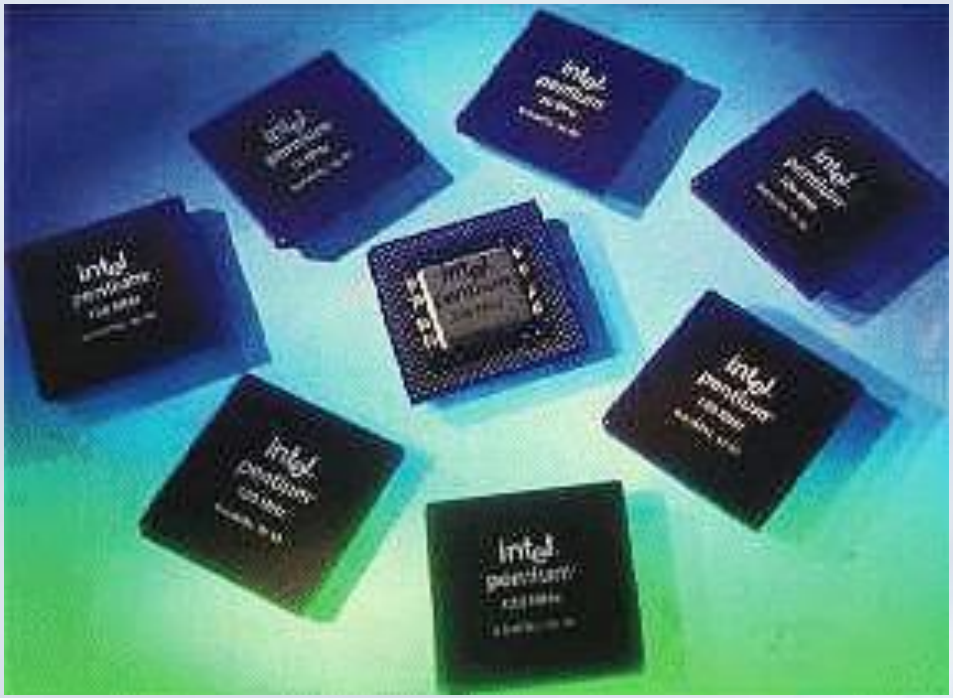
CPU (Central Processing Unit)

Microprocessor / CPU

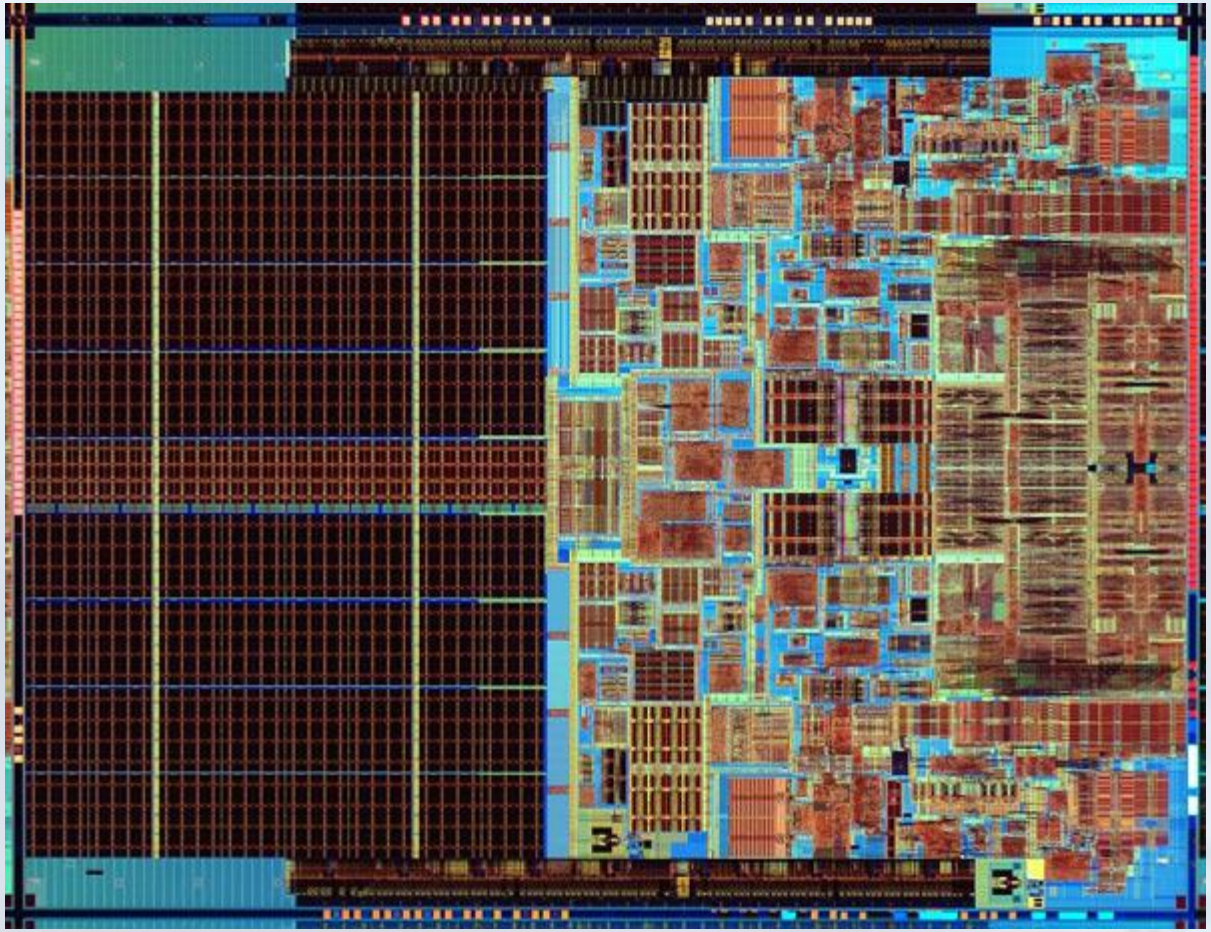
หน่วยประมวลผลกลาง (Central Processing Unit : CPU)

หน่วยประมวลผลกลางหรือซีพียู เรียกอีกชื่อหนึ่งว่า โปรเซสเซอร์ (Processor) หรือ ชิป (chip) นับเป็นอุปกรณ์ ที่มีความสำคัญมากที่สุด ของฮาร์ดแวร์เพราะมีหน้าที่ในการประมวลผลข้อมูลที่ผู้ใช้ป้อน เข้ามาทางอุปกรณ์อินพุต ตามชุดคำสั่งหรือโปรแกรมที่ผู้ใช้ต้องการใช้งาน หน่วยประมวลผลกลาง ประกอบด้วยส่วนประสำคัญ 3 ส่วน คือ ALU , Control Unit , Memory Unit

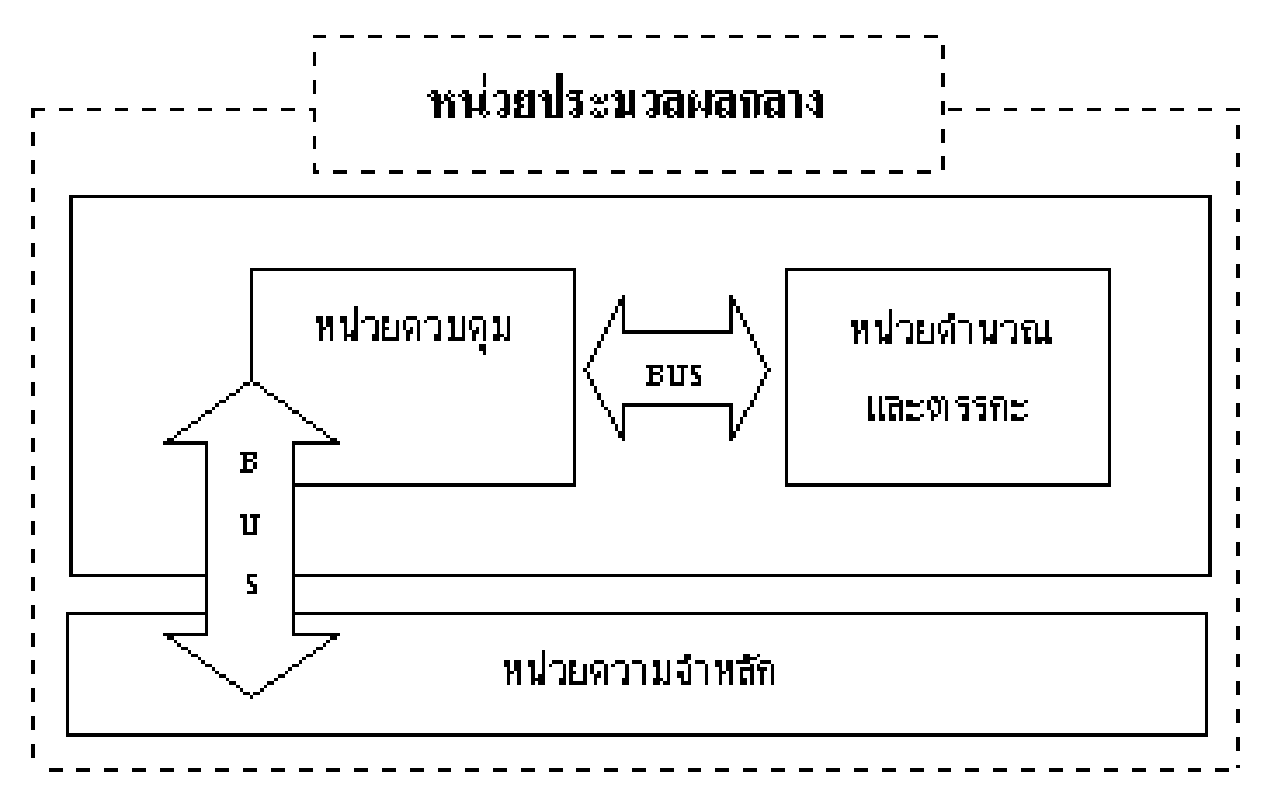
Microprocessor / CPU



Inside a Microprocessor



Inside a Microprocessor



Inside a Microprocessor

1. หน่วยคำนวณและตรรกะ (Arithmetic & Logical Unit : ALU)

หน่วยคำนวณตรรกะทำหน้าที่เหมือนกับเครื่องคำนวณอยู่ในเครื่องคอมพิวเตอร์โดยทำงานเกี่ยวข้องกับ การคำนวณทางคณิตศาสตร์ เช่น บวก ลบ คูณ หาร นอกจากนี้หน่วยคำนวณและตรรกะของคอมพิวเตอร์ ยังมีความสามารถอีกอย่างหนึ่งที่เครื่องคำนวณธรรมดาไม่มี คือ ความสามารถในเชิงตรรกะศาสตร์ หมายถึง ความสามารถในการเปรียบเทียบตามเงื่อนไข และกฎเกณฑ์ทางคณิตศาสตร์ เพื่อให้ได้คำตอบออกมาว่าเงื่อนไข นั้นเป็น จริง หรือ เท็จ เช่น เปรียบเทียบมากว่า น้อยกว่า เท่ากัน ไม่เท่ากัน ของจำนวน 2 จำนวน เป็นต้น ซึ่งการเปรียบเทียบนี้มักจะใช้ในการเลือกทำงานของเครื่องคอมพิวเตอร์จะทำตามคำสั่งใดของโปรแกรมเป็นคำสั่งต่อไป

Inside a Microprocessor

2. หน่วยควบคุม (Control Unit)

หน่วยควบคุมทำหน้าที่ควบคุมลำดับขั้นตอนการการประมวลผล และการทำงานของอุปกรณ์ต่างๆ ภายใน หน่วยประมวลผลกลาง และรวมไปถึงการประสานงานในการทำงานร่วมกันระหว่างหน่วยประมวลผลกลาง กับอุปกรณ์นำเข้าข้อมูล อุปกรณ์แสดงผล และหน่วยความจำสำรองด้วย เมื่อผู้ใช้ต้องการประมวลผล ตามชุดคำสั่งใด ผู้ใช้จะต้องส่งข้อมูลและชุดคำสั่งนั้น ๆ เข้าสู่ระบบ คอมพิวเตอร์เสียก่อน โดยข้อมูล และชุดคำสั่งดังกล่าวจะถูกนำไปเก็บไว้ในหน่วยความจำหลักก่อน จากนั้นหน่วยควบคุมจะดึงคำสั่งจาก ชุดคำสั่งที่มีอยู่ในหน่วยความจำหลักออกมาทีละคำสั่งเพื่อทำการแปล ความหมายว่าคำสั่งดังกล่าวสั่งให้ ฮาร์ดแวร์ส่วนใดทำงานอะไรกับข้อมูลตัวใด

Inside a Microprocessor

เมื่อทราบความหมายของ คำสั่งนั้นแล้ว หน่วยควบคุมก็จะส่งสัญญาณคำสั่งไปยังฮาร์แวร์ ส่วนที่ทำหน้าที่ ในการประมวลผลดังกล่าวให้ทำตามคำสั่งนั้น ๆ เช่น ถ้าคำสั่ง ที่เข้ามานั้นเป็นคำสั่งเกี่ยวกับการคำนวณ หน่วยควบคุมจะส่งสัญญาณ คำสั่งไปยังหน่วยคำนวณและตรรกะให้ทำงาน หน่วยคำนวณและตรรกะก็จะไปทำการดึงข้อมูลจากหน่วยความจำหลัก เข้ามาประมวลผล ตามคำสั่งแล้วนำผลลัพธ์ที่ได้ไปแสดงยังอุปกรณ์แสดงผล หน่วยควบคุมจึงจะส่งสัญญาณคำสั่งไปยังอุปกรณ์แสดงผลลัพธ์ ที่กำหนดให้ดึงข้อมูลจากหน่วยความจำหลักออกไปแสดงให้เห็นผลลัพธ์ดังกล่าว อีกต่อหนึ่ง

Inside a Microprocessor

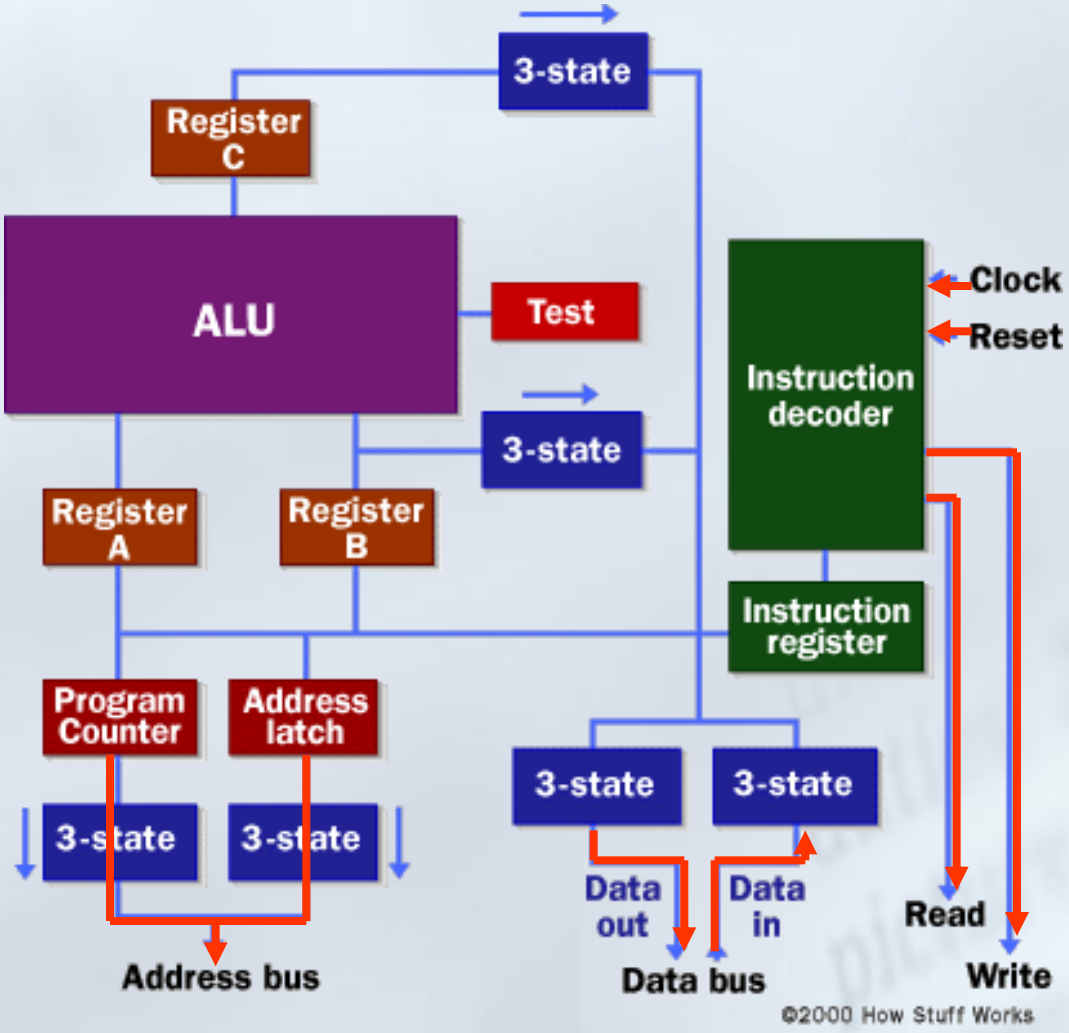
3. หน่วยความจำหลัก (Main Memory)

คอมพิวเตอร์จะสามารถทำงานได้เมื่อมีข้อมูล และชุดคำสั่งที่ใช้ในการประมวลผลอยู่ในหน่วยความ จำหลักเรียบร้อยแล้วเท่านั้น และหลังจากทำการประมวลผลข้อมูลตามชุดคำสั่งเรียบร้อบแล้ว ผลลัพธ์ที่ได้จะถูกนำไปเก็บไว้ที่หน่วยความจำหลัก และก่อนจะถูกนำออกไปแสดงที่อุปกรณ์แสดงผล

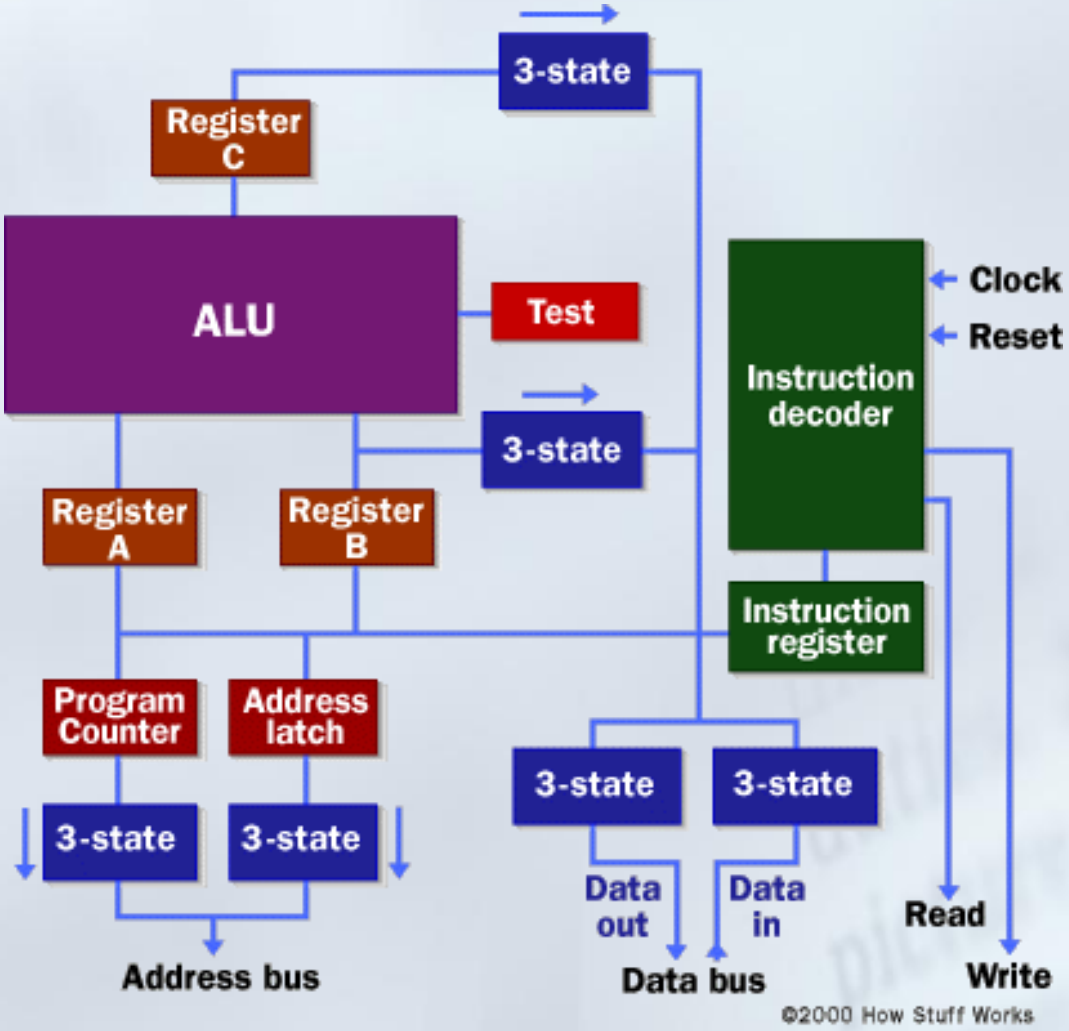
พื้นฐานการทำงานของไมโครโปรเซสเซอร์

- ใช้ ALU (Arithmetic/Logic Unit) เพื่อประมวลผลทางคณิตศาสตร์
- สามารถย้ายข้อมูลจากตำแหน่งในหน่วยความจำไปไว้ที่อื่น
- สามารถทำการตัดสินใจเพื่อกระโดดข้ามไปทำงานยังชุดคำสั่งใหม่

เส้นทางเดินข้อมูลของไมโครโปรเซสเซอร์



ส่วนประกอบของไมโครโปรเซสเซอร์



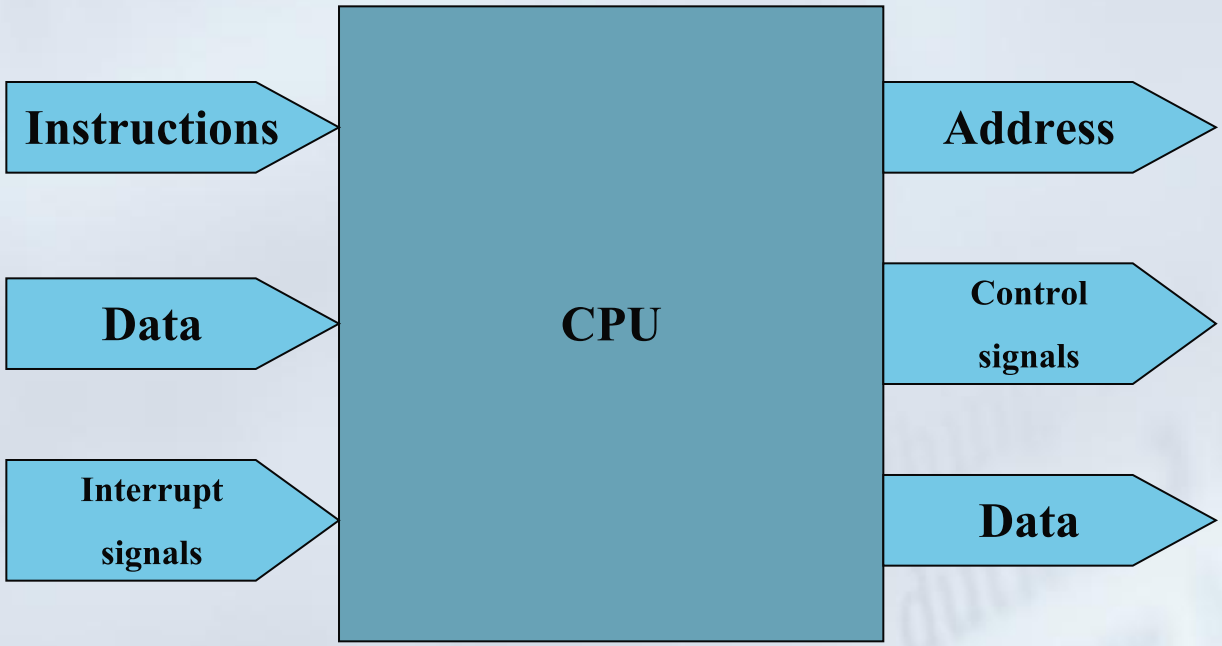
การทำงานของไมโครโปรเซสเซอร์ (1/2)

- อุปกรณ์นำเข้าข้อมูลจะนำคำสั่งและข้อมูลเข้ามาที่ CPU ตามลำดับ
- คำสั่งในโปรแกรมไม่ได้ทำงานเรียงตามลำดับเสมอไป บางครั้งอาจมีการข้าม (jump) ไปทำงานที่คำสั่งในลำดับอื่น
- การทำงานกับข้อมูลก็อาจจะมีการอ้างอิงถึงข้อมูลส่วนอื่นเพิ่มเติมนอกเหนือจากข้อมูลในลำดับที่กำหนดไว้ล่วงหน้า ดังนั้นจึงต้องมีที่พักชั่วคราวสำหรับเก็บข้อมูลและคำสั่ง หน่วยพักข้อมูลนี้เรียกว่าหน่วยความจำหลัก (Main memory)

การทำงานของไมโครโปรเซสเซอร์ (2/2)

- CPU แลกเปลี่ยนข้อมูลกับหน่วยความจำโดยการใช้ Register ที่สำคัญสองตัวคือ
  - MAR ทำหน้าที่กำหนดตำแหน่งข้อมูลในหน่วยความจำที่ต้องการอ่านหรือบันทึกข้อมูล
  - MBR ทำหน้าที่ในการเก็บข้อมูลที่อ่านมาจากหน่วยความจำหรือเก็บข้อมูลที่เตรียมไว้สำหรับบันทึกในหน่วยความจำ
- CPU แลกเปลี่ยนข้อมูลกับอุปกรณ์ภายนอกโดยใช้ Register สองตัวคือ
  - I/O AR ทำหน้าที่เก็บชื่อ (หมายเลข) ของหน่วย I/O ที่ต้องการอ่านหรือบันทึกข้อมูล
  - I/O BR ทำหน้าที่เก็บข้อมูลที่อ่านมาจาก I/O หรือเตรียมไว้สำหรับส่งไปยัง I/O ที่ต้องการ

Processor


Instructions
Data
Interrupt signals
CPU
Address
Control signals
Data

Over Clock

การโอเวอร์คล็อก (Over Clock)

เป็นขั้นตอนหรือวิธีการที่จะทำให้ซีพียู ไม่ทำงานที่ความเร็วตามที่ระบุไว้ในสเปกเดิมจากโรงงาน โดยทำให้ซีพียูทำงานในความเร็วสูงกว่าทั้งสัญญาณนาฬิกาภายในตัวซีพียู และ/หรือ ของระบบบัสในซีพียู การโอเวอร์คล็อกทำเพื่อเพิ่มประสิทธิภาพของระบบคอมพิวเตอร์ โดยให้เสียค่าใช้จ่ายน้อยที่สุด การโอเวอร์คล็อกโดยทั่วไปต้องการเพียงการเปลี่ยนค่าหรือเปลี่ยนการเซ็ตจัมเปอร์เล็กน้อยบนเมนบอร์ดเท่านั้น เราก็จะระบบคอมพิวเตอร์ที่มีความเร็วสูงขึ้น บางกรณีอาจต้องซื้ออุปกรณ์เพิ่ม อย่างระบบระบายความร้อนเช่น พัดลม เป็นต้น

Over Clock

การโอเวอร์คล็อกอาจทำให้เกิดปัญหาที่ตามมามากมาย เช่น ความเสียหายที่เกิดกับซีพียู เนื่องจากโอเวอร์คล็อกไม่ถูกวิธี เป็นต้น ฉะนั้นการโอเวอร์คล็อกจึงควรระมัดระวัง เพราะอาจทำให้ต้องสูญเสียข้อมูลหรือทรัพยากรอันมีค่าไป

64-bit Processors

- มีมาตั้งแต่ปี 1992 และในศตวรรษที่ 21 ก็เริ่มอยู่ในสายการผลิตของ Intel และ AMD รวมทั้ง Mac G5 ก็ใช้ CPU แบบ 64 บิต ด้วย
- โปรเซสเซอร์แบบ 64 บิต ประกอบไปด้วย ALU แบบ 64 บิต รีจิสเตอร์แบบ 64 บิต และบัสแบบ 64 บิต
- ชิพ 32 บิต สามารถเข้าถึงหน่วยความจำได้ขนาดสูงสุด 2 GB หรือ 4 GB ในขณะที่ชิพ 64 บิต สามารถเข้าถึงหน่วยความจำได้สูงสุดประมาณพันล้าน GB
- ด้วยบัสขนาด 64 บิต ช่วยให้สามารถเข้าถึงอุปกรณ์ I/O ได้เร็วขึ้น ทำให้ประสิทธิภาพของเครื่องคอมพิวเตอร์เพิ่มขึ้น

Memory Unit

Memory Unit

หน่วยความจำ (Memory Unit)

หน่วย หรือ อุปกรณ์ที่ใช้ในการจัดเก็บข้อมูลของระบบคอมพิวเตอร์

Memory Unit

PC-133 COMPLIANT
PC133U-333-542

maxell
SUPER DISK
120MB

650MB
SONY

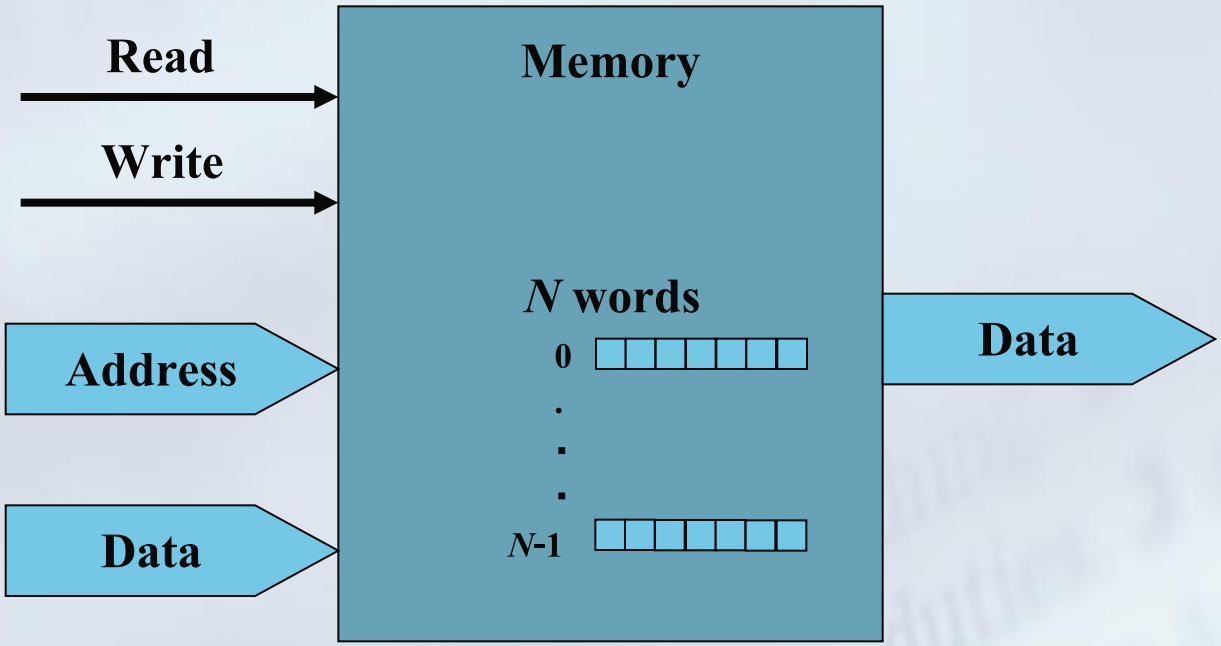
Memory Address

- หน่วยความจำประกอบไปด้วยตำแหน่งของข้อมูลในหน่วยความจำ (Memory address) ซึ่งกำหนดโดยการใช้ตัวเลขที่เรียงลำดับต่อเนื่องจากเลขต่ำสุดคือ 0 เพิ่มขึ้นไปเรื่อยๆ ตามหมายเลขสูงสุด ซึ่งขึ้นอยู่กับชิพหน่วยความจำในคอมพิวเตอร์เครื่องนั้น
- โดยปกติหมายเลขที่ใช้จะเป็นเลขฐานสอง สิ่งที่เก็บอยู่ในแต่ละตำแหน่งในหน่วยความจำนั้นก็เป็นเลขฐานสอง ซึ่งอาจจะหมายถึงข้อมูลหรือคำสั่งก็ได้

โครงสร้างการเชื่อมต่อของหน่วยความจำ

- โดยทั่วไป หน่วยความจำประกอบด้วยเซลล์ความจำขนาด N หน่วย (words) ซึ่งแต่ละหน่วยจะต้องมีขนาดเท่ากันทั้งหมด แต่ละหน่วยจะถูกตั้งชื่อสำหรับการอ้างอิงโดยใช้หมายเลข (0, 1, 2, …)
- ข้อมูลจะถูกอ่านหรือบันทึกลงในหน่วยความจำครั้งละหนึ่งหน่วย ลักษณะการทำงานของหน่วยความจำจะถูกควบคุมด้วยสัญญาณ Read หรือ Write
- ตำแหน่งของข้อมูลที่ถูกอ่านหรือถูกบันทึกจะกำหนดโดย Address และข้อมูลจะถูกส่งเข้ามาหรือส่งออกไปผ่าน Data

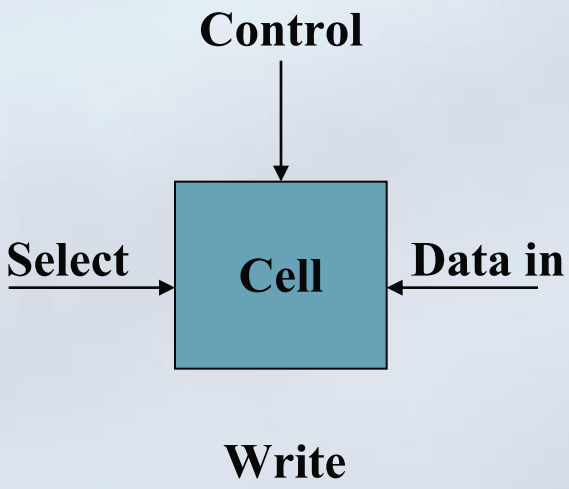
Memory



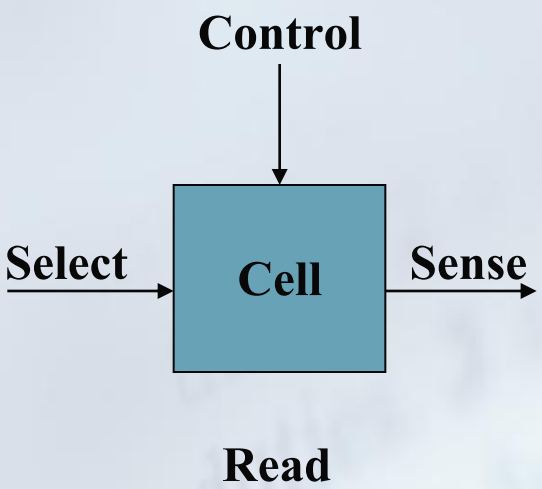
คุณลักษณะของเซลล์หน่วยความจำ

- เซลล์หน่วยความจำเป็นหน่วยย่อยเล็กที่สุดของเซมิคอนดักเตอร์
- เซลล์มีสถานะที่เสถียรสองสถานะ ซึ่งนำมาใช้แทนความหมายบิต 0 และบิต 1
- เซลล์สามารถที่จะถูกทำให้เปลี่ยนสถานะ (Write) ได้
- เซลล์สามารถตรวจสอบสถานะ (Read) ได้

การทำงานของเซลล์หน่วยความจำ



Write



Read

Random Access Memory (RAM)

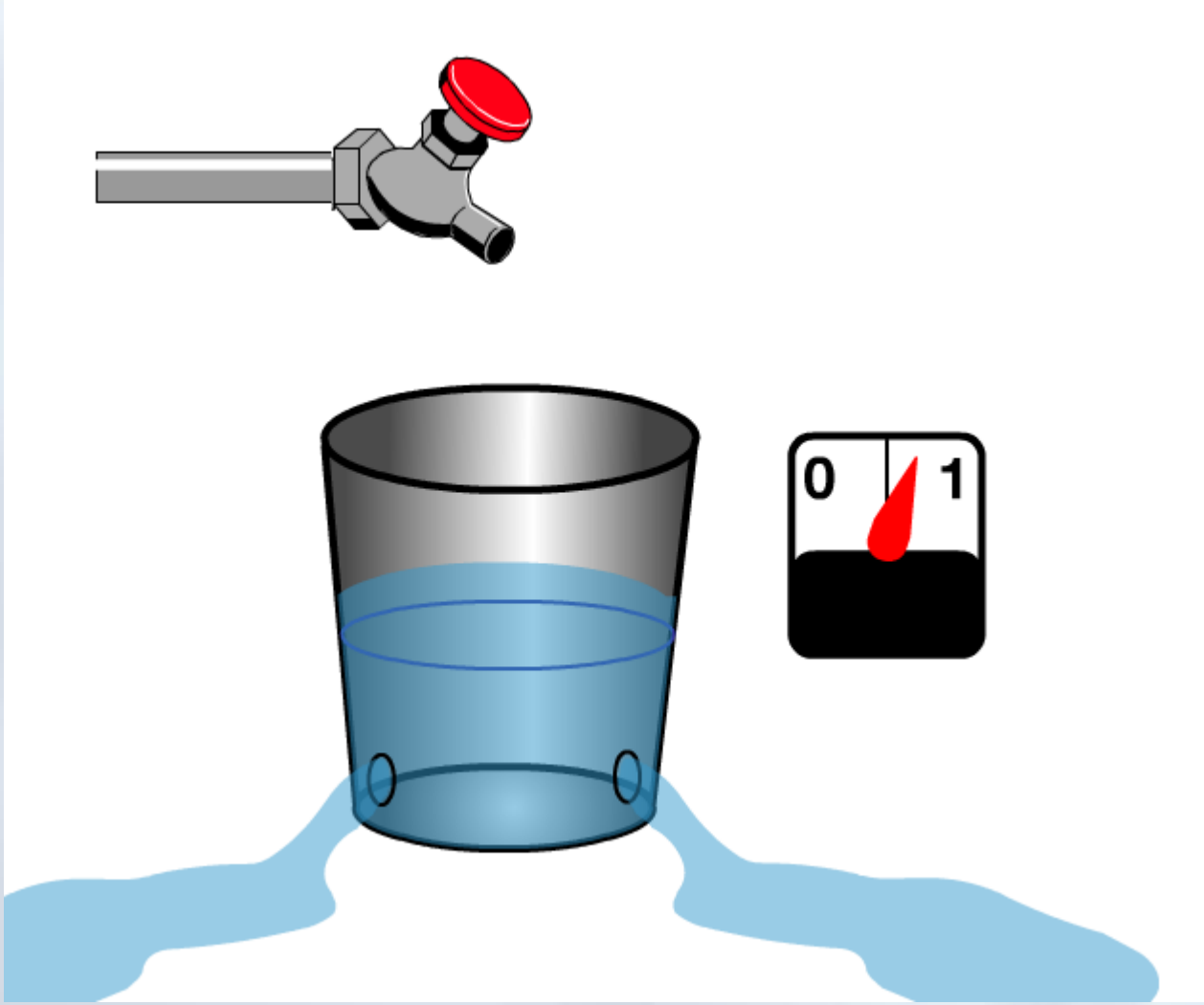
- สามารถอ่านหรือบันทึกข้อมูลได้อย่างง่ายดายและรวดเร็วด้วยการใช้สัญญาณไฟฟ้า
- ข้อมูลที่เก็บอยู่นั้นเป็นการเก็บไว้ชั่วคราว หน่วยความจำ RAM จะต้องได้รับพลังงานไฟฟ้ามาป้อนอยู่อย่างต่อเนื่องตลอดเวลา
- RAM แบ่งออกเป็นสองประเภท
  - DRAM (Dynamic RAM)
  - SRAM (Static RAM)

Dynamic RAM (DRAM)

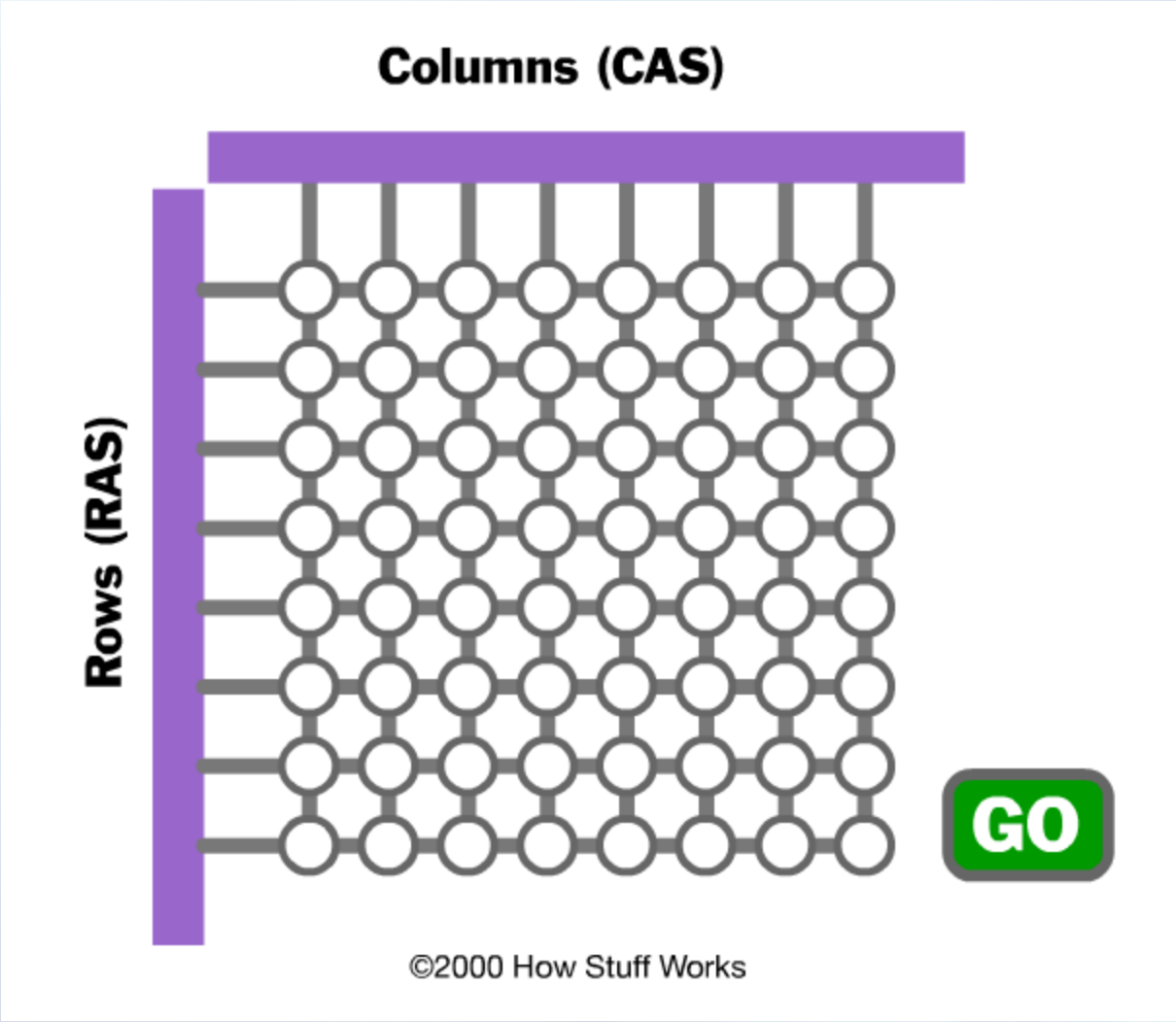
- DRAM ประกอบด้วยเซลล์ที่ใช้เก็บข้อมูลที่ใช้วิธีการอัดประจุไฟฟ้าเข้าไปเก็บไว้ในตัว Capacitor เพื่อใช้แทนความหมายบิต 0 หรือบิต 1
- Capacitor มีคุณสมบัติทางธรรมชาติที่จะสูญเสียประจุไฟฟ้าในตัวไปได้เอง ดังนั้นจะต้องการอัดประจุไฟฟ้าเข้ามาในตัวเองเป็นระยะๆ อยู่เสมอ เพื่อที่จะได้สามารถเก็บรักษาข้อมูลเอาไว้ได้ เรียกว่าการ Refresh
- สำหรับการบันทึกข้อมูลจะมีกระแสไฟฟ้าเข้ามา กระแสไฟฟ้าแรงดันสูงจะหมายถึงบิต 1 และกระแสไฟฟ้าแรงดันต่ำจะหมายถึงบิต 0
- สำหรับการอ่านข้อมูลเซลล์จะถ่ายทอดประจุไฟฟ้าไปยังตัวรับสัญญาณเพื่อเปรียบเทียบแล้วแปลให้เป็นสัญญาณบิต 0 หรือบิต 1

Capacitor

Recharging capacitor



การบันทึกข้อมูลในหน่วยความจำ



การอ่านข้อมูลในหน่วยความจำ

- Sense-amplifier จะตรวจสอบระดับของกระแสไฟฟ้าใน Capacitor
- ถ้าใน Capacitor มีกระแสไฟฟ้ามากกว่า 50% จะอ่านค่าได้เป็น 1 ถ้าต่ำกว่านั้นจะอ่านค่าเป็น 0
- ขั้นตอนทั้งหมดจะใช้ระยะเวลาที่สั้นมากในระดับ nanoseconds
- ถ้า memory chip rating เป็น 70 ns หมายความว่าใช้เวลา 70 ns ในกระบวนการอ่านและ recharge ข้อมูลในแต่ละเซลล์

รูปแบบของหน่วยความจำ

SIMM (Single in-line memory module)

DIMM (Dual in-line memory module)

SODIMM (Small outline dual in-line memory module)

SIMM (Single in-line memory module)

- มีจำนวนขาเชื่อมต่อขนาด 30 พิน และมีขนาดประมาณ 9 x 2 ซม.
- ในการติดตั้ง SIMM จะต้องติดตั้งเป็นคู่ เพราะความกว้างของบัสมากกว่า SIMM 2 เท่า (บัส 16 บิต, SIMM 8 บิต)
- SIMM รุ่นล่าสุดมีขนาดใหญ่ขึ้นคือ 11 x 2.5 ซม. มีตัวเชื่อมต่อแบบ 72 พิน ช่วยเพิ่มแบนวิดธ์ (Bandwidth) และเพิ่มความจุต่อโมดูลเป็น 256 MB

DIMM (Dual in-line memory module)

- มีขนาด 14 x 2.5 ซม. และมีขาเชื่อมต่อขนาด 168 พิน หรือ 184 พิน
- ความจุต่อโมดูลตั้งแต่ 8 MB ถึง 1 GB
- สามารถติดตั้งตัวเดียวได้
- ใช้กระแสไฟฟ้า 2.5 โวลต์

SODIMM (Small outline DIMM)

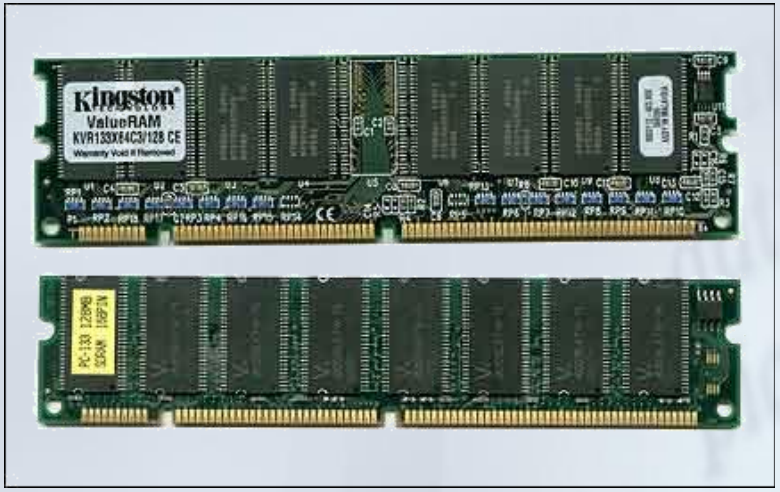
- เป็นหน่วยความจำของเครื่องคอมพิวเตอร์ Notebook
- มีขนาดเล็กมาก ประมาณ 5 x 2.5 ซม. และมีขาเชื่อมต่อขนาด 144 พิน หรือ 200 พิน
- มีความจุต่อโมดูลตั้งแต่ 16 MB ถึง 1 GB

EDODRAM (Extended data-out DRAM)

- หลังจากที่ได้รับข้อมูลบิตแรกแล้วก็จะรับข้อมูลบิตถัดไปทันที โดยไม่รอให้ประมวลผลคำสั่งเสร็จ
- อัตราการส่งข้อมูลไปยัง L2 Cache สูงสุดประมาณ 264 MBps

SDRAM (Synchronous DRAM)

- ใช้ประโยชน์จากแนวคิดแบบ Burst mode เพื่อเพิ่มประสิทธิภาพในการทำงาน แนวคิดนี้จะมีประสิทธิภาพสูงสุดก็ต่อเมื่อคำสั่งเป็นแบบเรียงลำดับ ไม่มีคำสั่งข้ามการทำงาน
- SDRAM ทำงานเร็วกว่า EDORAM ประมาณ 5%
- อัตราการส่งข้อมูลไปยัง L2 Cache สูงสุดประมาณ 528 MBps



DDR SDRAM (Double data rate SDRAM)

- คล้ายกับ SDRAM แต่มีแบนด์วิดธ์ที่กว้างกว่า ทำให้มีความเร็วสูงกว่า
- มีความเร็วเป็น 2 เท่าของ SDRAM โดยทำงานทั้งขาขึ้นและขาลงของสัญญาณนาฬิกา
- อัตราการส่งข้อมูลไปยัง L2 Cache สูงสุดประมาณ 1,064 MBps (สำหรับ DDR SDRAM 133 MHz)

RDRAM (Rambus DRAM)

- พัฒนาโดยบริษัท Rambus
- ใช้ RIMM (Rambus in-line memory module) ซึ่งมีขนาดและจำนวนขาเชื่อมต่อใกล้เคียงกับ DIMM
- RDRAM ใช้บัสความเร็วสูงที่ออกแบบมาพิเศษ เรียกว่า Rambus Channel
- ชิพหน่วยความจำทำงานแบบขนานมีอัตราเร็วอยู่ที่ 800 MHz หรือ 1,600 MBps
- เนื่องจาก RDRAM ทำงานด้วยความเร็วสูงจึงทำให้เกิดความร้อนสูงมาก จึงต้องมีอุปกรณ์สำหรับระบายความร้อน

ตารางเปรียบเทียบประเภทของ RAM

ประเภท	Bandwidth x จำนวนครั้ง x ความถี่	อัตราการส่งข้อมูล
PC-133 (SDRAM)	8 x 1 x 133	1,064 MB/s
PC-2100 (DDR SDRAM)	8 x 2 x 133	2,128 MB/s
PC-2700 (DDR SDRAM)	8 x 2 x 166	2,656 MB/s
PC-3200 (DDR SDRAM)	8 x 2 x 200	3,200 MB/s
PC-3600 (DDR SDRAM)	8 x 2 x 225	3,600 MB/s
PC-4000 (DDR SDRAM)	8 x 2 x 250	4,000 MB/s
PC-800 (RDRAM)	2 x 2 x 400	1,600 MB/s
PC-1066 (RDRAM)	4 x 2 x 533	4,264 MB/s

Static RAM (SRAM)

- SRAM มีการจัดเรียงอุปกรณ์ภายในเป็นลักษณะเดียวกันกับโครงสร้างของโปรเซสเซอร์
- ค่าไบนารี่ (0 หรือ 1) จะถูกเก็บไว้ด้วย Flip-flop logic gate ซึ่งสามารถเก็บค่าไว้ในตัวเองได้นานตราบเท่าที่มีพลังงานไฟฟ้าป้อนให้อย่างต่อเนื่อง
- ในการเก็บข้อมูลจะใช้ทรานซิสเตอร์จำนวน 2 ตัว
  - ที่สถานะ “1” C_1 มีแรงดันไฟฟ้าสูง ส่วน C_2 มีแรงดันไฟฟ้าต่ำ
  - ที่สถานะ “0” C_1 มีแรงดันไฟฟ้าต่ำ ส่วน C_2 มีแรงดันไฟฟ้าสูง

เปรียบเทียบระหว่าง SRAM กับ DRAM

- DRAM มีโครงสร้างที่ง่ายกว่าและมีขนาดเล็กกว่า SRAM จึงสามารถบรรจุจำนวนเซลล์ต่อพื้นที่ได้มากกว่า ทำให้มีราคาถูกกว่า SRAM ที่มีความจุเท่าๆ กัน
- DRAM จำเป็นต้องมีวงจรสำหรับการกระตุ้น (refresh cycle) เป็นระยะๆ อย่างสม่ำเสมอ
- SRAM มีความเร็วในการทำงานสูงกว่า DRAM
- ระบบคอมพิวเตอร์ส่วนใหญ่นำ DRAM มาใช้เป็นหน่วยความจำหลักและนำ SRAM มาใช้เป็นหน่วยความจำ Cache

Read Only Memory (ROM)

- ประกอบไปด้วยข้อมูลที่ถูกบันทึกไว้เป็นการถาวร (Nonvolatile) ซึ่งไม่สามารถแก้ไขได้
- ROM ไม่จำเป็นต้องอาศัยพลังงานไฟฟ้าในการจัดเก็บข้อมูล
- ข้อดีของ ROM คือการที่ข้อมูลหรือโปรแกรมถูกจัดเก็บไว้อย่างถาวรในหน่วยความจำ จึงไม่มีความจำเป็นต้องถูกอ่านขึ้นมาจากอุปกรณ์ I/O ภายนอก ซึ่งจะต้องเสียเวลาเป็นอย่างมาก
- ROM ถูกสร้างขึ้นมาจากแผงวงจรรวม (IC) โดยจัดการบันทึกข้อมูลหรือโปรแกรมลงไปพร้อมกัน ทำให้เกิดค่าใช้จ่ายสูง
- ในระหว่างกระบวนการผลิตจะมีข้อผิดพลาดเกิดขึ้นไม่ได้เลย

Flash memory

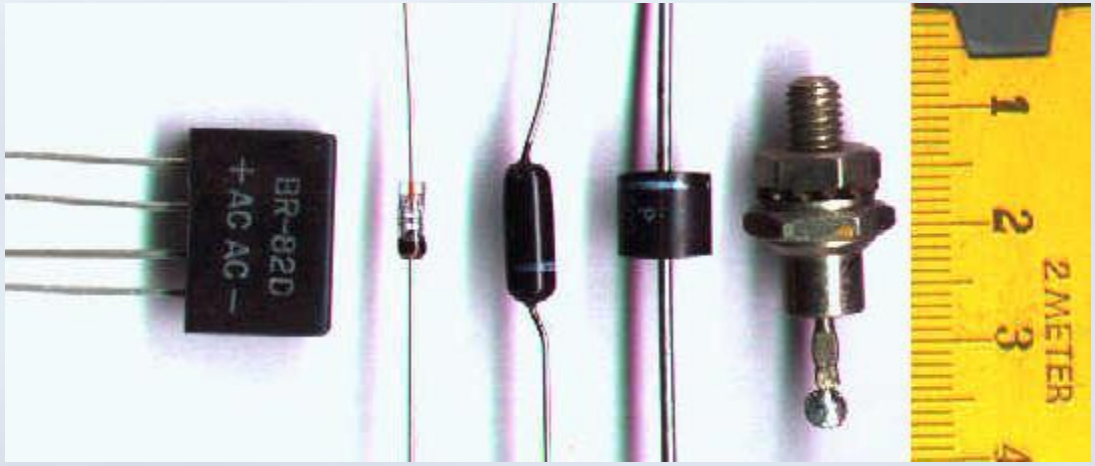
Flash Memory

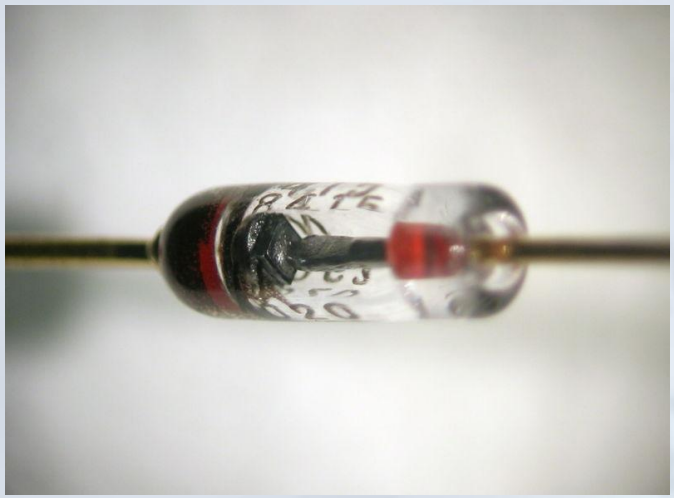
เป็นหน่วยความจำชนิดที่ไม่มีการเปลี่ยนแปลงข้อมูลหรือการสูญหายของข้อมูลแม้ว่าขณะนั้นจะไม่มีกระแสไฟฟ้าไปหล่อเลี้ยงวงจรก็ตาม ไม่เหมือนกับ EEPROM ซึ่งเป็นชิพที่ฝังตัวอยู่ในเครื่องคอมพิวเตอร์ เพราะฉะนั้น Flash Memory จึงกลายเป็นเทคโนโลยีที่โดดเด่น ลักษณะของแฟลชเมมอรีเป็นหน่วยความจำที่มีส่วนประกอบที่ไม่มีการขับเคลื่อนเหมือนกับฮาร์ดดิสก์จึงเป็นที่ต้องการของผู้ใช้ทั่วไป

การทำงานของ ROM

- ชิพของ ROM จะคล้ายกับ RAM คือเก็บข้อมูลในลักษณะที่เป็นคอลัมน์และแถว แต่จุดที่คอลัมน์และแถวตัดกันจะแตกต่างกัน
- RAM ใช้ทรานซิสเตอร์เพื่อเปิดหรือปิดการเข้าถึง Capacitor
- ROM ใช้ Diode ในการเชื่อมต่อสายสัญญาณถ้าข้อมูลมีค่าเป็น 1 ถ้าข้อมูลมีค่าเป็น 0 สายสัญญาณจะไม่ถูกเชื่อมต่อ

Diode

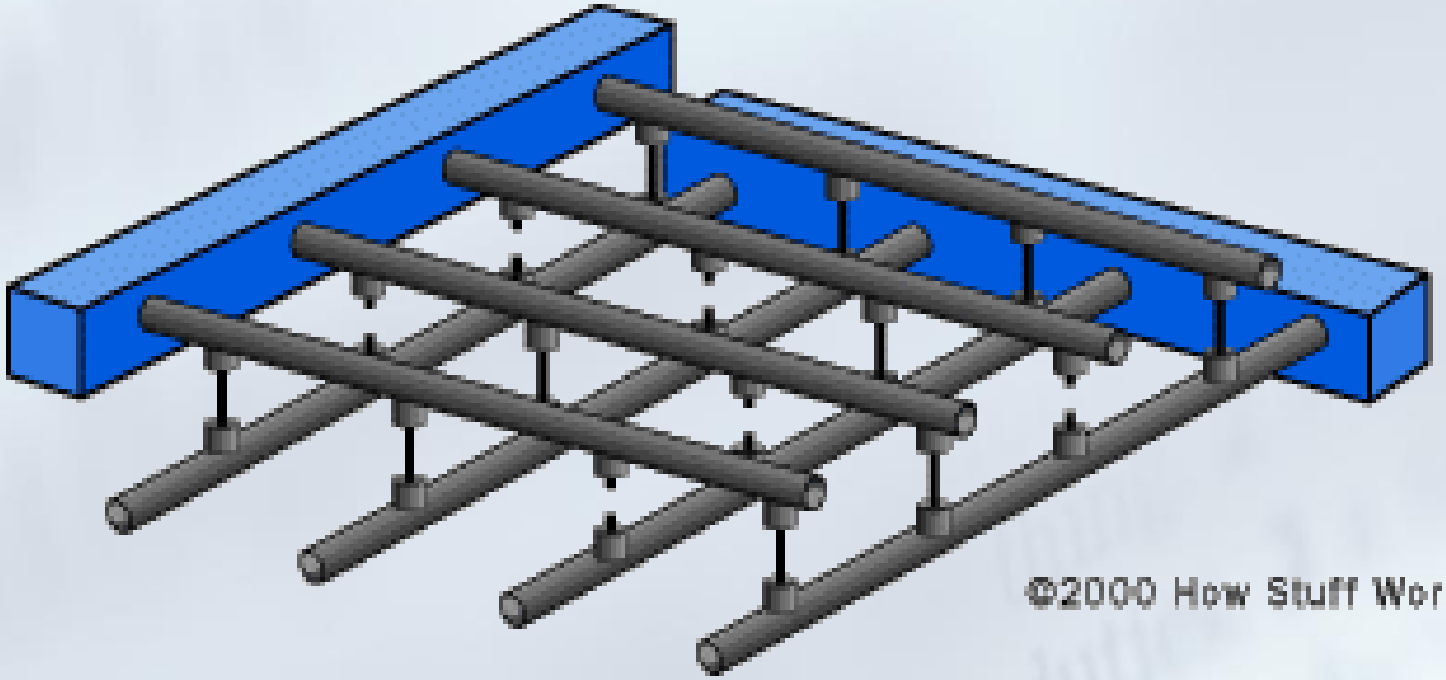




PROM (Programmable ROM)

- เป็นชิพ ROM แบบว่างเปล่า มีราคาไม่แพง สามารถบรรจุโปรแกรมลงไปได้โดยใช้เครื่องมือพิเศษที่เรียกว่า Programmer
- PROM แตกต่างจาก ROM ทั่วไปคือส่วนที่ตัดกันของคอลัมน์และแถวจะมีฟิวส์อยู่ ซึ่งจะมีค่าเป็น 1
- หากต้องการเปลี่ยนค่าให้เป็น 0 ก็จะปล่อยกระแสไฟฟ้าแรงดันสูงเข้าไปเพื่อให้ฟิวส์ขาดออกจากกัน กระบวนการนี้เรียกว่า Burning the PROM
- PROM สามารถบรรจุโปรแกรมลงไปได้ครั้งเดียวเท่านั้น

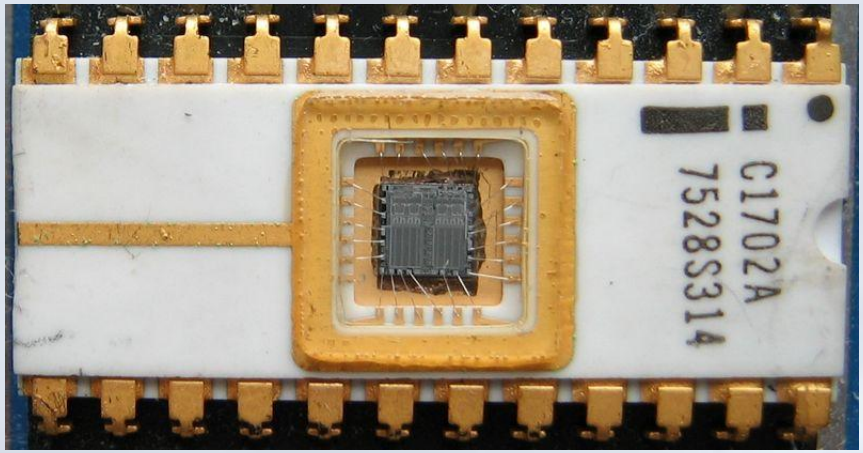
PROM Fuse



PROM

EPROM (Erasable PROM)

- EPROM สามารถเขียนข้อมูลลงไปใหม่ได้หลายครั้ง
- การลบข้อมูลใน EPROM ต้องใช้เครื่องมือพิเศษในการส่งคลื่นความถี่ของแสง Ultraviolet (UV) ลงไปในช่องรับแสงของชิพ



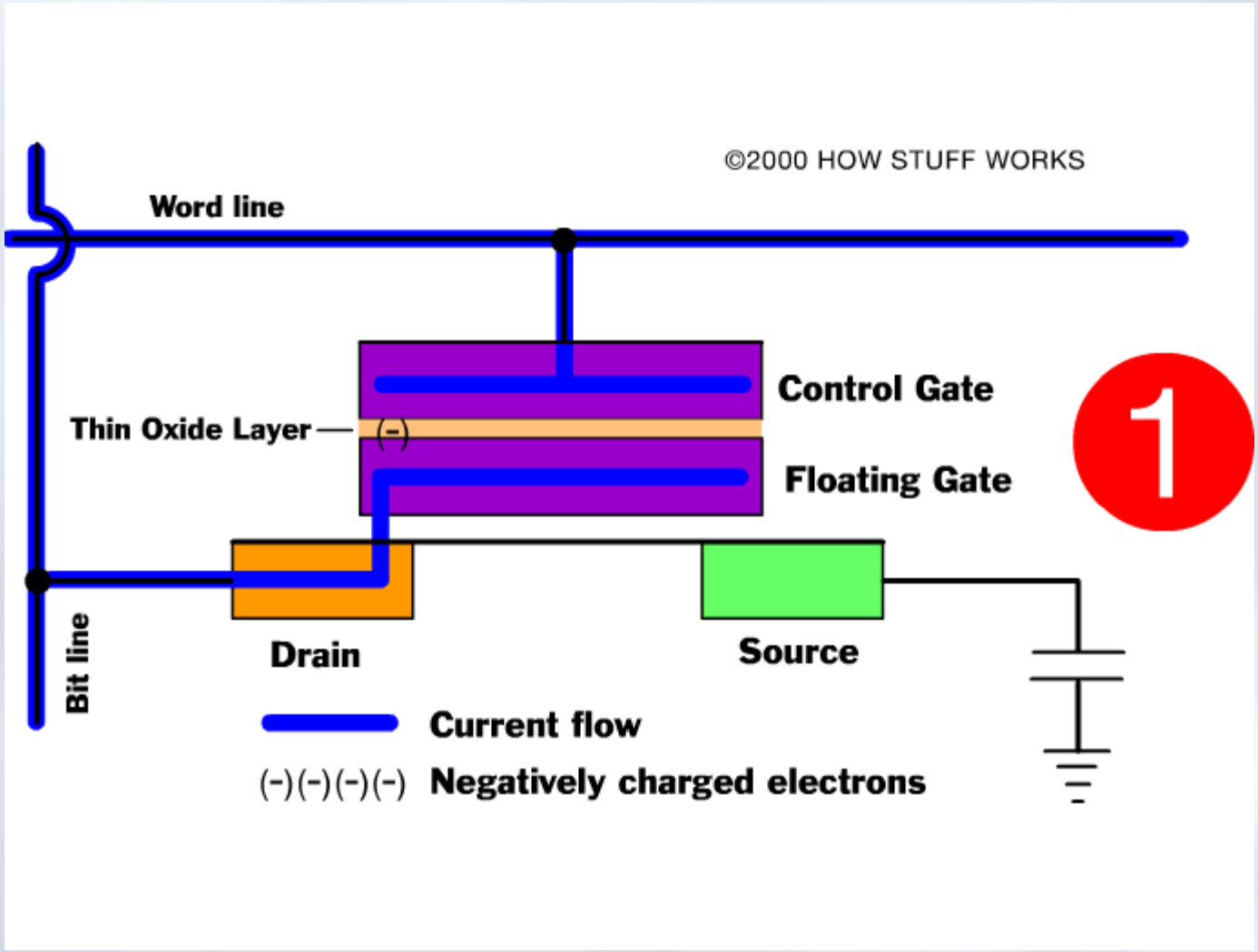
การบันทึกข้อมูลของ EPROM

- ส่วนที่ตัดกันของคอลัมน์และแถวจะมีทรานซิสเตอร์สองตัว ซึ่งถูกกั้นด้วยชั้นของ thin oxide ตัวหนึ่งเรียกว่า Floating gate อีกตัวหนึ่งเรียกว่า Control gate
- Floating gate จะถูกเชื่อมต่อเข้ากับแถว (Word line) ซึ่งถูกส่งผ่าน Control gate
- เซลล์จะมีค่าเป็น 1 หากต้องการเปลี่ยนให้เป็น 0 การประมวลผลแบบพิเศษที่เรียกว่า Fowler-Nordheim tunneling
- Tunneling ใช้สำหรับเปลี่ยนสถานะของอิเล็กตรอนใน Floating gate

การชาร์จกระแสไฟฟ้า

- ในการชาร์จจะใช้กระแสไฟฟ้าประมาณ 10 ถึง 13 โวลต์ ซึ่งกระทำจาก Bit line ผ่าน Floating gate และระบายลง Ground
- การชาร์จด้วยอิเล็กตรอน (ขั้ว -) จะทำหน้าที่เหมือนกับกำแพงกั้นระหว่าง Control gate และ Floating gate
- Cell sensor จะคอยตรวจสอบดูระดับของการชาร์จผ่าน Floating gate ถ้ามีค่ามากกว่า 50% จะมีค่าเป็น 1 ถ้าระดับลดลงน้อยกว่า 50% จะมีค่าเป็น 0
- EPROM เปล่า ทุก Gate จะถูกเปิดเอาไว้ ทำให้มีค่าเป็น 1

การเปลี่ยนสถานะของบิต



การลบข้อมูลใน EPROM

- ในการลบจะต้องใช้พลังงานมากพอที่จะทำลายอิเล็กตรอนที่กั้น Floating gate อยู่
- สิ่งที่ดีที่สุดคือลำแสง UV ที่ความถี่ 253.7 เพราะความถี่นี้จะไม่สามารถทะลุผ่านพลาสติกหรือแก้วได้
- ไม่สามารถเลือกลบข้อมูลบางส่วนได้ ต้องลบทั้งหมดเท่านั้น
- ในการลบจะต้องถอดออกจากอุปกรณ์ไปไว้ภายใต้ลำแสง UV เป็นเวลาหลายนาที แต่ถ้าใช้เวลานานเกินไปจะทำให้เกิด Over-erased คือ Floating gate จะไม่สามารถเก็บประจุอิเล็กตรอนได้

EEPROM (Electrical EPROM)

- ไม่ต้องถอดออกจากอุปกรณ์เพื่อทำการลบข้อมูล
- ไม่สามารถลบข้อมูลแบบสมบูรณ์เพื่อเปลี่ยนแปลงวัตถุประสงค์ในการใช้งานได้
- การเปลี่ยนแปลงข้อมูลไม่ต้องใช้อุปกรณ์ใดๆ เพิ่มเติม

การลบข้อมูล

- สามารถเปลี่ยนค่าของอิเล็กตรอนภายในเซลล์ได้โดยใช้ซอฟต์แวร์เฉพาะ
- EEPROM สามารถเปลี่ยนค่าของข้อมูลครั้งละ 1 ไบต์ ซึ่งจะคงอยู่แบบถาวรแต่ช้า
- EEPROM ช้าเกินไปที่จะนำไปใช้ในอุปกรณ์ที่ต้องการเปลี่ยนแปลงข้อมูลในชิพอย่างรวดเร็ว

การตรวจสอบและแก้ไขข้อผิดพลาดของหน่วยความจำ

Hard failure

- หมายถึง ข้อผิดพลาดที่เกิดขึ้นเนื่องจากชิพที่ผลิตขึ้นมานั้นไม่สามารถทำงานได้ตามที่ต้องการ
- เป็นความผิดปกติแบบถาวร เช่น เซลล์หน่วยความจำไม่สามารถเก็บข้อมูลไว้ได้ หรือเก็บเพียงข้อมูล “0” หรือ “1” แต่เพียงอย่างเดียว
- อาจเกิดขึ้นจากสภาพการใช้งานที่ไม่เหมาะสม การใช้งานผิดประเภท หรือการเสื่อมสภาพตามอายุการใช้งาน

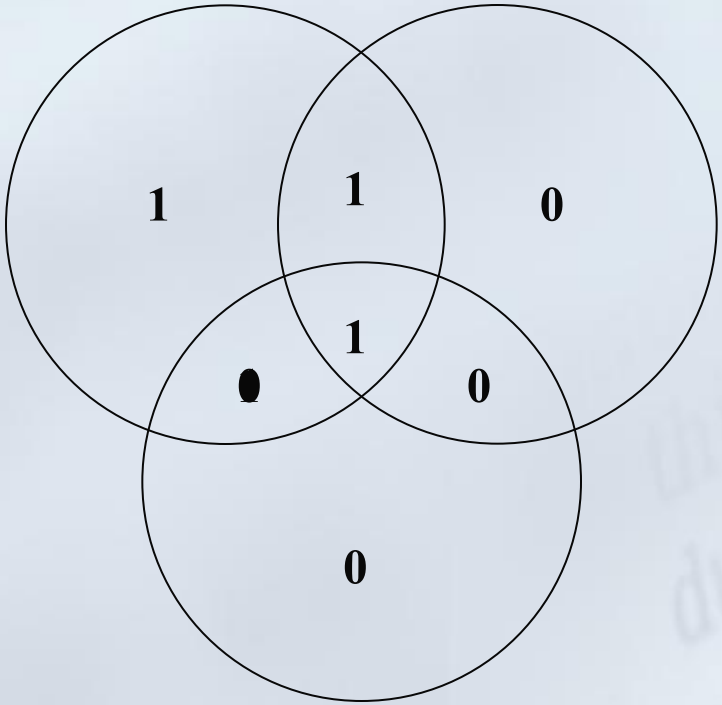
Soft error

- เป็นข้อผิดพลาดที่เกิดขึ้นอย่างไม่สามารถคาดเดาได้
- เกิดจากเหตุการณ์บางอย่างที่ทำให้ข้อมูลในเซลล์หน่วยความจำบางเซลล์หรือบางกลุ่มเปลี่ยนไปจากเดิม โดยไม่ได้ทำลายเซลล์หน่วยความจำอย่างถาวร
- อาจเกิดขึ้นเนื่องจากกระแสไฟฟ้าที่ไม่สม่ำเสมอ หรือการคายประจุไฟฟ้าสถิตที่มีอยู่ทุกหนทุกแห่ง

ฟังก์ชั่นในการแก้ไขข้อผิดพลาด

- ถ้าตรวจไม่พบข้อผิดพลาดใดๆ ข้อมูลนั้นก็จะถูกส่งไปใช้งาน
- สามารถตรวจพบข้อผิดพลาดและสามารถที่จะแก้ไขได้ ข้อมูลและโค้ดรหัสจะถูกส่งไปให้หน่วยแก้ไขข้อมูล ซึ่งจะสร้างข้อมูลใหม่ที่ถูกต้อง และส่งไปใช้งานต่อไป
- สามารถตรวจพบข้อผิดพลาดแต่ไม่สามารถแก้ไขได้ ก็จะแสดงรายงานให้ผู้ใช้ทราบ
- โค้ดรหัสที่สามารถแก้ไขข้อผิดพลาดได้เรียกว่า Error-correcting code (ECC)

วิธีแก้ไขข้อผิดพลาดของ Hamming *



* Richard Hamming, Bell Laboratories, USA

Syndrome word

- คือ ผลที่ได้จากการเปรียบเทียบข้อมูลทางตรรกะ (Comparison logic) โดยใช้ฟังก์ชั่น exclusive OR (ถ้าบิตเหมือนกันจะได้ค่าเป็น 0 ถ้าบิตต่างกันได้ค่าเป็น 1)
- เนื่องจากข้อผิดพลาดอาจเกิดขึ้นในบิตใดๆ ก็ได้จากข้อมูล M บิต และข้อมูลสำหรับการตรวจสอบอีก K บิต ดังนั้นจะได้ความสัมพันธ์

$$2^k - 1 >= M + K$$

- สมมติว่ามีข้อมูลขนาด 8 บิต ($M = 8$) จะได้ว่า
  - ถ้า $K = 3$ จะได้ว่า $2^3 - 1 < 8 + 3$
  - ถ้า $K = 4$ จะได้ว่า $2^4 - 1 > 8 + 4$

การสร้าง Syndrome word ขนาด 4 บิต สำหรับข้อมูล 8 บิต

- ถ้า Syndrome word มีค่าเป็น “00000000” แสดงว่าไม่มีข้อผิดพลาดในข้อมูลนั้น
- ถ้า Syndrome word มี “1” แทรกอยู่เพียงบิตเดียว แสดงว่าข้อผิดพลาดเกิดขึ้นในบิตตรวจสอบ (1 ใน 4 บิต) จึงไม่จำเป็นต้องแก้ไข
- ถ้า Syndrome word มี “1” มากกว่า 1 บิต แล้ว ค่าของบิต Syndrome word จะบอกตำแหน่งข้อมูลที่ผิดพลาด ซึ่งจะต้องกลับบิตที่ตำแหน่งนั้นเป็นบิตตรงกันข้าม

ตำแหน่งของบิตข้อมูลและบิตตรวจสอบ

Bit position	12	11	10	9	8	7	6	5	4	3	2	1
Position Number	1100	1011	1010	1001	1000	0111	0110	0101	0100	0011	0010	0001
Data bit	D8	D7	D6	D5		D4	D3	D2		D1		
Check bit					C8				C4		C2	C1

การคำนวณบิตตรวจสอบ

$C1 = D1 \oplus D2 \oplus D4 \oplus D5 \oplus D7$

$C2 = D1 \oplus D3 \oplus D4 \oplus D6 \oplus D7$

$C4 = D2 \oplus D3 \oplus D4 \oplus D8$

$C8 = D5 \oplus D6 \oplus D7 \oplus D8$

ตัวอย่างการคำนวณบิตตรวจสอบ (1/3)

- สมมติว่าบิตข้อมูลขนาด 8 บิตที่ต้องการนำมาใช้คือ “00111001” โดยที่ D1 คือบิตทางขวาสุด ดังนั้นบิตตรวจสอบ คือ

  $C1 = 1 \oplus 0 \oplus 1 \oplus 1 \oplus 0 = 1$

  $C2 = 1 \oplus 0 \oplus 1 \oplus 1 \oplus 0 = 1$

  $C4 = 0 \oplus 0 \oplus 1 \oplus 0 = 1$

  $C8 = 1 \oplus 1 \oplus 0 \oplus 0 = 0$

- บิตตรวจสอบ = 0111
- ข้อมูลที่จะส่ง = 0 0 1 1 0 1 0 0 1 1 1 1

D8 D7 D6 D5 C8 D4 D3 D2 C4 D1 C2 C1

ตัวอย่างการคำนวณบิตตรวจสอบ (2/3)

- สมมติว่าบิตที่ 3 (D3) เป็นบิตที่ผิดพลาด ข้อมูลที่ได้รับจะเป็น “00111101” บิตตรวจสอบใหม่จะเป็น

$C1 = 1 \oplus 0 \oplus 1 \oplus 1 \oplus 0 = 1$

$C2 = 1 \oplus 1 \oplus 1 \oplus 1 \oplus 0 = 0$

$C4 = 0 \oplus 1 \oplus 1 \oplus 0 = 0$

$C8 = 1 \oplus 1 \oplus 0 \oplus 0 = 0$

- บิตตรวจสอบ = 0001

ตัวอย่างการคำนวณบิตตรวจสอบ (3/3)

- เมื่อนำบิตตรวจสอบเก่าและบิตตรวจสอบใหม่มาเปรียบเทียบกันจะได้ว่า

C8	C4	C2	C1	
0	1	1	1	⊕
0	0	0	1	
0	1	1	0	

- ผลลัพธ์ที่ได้คือ “0110” แสดงว่าบิตตำแหน่งที่ 6 (Bit position) หรือ D3 เป็นบิตที่ไม่ถูกต้อง
- โค้ดรหัสที่กล่าวถึงจัดอยู่ในประเภท single-error-correcting (SEC) ซึ่งจะไม่สามารถตรวจสอบหรือแก้ไขใดๆ ถ้ามีบิตที่ผิดพลาดพร้อมกันมากกว่า 2 บิต