ระบบคอมพิวเตอร์และสถาปัตยกรรม
(Computer System and Architecture)

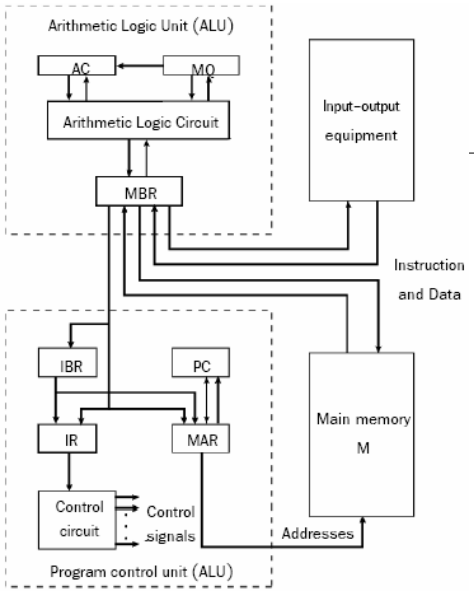
บทที่ 7
หน่วยความจำหลัก
(Main Memory)

หน่วยความจำหลัก

ประสิทธิภาพของโปรเซสเซอร์ หรือซีพียูนั้นถูกวัดจากจำนวนคำสั่ง (Instruction) ที่ซีพียูสามารถเรียกใช้งานได้ใน 1 วินาที และความเร็วดังกล่าวนี้ก็มีการเพิ่มขึ้นเป็น 2 เท่า ทุก ๆ 18 เดือน หน่วยความจำของคอมพิวเตอร์เองก็มีการพัฒนาเพิ่มขึ้นเช่นเดียวกันแต่ในทิศทางที่แตกต่าง คือหน่วยความจำนั้นจะมีการเพิ่มขนาดเป็น 4 เท่าในทุก ๆ 36 เดือน ในขณะที่ราคาคงที่ แต่ความเร็วของหน่วยความจำมีการเพิ่มขึ้นประมาณ 10% ต่อปีเท่านั้น ด้วยเหตุที่การเพิ่มความเร็วของโปรเซสเซอร์ที่มีอัตราสูงกว่าการเพิ่มความเร็วของหน่วยความจำ ทำให้ช่องว่างระหว่างความเร็วของโปรเซสเซอร์และความเร็วของหน่วยความจำเพิ่มมากขึ้น การออกแบบสถาปัตยกรรมคอมพิวเตอร์จึงต้องพิจารณาการลดช่องว่างนี้

การทำงานของหน่วยความจำ

หน่วยความจำหลัก เป็นศูนย์กลางของการทำงานต่าง ๆ ของระบบคอมพิวเตอร์ในปัจจุบัน หน่วยความจำหลัก คือพื้นที่เก็บข้อมูลขนาดใหญ่ที่ประกอบไปด้วยพื้นที่เก็บข้อมูลย่อยที่มีขนาดเป็น ไบต์ (Byte) โดยแต่ละไบต์จะมีแอ็ดเดรส (Address) บอกตำแหน่งของตัวเองด้วย ซึ่งการทำงานของระบบคอมพิวเตอร์นั้นจะเริ่มต้นจากการดึงหรือเฟ็ตช์คำสั่ง (Fetch) ออกมาจากหน่วยความจำแล้วเอ็กซิคิวต์ (Execute) คำสั่งเหล่านั้นทีละคำสั่ง

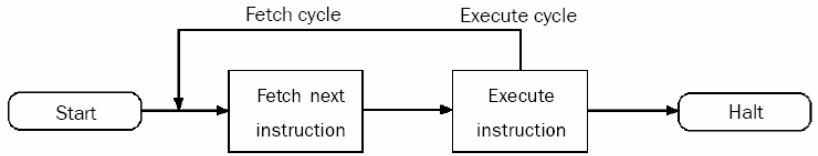


รีจิสเตอร์ (Register)

- **เมมโมรีบัพเฟอร์รีจิสเตอร์** (MBR : Memory buffer register) : เป็นหน่วยเก็บข้อมูลที่จะส่งออกไปเก็บ หรือที่ดึงมาจากหน่วยความจำ
- **เมมโมรีแอ็ดเดรสรีจิสเตอร์** (MAR : Memory address register) : เป็นตัวกำหนดที่อยู่ของหน่วยความจำที่ MBR จะทำการอ่าน หรือ เก็บข้อมูล
- **อินสตรัคชันรีจิสเตอร์** (IR : Instruction register) : เป็นส่วนที่เก็บคำสั่ง (Opcode) ที่จะถูกเรียกใช้งาน
- **โปรแกรมเคาน์เตอร์** (PC : Program counter) : เป็นส่วนที่เก็บที่อยู่ของคำสั่งถัดไปที่จะถูกดึงออกมาจากหน่วยความจำหลัก
- **แอ็กคิวมูเลเตอร์** (AC : Accumulator) : จะใช้ในเก็บค่าชั่วคราวของโอเปอแรนด์ (Operands) และผลลัพธ์ของการคำนวณในส่วน ALU

การทำงานของระบบคอมพิวเตอร์

ระบบคอมพิวเตอร์จะทำงานโดยการเรียกใช้คำสั่งจากหน่วยความจำหลักเข้ามาทำงาน หรือที่เราเรียกว่า “วงรอบคำสั่ง” (Instruction Cycle) ซึ่งในแต่ละวงรอบจะประกอบไปด้วย 2 วงรอบการทำงานย่อย คือ “วงรอบเฟ็ตช์” (Fetch Cycle) และวงรอบเอ็กซิคิวต์ (Execute Cycle)



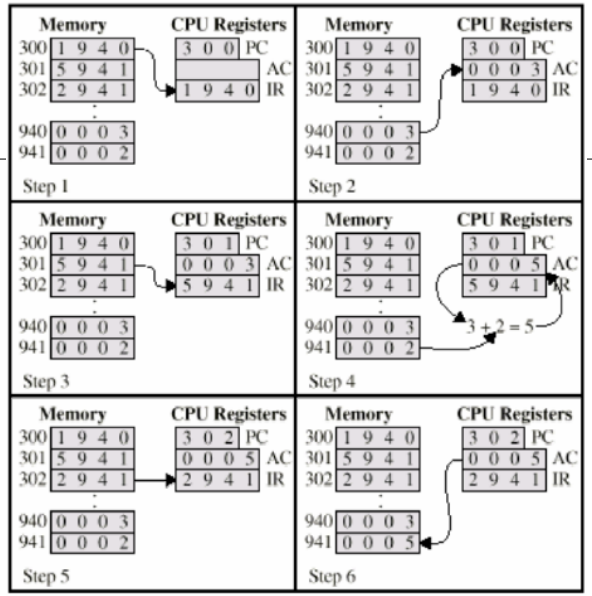
Fetch Cycle

เริ่มทำงานโดยการโหลดคำสั่งเข้าไปในรีจิสเตอร์ IR และส่วนของแอ็ดเดรสของหน่วยความจำจะถูกโหลดเข้าไปใน รีจิสเตอร์ MAR ซึ่งคำสั่งดังกล่าวนี้อาจจะมาจากรีจิสเตอร์ IBR หรืออาจจะมาจากหน่วยความจำหลักที่ถูกโหลดเข้าไปในรีจิสเตอร์ MBR และส่งผ่านไปยังรีจิสเตอร์ IBR อีกทีก็ได้

- การเคลื่อนย้ายข้อมูลระหว่างโปรเซสเซอร์กับหน่วยความจำหลัก (Processor-memory)
- การเคลื่อนย้ายข้อมูลระหว่างโปรเซสเซอร์กับอุปกรณ์อินพุต/เอาต์พุต (Processor-I/O)
- การประมวลผลข้อมูล (Data Processing)
- การควบคุมการทำงาน (Control)

Fetch Cycle

เมื่อคำสั่งถูกโหลดเข้ามาใน รีจิสเตอร์ IR วงรอบเฟ็ตช์จะเริ่มทำงาน โดยแผงวงจรอิเล็กทรอนิกส์ในโปรเซสเซอร์จะทำการอ่านคำสั่งที่ให้ทำงาน และเรียกใช้โดยการส่งสัญญาณไปยังส่วนควบคุม (Control) เพื่อเคลื่อนย้ายข้อมูล หรือทำการคำนวณในส่วน ALU โดยการเรียกใช้คำสั่งนั้นจะทำงานตลอดเวลา และจะหยุดก็ต่อเมื่อเครื่องคอมพิวเตอร์ถูกปิด หรือเกิดข้อผิดพลาดที่ไม่สามารถกู้คืนได้ หรือมีคำสั่งในโปรแกรมให้หยุดการทำงานของเครื่อง (Halt)



คุณลักษณะของหน่วยความจำ

- **ตำแหน่ง** (Location)
- **ความจุ** (Capacity)
- **หน่วยของการถ่ายโอนข้อมูล** (Unit of Transfer)
  - **เวิร์ด** (word)
  - **หน่วยที่ใช้ในการอ้างถึงตำแหน่งข้อมูล**
- **วิธีการเข้าถึงข้อมูล** (Access method)
  - **การเข้าถึงข้อมูลแบบลำดับ** (Sequential access)
  - **การเข้าถึงข้อมูลแบบโดยตรง** (Direct access)
  - **การเข้าถึงข้อมูลแบบสุ่ม** (Random access)

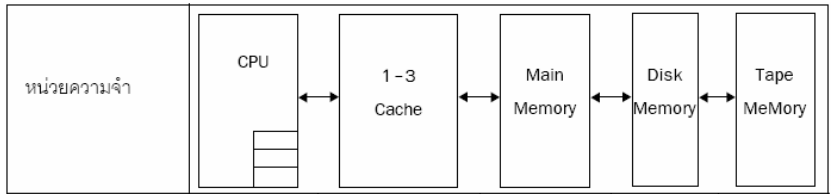
คุณลักษณะของหน่วยความจำ

- **ประสิทธิภาพของหน่วยความจำ** (Performance)

ชื่อ	หน่วย	คำอธิบาย
Access time	วินาที	เวลาที่ใช้ในการเข้าถึงเวิร์ดในหน่วยความจำ
Cycle time	วินาที	เวลาที่ใช้เริ่มนับจากหน่วยความจำอ่านข้อมูลจนถึงจุดเริ่มของคำสั่งถัดไป
Block size	เวิร์ด	จำนวนของเวิร์ดใน 1บล็อกข้อมูล
Transfer Rate	เวิร์ด/วินาที	อัตาการถ่ายโอนข้อมูล
Latency	วินาที	เวลาที่ใช้ในการเข้าถึงเวิร์ดแรกของข้อมูล
Block access time	วินาที	เวลาที่ใช้ในการเข้าถึงบล็อกของเวิร์ดทั้งหมด

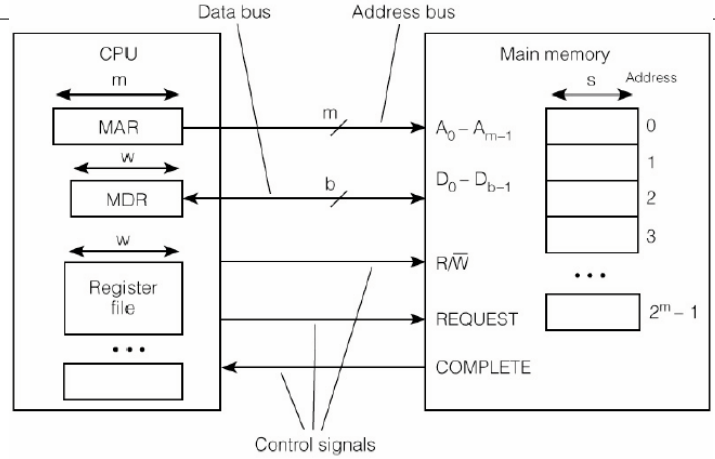
- **ชนิดของสื่อที่ใช้บันทึก** (Physical Type)

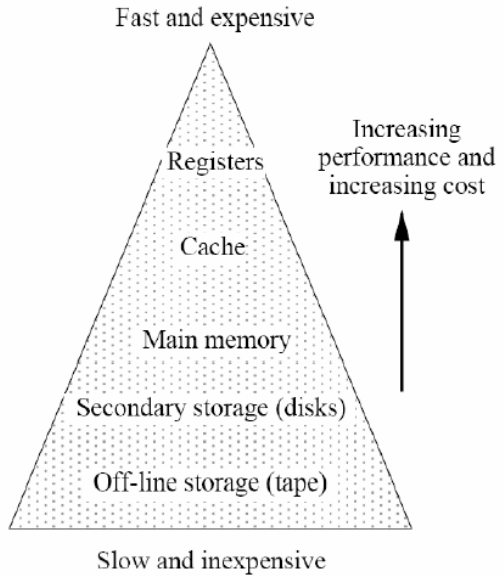
การแบ่งหน่วยความจำออกเป็นลำดับชั้น



หน่วยความจำ					
การเข้าถึงข้อมูล	แบบสุ่ม	แบบสุ่ม	แบบสุ่ม	โดยตรง	แบบลำดับ
ความจุ (ไบต์)	64-1024	8KB-4MB	64MB-2GB	10-200GB	1TB
เวลาแฝง	4-10ns	0.4-20ns	10-50ns	10ms	10ms-10s
ขนาดของเวิร์ด	1 word	16 words	16 words	4 KB	4 KB
อัตราการโอนข้อมูล	ตามสัญญาณนาฬิการะบบ	สัญญาณนาฬิกา -80MB/s	10-400 MB/s	50 MB-s	1 MB/s
ราคาต่อ MB	สูง	450 บาท	10 บาท	0.08 บาท	0.4 บาท

การเชื่อมต่อระหว่างซีพียูกับหน่วยความจำ

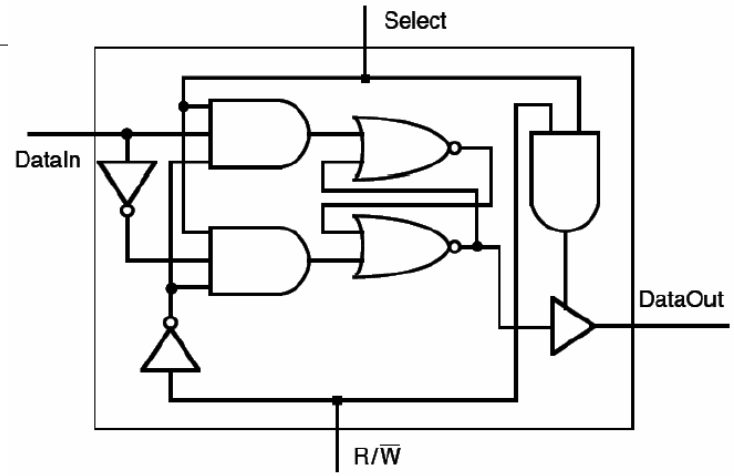




หน่วยความจำ RAM

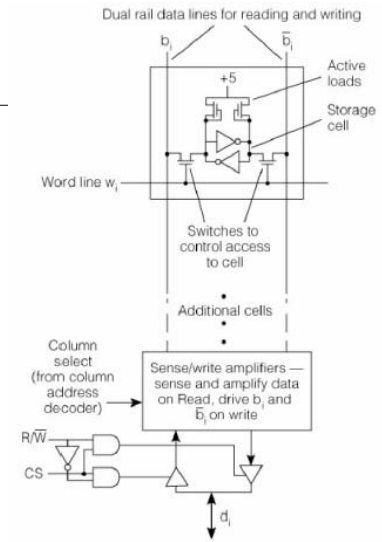
หน่วยความจำ RAM มีการทำงานเหมือนกับเกต D-ฟลิบฟล็อบ (D flip-flop) และมีตัวควบคุมที่จะอนุญาตให้เซลนั้นถูกเลือก, อ่าน, หรือเขียนลงไปได้ ในเซลนั้นจะมีสายสัญญาณนำข้อมูลเข้าและออก และเมื่อเรากล่าวถึงชิปของ RAM เราจะนำเซลของหน่วยความจำที่คล้ายกับในรูปมาใช้ อย่างไรก็ตาม เซลของหน่วยความจำจริง ๆ นั้นไม่จำเป็นต้องเหมือนดังในรูปก็ได้

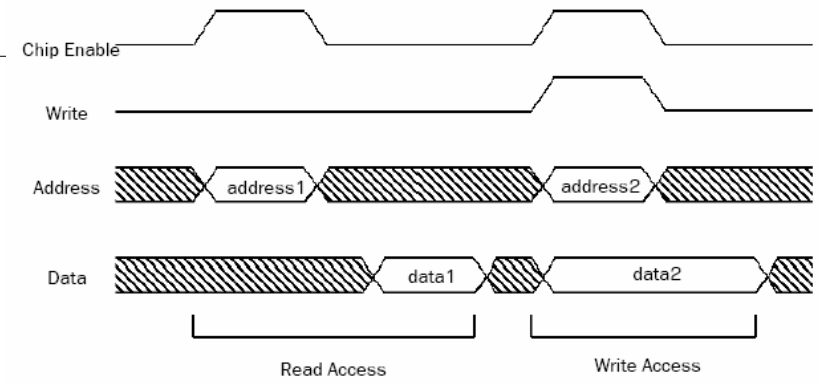
หน่วยความจำ RAM



SRAM

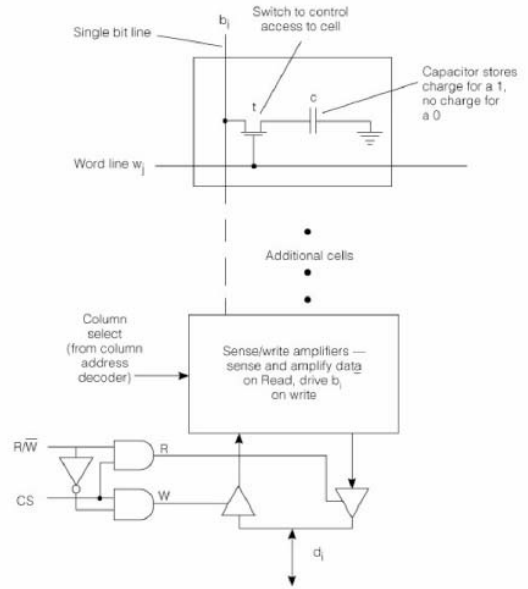
เซลของหน่วยความจำ SRAM นั้นจะประกอบไปด้วยตัวกลับสัญญาณ (Inverter) 2 ตัวที่เชื่อมต่อกันแบบกลับไปกลับมาเป็นวงแหวน เมื่อค่าเซลของหน่วยความจำถูกกำหนดค่า โครงสร้างการเชื่อมต่อดังกล่าวของตัวกลับสัญญาณจะเป็นตัวรักษาค่าที่ถูกบันทึกไว้ เพราะสัญญาณไฟนำเข้าของตัวกลับสัญญาณจะมาจากตัวกลับสัญญาณที่อยู่ตรงข้ามกัน และส่งสัญญาณสลับกันไปมา และด้วยเหตุนี้ SRAM จึงถูกเรียกว่าเป็นหน่วยความจำที่เป็นแบบสถิต (Static RAM) ซึ่งค่าที่บรรจุอยู่ในแรมจะคงที่ตราบเท่าที่ยังมีกระแสไฟฟ้าหล่อเลี้ยงระบบ



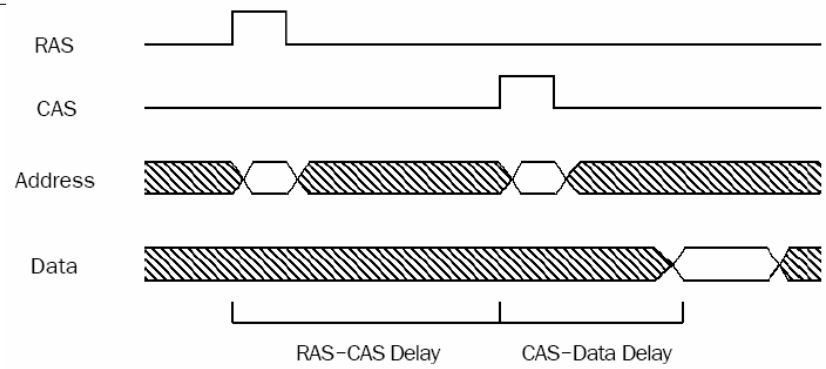


DRAM

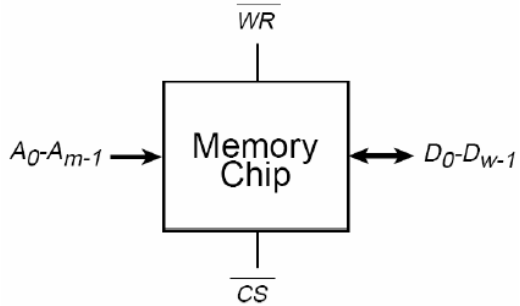
เซลของ DRAM นั้นสามารถแสดงได้ ซึ่งเราจะเห็นว่า DRAM มีการใช้ตัวเก็บประจุ (capacitor) แทนตัวกลับสัญญาณ (inverter) ที่ใช้ในเซลที่เก็บข้อมูล เมื่อสายสัญญาณเวิร์ดถูกจองสิทธิเรียกใช้งาน ตัวเก็บประจุก็จะถูกเชื่อมต่อเข้ากับสายสัญญาณบิต ทำให้ค่าที่เก็บไว้ในเซลถูกอ่านออกมาโดยเราสามารถเช็คค่าที่เก็บไว้ในตัวเก็บประจุได้ว่ามีค่าบิตเป็นเท่าไหร่ หรือเราจะทำการเขียนข้อมูลโดยประจุค่าต่างศักย์ใหม่ให้กับตัวเก็บประจุก็ได้

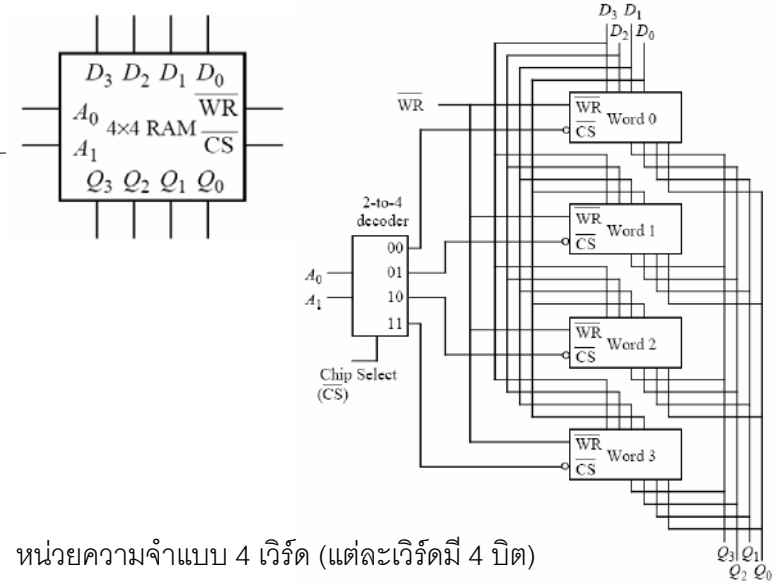


DRAM



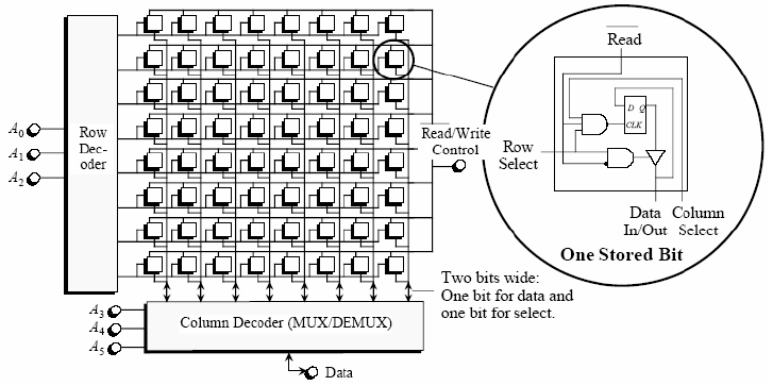
การจัดโครงสร้างภายในชิบ RAM





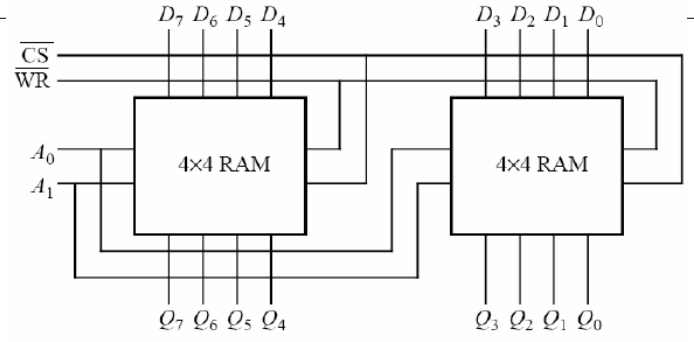
หน่วยความจำแบบ 4 เวิร์ด (แต่ละเวิร์ดมี 4 บิต)

การจัดโครงสร้างภายในชิบ RAM

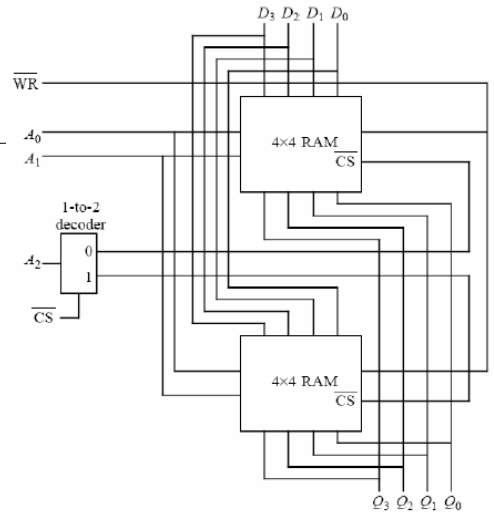


การจัดแบบรูปแบบ 2-1/2D ของ 64 เวิร์ด x 1 บิตแรม

การสร้าง RAM ขนาดใหญจาก RAMขนาดเล็ก



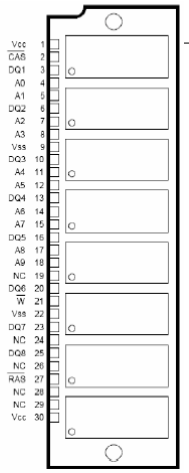
แรมแบบ 4 เวิร์ด x 8 บิต ที่เกิดจากแรม 4 บิต x 4 บิต เชื่อมต่อกัน



RAMแบบ 8 เวิร์ด x 4 บิต ที่เกิดจากRAM 4 บิต x 4 บิต จำนวน 2 ตัวเชื่อมต่อกัน

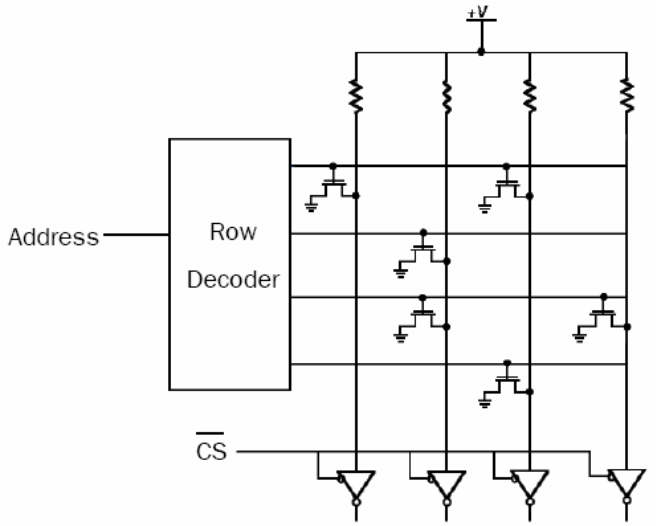
โมดูลของRAMที่มีขายตามท้องตลาด

PIN NOMENCLATURE	
A0-A9	Address Inputs
$\overline{CAS}$	Column-Address Strobe
DQ1-DQ8	Data In/Data Out
NC	No Connection
$\overline{RAS}$	Row-Address Strobe
V_{cc}	5-V Supply
V_{ss}	Ground
$\overline{W}$	Write Enable



หน่วยความจำรอม (ROM : Read Only Memory)

เมื่อมีความต้องการที่จะบรรจุข้อมูลที่ถาวรลงไปในหน่วยความจำ เช่น ค่าติดตั้งต่าง ๆ ของตัวเครื่องหรือเมนบอร์ด หน่วยความจำชนิดอ่านอย่างเดียวหรือ ROM จะถูกนำมาใช้ ROM นั้นเป็นหน่วยความจำแบบไม่ลบเลือน (non-volatile) ซึ่งจะคงค่าของข้อมูลถึงแม้ว่าไม่มีกระแสไฟฟ้าหล่อเลี้ยงก็ตาม เรานิยมใช้ ROM เพื่อเก็บข้อมูลที่ไม่ต้องการการปรับปรุงข้อมูลมากนัก เช่น ใช้สำหรับเก็บค่าต่างๆที่ใช้ในการควบคุมเครื่องยนต์ หรือใช้เก็บบันทึกค่าการทำงานของเครื่องยนต์



ROM Chip แบบ 2-D CMOS

ประเภทของหน่วยความจำรอม

- **PROM** (Programmable Read Only Memory)
- **EPROM** (Erasable Programmable Read Only Memory)
- **Flash EPROM**
- **EEPROM** (Electrically Erasable PROM)

คุณสมบัติรอมประเภทต่าง ๆ

ชนิดของROM	ราคา	เขียนข้อมูลลงได้	เวลาที่ใช้ในการเขียน	เวลาที่ใช้ในการลบ
Mask-programmed ROM	ถูกมาก	เขียนที่โรงงานเท่านั้น	หลายสัปดาห์	ไม่สามารถลบได้
PROM	ถูก	ครั้งเดียวที่ฝั่งผู้ใช้	หลายวินาที	ไม่สามารถลบได้
EPROM	ปานกลาง	หลายครั้ง	หลายวินาที	20 นาที
Flash EPROM	แพง	หลายครั้ง	100 µs	1 วินาที, บล็อกข้อมูลขนาดใหญ่
EEPROM	แพงมาก	หลายครั้ง	100 µs	10 ms, ไบต์