หอสมุดกลาง สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

# บทที่ 2

# แนวความคิดและทฤษฎีที่เกี่ยวข้อง

## 2.1 ข้อมูลและการจัดข้อมูล[1]

ข้อมูล เป็นข้อความจริงเกี่ยวกับสิ่งต่าง ๆ ทั้งที่สามารถจับต้องได้ และสามารถสร้างภาพพจน์ได้ ซึ่งมีอยู่ในโลกแห่งความเป็นจริง (real world) อันเป็นแหล่งกำเนิดข้อมูลต่าง ๆ หน่วยของข้อมูลในขอบเขตนี้เรียกว่า เอนทิตี(entities) ในเอนทิตีหนึ่งมีคุณสมบัติ(properties) มากมายประกอบกันเป็น เอนทิตีนั้น เช่น บุคคล เป็น เอนทิตีหนึ่ง มีคุณสมบัติซึ่งประกอบด้วย ชื่อ นามสกุล เพศ อาชีพ ฯลฯ มีคุณสมบัติอย่างน้อย 1 อย่างแสดงให้เราทราบว่าเป็นเอนทิตีใดหรือบุคคลใด ในที่นี้คือ ชื่อ และนามสกุล เป็นต้น กลุ่มของเอนทิตี เรียกว่า เอนทิตี เซ็ท (entity set) ในการนำเอาข้อมูล หรือ เอนทิตี ต่าง ๆ จากโลกแห่งความเป็นจริงมาประมวลเป็นข้อสนเทศ(information) เพื่อใช้ในการบอกเล่า หรือเก็บบันทึกไว้เป็นหลักฐานก็ตาม เราไม่สามารถนำเอาเหตุการณ์ที่เกิดขึ้นจริง ๆ มาแสดงหรือเก็บบันทึกไว้ได้ แต่เราจะใช้สัญญลักษณ์แทนคุณสมบัติต่าง ๆ ของเอนทิตี เรียกว่า คุณลักษณะ(attributes) และให้ค่า(value) แก่คุณลักษณะดังกล่าวนั้น

ในการจัดเก็บข้อมูล ในสื่อบันทึกข้อมูลของคอมพิวเตอร์จะต้องแปลงข้อมูลให้อยู่ในรูปแบบที่คอมพิวเตอร์อ่านได้ นั่นคือ จะจัดเก็บอยู่ในรูปของ แถวบิท (bit string) หรือ แถวอักษร (Characters string) ซึ่งสามารถจัดข้อมูลได้เป็น 2 ลักษณะคือ การจัดข้อมูลทางตรรกภาพ (Logical Organization) และการจัดข้อมูลทางกายภาพ(physical Organization)

2.1.1 การจัดข้อมูลทางตรรกภาพ เป็นการจัดข้อมูลของโปรแกรมเมอร์หรือการจัดข้อมูลที่ใช้ในโปรแกรมประมวลผล(processing program) โดยมีหน่วยข้อมูลที่เล็กที่สุดเรียกว่า เขตข้อมูล(data fields หรือ data items) เป็นกลุ่มของบิท หรืออักษร ที่เล็กที่สุดที่สื่อ

---

[1] James Martin, Computer Data-Base Organization, 2 d ed,(New Jersey : Prentice-Hall, Inc., 1977) , pp. 48-50

ความหมายได้ เขตข้อมูลหนึ่งก็คือคุณลักษณะอย่างหนึ่งของ เอนทิตีนั่นเอง กลุ่มของเขตข้อมูลรวมกันเรียกว่า ระเบียนข้อมูลหรือระเบียนข้อมูลตรรก(Logical record) ซึ่งหมายถึง เอนทิตีนั่นเอง และกลุ่มของระเบียนข้อมูลรวมกันเป็น แฟ้มข้อมูล(data file)

2.1.2 การจัดข้อมูลทางกายภาพ เป็นการจัดเก็บข้อมูลในสื่อบันทึกข้อมูลจริง ๆ เช่น เทป(Tape) จานบันทึกข้อมูล(Disk) เป็นต้น หน่วยของข้อมูลที่เล็กที่สุด ที่ใช้ในการโยกย้ายข้อมูลระหว่างหน่วยความจำหลัก(main memory) กับสื่อบันทึกข้อมูล เรียกว่า ระเบียนข้อมูลกายภาพ(physical records) หรือ บลอค(blocks) ซึ่งทำให้การจัดรูปแบบของระเบียนข้อมูลตรรก แบ่งออกเป็น 3 แบบ[1] คือ

2.1.2.1 ในบลอคประกอบด้วย 1 ระเบียนตรรก เรียกว่า อันบลอค (Unblocked Record)

2.1.2.2 ในบลอคประกอบด้วย 2 ระเบียนตรรกขึ้นไป เรียกว่า บลอค (Blocked Record)

2.1.2.3 ในบลอคประกอบด้วย บางส่วนของระเบียนตรรก เรียกว่า ระเบียนข้อมูลขยาย(Spanned record)

ในการประมวลผลข้อมูล โปรแกรมประมวลผลต้องการข้อมูลครั้งละระเบียนข้อมูลตรรก และโปรแกรมที่ทำหน้าที่ในการปฏิบัติการ อ่าน หรือบันทึกข้อมูล ก็คือ โปรแกรมย่อยของ โอเอส วิธีการนำระเบียนข้อมูลไปจัดในบลอค(blcoking) และนำไปบันทึกในสื่อบันทึกข้อมูล หรืออ่านจากสื่อบันทึกข้อมูลมา และแยกระเบียนข้อมูลออกจากบลอค(deblocking) ส่งไปยังโปรแกรมประมวลผลนั้น จะขึ้นอยู่กับความต้องการในการเข้าถึงข้อมูล(access) ของโปรแกรมเมอร์ ซึ่งจะต้องมีการวางแผนไว้ก่อน เพื่อว่าจะได้เลือกสื่อบันทึกข้อมูล การจัดเรียงระเบียนข้อมูลตรรกในแฟ้มข้อมูล และเลือกใช้วิธีการเข้าถึงข้อมูล(Access Method) ได้อย่างถูกต้องและเหมาะสมต่อไป

---

[1] Katzan, Jr., Computer Data Management and Data Base Technology, 92.

2.2 จานบันทึกข้อมูล(Disk)[1]

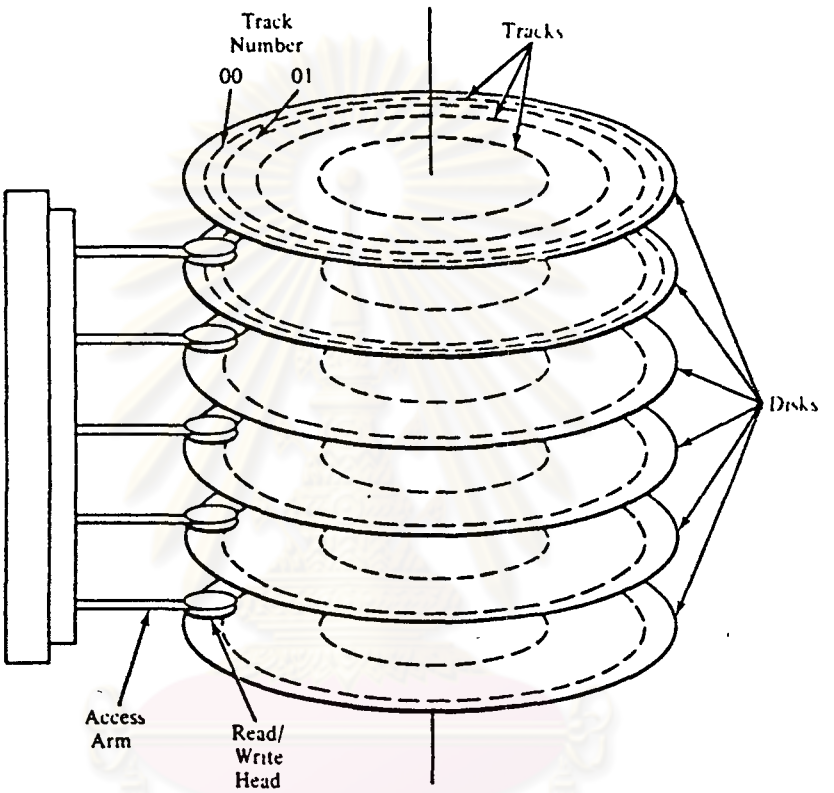
เป็นสื่อที่ใช้ในการบันทึกข้อมูลของระบบคอมพิวเตอร์ ประกอบด้วยแผ่นจานแม่เหล็กตั้งแต่ 1 แผ่นขึ้นไป วางเรียงซ้อนกัน โดยมีแกนเหล็กยึดอยู่ตรงกลาง จานแต่ละแผ่นฉาบด้วยสารแม่เหล็กทั้ง 2 ด้าน เพื่อใช้สำหรับบันทึกข้อมูล(ยกเว้นด้านบนของแผ่นบนสุด และด้านล่างของแผ่นล่างสุด) ระหว่างจานแม่เหล็กแต่ละแผ่นมีหัว อ่าน/เขียน(read/write heads)) ซึ่งเชื่อมอยู่กับแขนอ่าน(access arms) แขนแต่ละแขนจะมีหัว อ่าน/เขียน 2 หัว(สำหรับอ่านและเขียนข้อมูลบนแผ่นจานแม่เหล็ก ด้านล่างของแผ่นบน และด้านบนของแผ่นล่าง ตามลำดับ) ในการอ่านหรือเขียนข้อมูล เครื่องกลไกของจานแม่เหล็กจะหมุนจานแม่เหล็กด้วยความเร็วคงที่ และเลื่อนหัวอ่าน/เขียนไปยังตำแหน่งที่ต้องการบนจานแม่เหล็ก จานแม่เหล็ก 1 ชุด เรียกว่า วอลลุม (disk volume)และถ้าวอลลุมนั้นสามารถเคลื่อนย้ายได้ เรียกว่าแพค(disk pack)

บนพื้นผิวของจานแม่เหล็กทั้ง 2 ด้าน แบ่งเป็นวงกลมหลายวงซ้อนกัน โดยมีจุดศูนย์กลางร่วมกัน วงกลมแต่ละวงเรียกว่า แทรค(track) และในการอ่านหรือเขียนข้อมูในช่วงระยะเวลาหนึ่ง หัวอ่าน/เขียน ทั้งหมดจะเคลื่อนที่ไปด้วยกัน และจะมีหัวอ่าน/เขียนเพียงหัวเดียวเท่านั้นที่ทำงาน ชุดของแทรคทั้งหมดที่อยู่ในตำแหน่งของหัวอ่าน/เขียน ในช่วงเวลาหนึ่ง เรียกว่า ไซเลนเดอร์(Cylender) ดังแสดงในรูปที่ 2.1 ในจานแม่เหล็ก วอลลุมหนึ่งจะประกอบด้วยจำนวนไซเลนเดอร์เท่ากับจำนวนวงกลมบนผิวของจาน ซึ่งจำนวนนี้จะขึ้นอยู่กับรุ่นของจานแม่เหล็ก

ในแต่ละแทรคจะมีเครื่องหมาย(index marker) บอกจุดเริ่มต้น และจุดสุดท้ายของแทรค ข้อมูลบันทึกในแทรคเป็น ไบต์ (bytes)เรียงต่อกัน เรียกว่า บลอค(block) คั่นด้วยช่วงว่างระหว่างบลอค(Interblock gaps) ดังรูปที่ 2.2

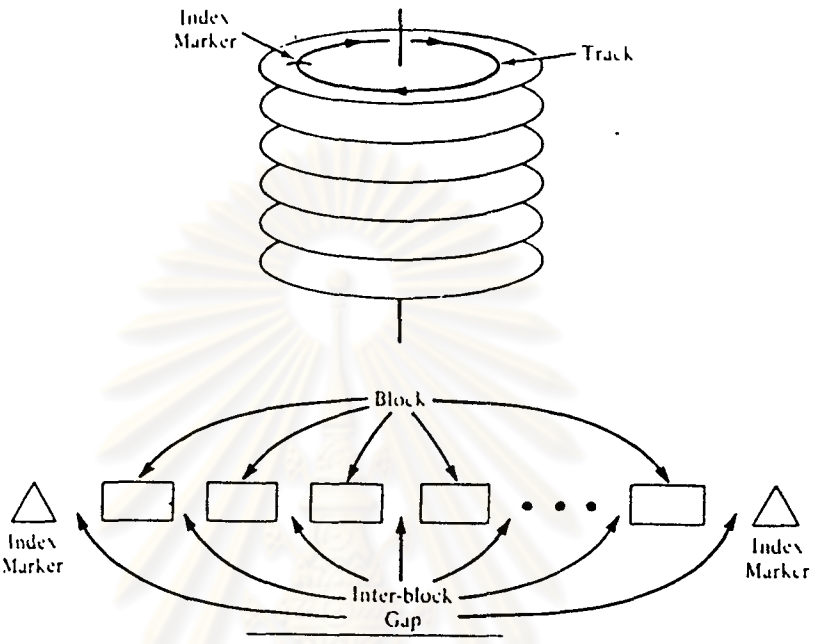
---

[1] Ibid., PP. 41 - 42



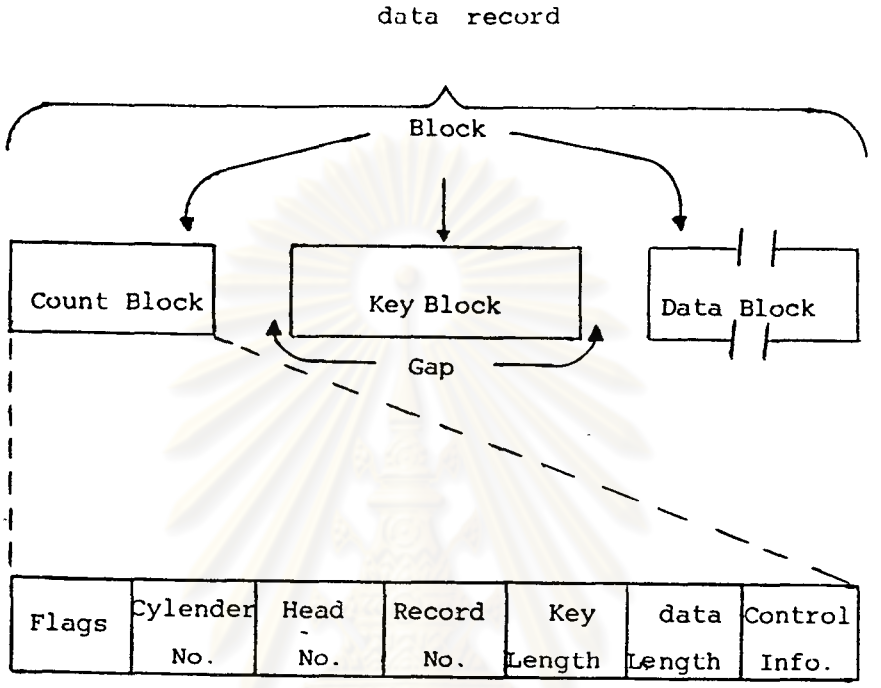

รูปที่ 2.1 จานบันทึกข้อมูล

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย




รูปที่ 2.2 แสดงรูปแบบของแทรค

ข้อมูลที่บันทึกลงในแต่ละแทรค แบ่งออกเป็น 2 ส่วน คือ ส่วนแรกใช้บันทึกสถานะของแทรค(Capacity record or $R_0$)และอีกส่วนใช้บันทึกข้อมูล ในส่วนข้อมูลประกอบด้วย 3 บล็อค คือ เคานบล็อค(count block) คีย์ บลอค (Key block) และ ดาต้า บลอค (data block) รวมเป็นระเบียนข้อมูลกายภาพ(physical record) ดังแสดงในรูปที่ 2.3 ในส่วนของเคานมีความยาว 8 ไบต์ ใช้แสดงตำแหน่งของระเบียนข้อมูลซึ่งประกอบด้วย หมายเลขไซเลนเดอร์ หมายเลขหัวอ่าน หมายเลขระเบียนข้อมูล ความยาวของคีย์(เป็น 0 ถ้าไม่มีคีย์) และความยาวของข้อมูล ถ้าหากระเบียนข้อมูลบันทึกด้วยคีย์ ในพื้นที่คีย์(1-255 ไบต์) จะบันทึกคีย์ของระเบียนข้อมูลไว้ ถ้าไม่มีคีย์จะไม่มีส่วนคีย์ บลอค สำหรับในส่วนข้อมูลจะบันทึกข้อมูลจริง




รูปที่ 2.3 บลอค หรือระเบียนข้อมูลกายภาพ ในจานบันทึกข้อมูล

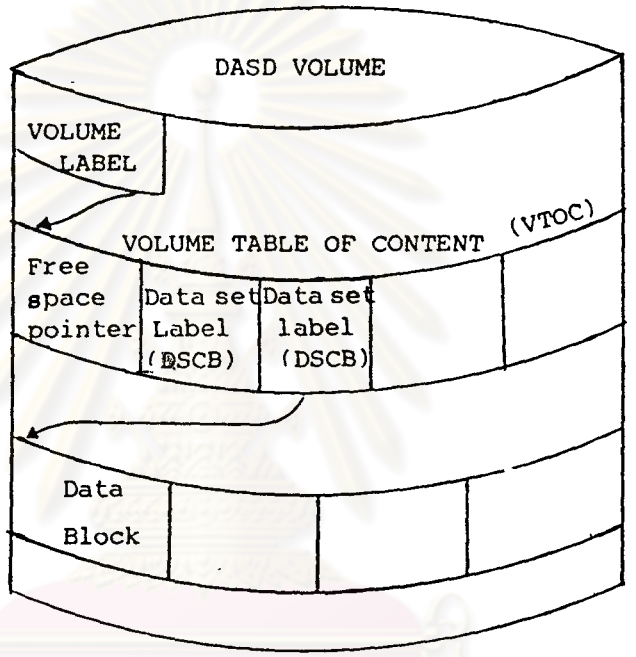
## 2.3 การใช้พื้นที่ในจานบักทึกข้อมูล[1]

ในจานบันทึกข้อมูล วอลลุมหนึ่งจะแบ่งพื้นที่สำหรับบันทึกข้อมูล และข่าวสารที่ใช้ในการควบคุม(Control Information) ออกเป็น 3 ส่วน ดังแสดงในรูปที่ 2.4

2.3.1 วอลลุม เลเบล(Volume label) อยู่ในแทรค 0 ไซเลนเดอร์ 0 ของจานบันทึกข้อมูล(ต่อจาก IPL Record) ระเบียนวอลลุม เลเบล นี้มีความยาว 80 ไบต์ มีรายละเอียดดังแสดงในรูปที่ 2.5 ในวอลลุมหนึ่งอาจมีวอลลุม เลเบลได้ถึง 7 วอลลุม เลเบล

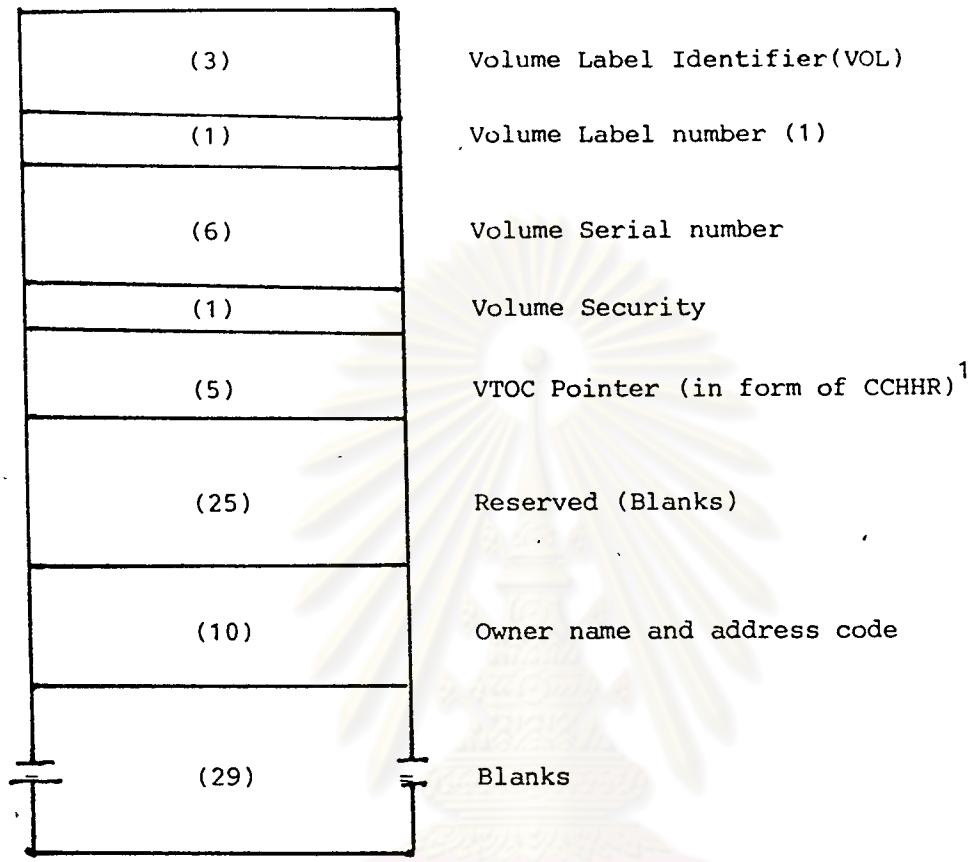
---

[1] IBM, Fundamentals of OS/VS Concepts and Facilities, (Bangkok : IBM co., Ltd, (Thailand)), P. 7 - 14




รูปที่ 2.4 ข้อมูลและข่าวสารควบคุม ในจานบันทึกข้อมูล

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย




รูปที่ 2.5 วอลลุม เลเบล (Initial Volume Label)

2.3.2 วีทอค (VTOC หรือ Volume Table of Content) เป็นที่เก็บข่าวสารเกี่ยวกับข้อมูลที่บันทึกในจานบันทึกข้อมูล ข้อมูลทุกแฟ้มข้อมูลที่บันทึกในจานบันทึกข้อมูล จะเก็บข่าวสารเกี่ยวกับ ชื่อแฟ้มข้อมูล ที่อยู่(address) และการจัดองค์การข้อมูลไว้ใน วีทอค ซึ่งเรียกว่า ดีเอสซีบี (DSCB ย่อมาจาก Data Set Control Blocks) เราเรียกแฟ้มข้อมูลอีกชื่อหนึ่งว่า ดาต้า เซ็ท [1] (Data Set) ทุกดาต้า เซ็ท ต้องมี ดีเอสซีบี และ ดีเอสซีบี จะกำหนดการใช้พื้นที่ในจานบันทึกข้อมูล

---

[1]data set มีความหมายรวมถึง แฟ้มข้อมูล และโปรแกรมต่าง ๆ ที่อยู่ในไลบรารี (library) ด้วย

2.3.3 พื้นที่ที่เหลือจากการบันทึกวอลลุม เลเบล และวีทอค เป็นพื้นที่ว่าง และพื้นที่ที่ใช้บันทึก ดาต้า เซ็ท ไปแล้ว รายละเอียดของการใช้พื้นที่ในจานบันทึกข้อมูลจะบันทึกอยู่ในวีทอค

## 2.4 การประมวลผลดาต้า เซ็ท

ในการประมวลผลดาต้า เซ็ท ข่าวสารเกี่ยวกับ ดาต้า เซ็ท จะถูกรวบรวมไว้ที่หน่วยความจำหลัก ในพื้นที่ที่เรียกว่า ดีซีบี (DCB ย่อมาจาก Data Control Block ) ซึ่งเป็นบล็อคข่าวสารที่ใช้ในการควบคุมการนำเข้า/นำออก( Input/Output) ในการเขียนโปรแกรมด้วยภาษาระดับสูง(high level language ) เราไม่จำเป็นต้องยุ่งเกี่ยวกับ ดีซีบี โดยตรง แต่ถ้าใช้แอสแซมบลีแล้ว จะต้องเตรียมข่าวสารสำหรับ ดีซีบี ซึ่งทำได้ 3 ทาง คือ

- เขียนด้วย ดีซีบี แมโคร (DCB macro ) ในโปรแกรม
- เขียนใน ดีซีบี โอเปอร์แรน(DCB operand ) ในบัตร DD[1] ของเจซีแอล
- ได้ข่าวสารจาก ดีเอสซีบี ใน วีทอค

การอ่านข้อมูลเพื่อให้ได้ระเบียนข้อมูลที่ต้องการนั้น จะมีการสร้างโปรแกรมมาตรฐานไว้ซึ่งเป็นโปรแกรมย่อยของ โอเอส เรียกว่า โปรแกรมการเข้าถึงข้อมูล (Access Method Routine) สำหรับระบบ โอเอส/วีเอส 1 . มีหลายโปรแกรมให้เลือกใช้ ตามลักษณะการจัดองค์การข้อมูล และเทคนิคการเข้าถึงข้อมูล ได้แก่ แซม (SAM ย่อมาจาก Sequential Access Method) แดม (DAM ย่อมาจาก Direct Access Method) ไอ แซม.(ISAM ย่อมาจาก Index Sequential Access Method ) และ . วีแซม (VSAM ย่อมาจาก Virtual Storage Access Method ) เป็นต้น จากการศึกษาครั้งนี้จะกล่าวถึงเฉพาะ วีแซม เท่านั้น ซึ่งจะกล่าวถึงในหัวข้อต่อไป

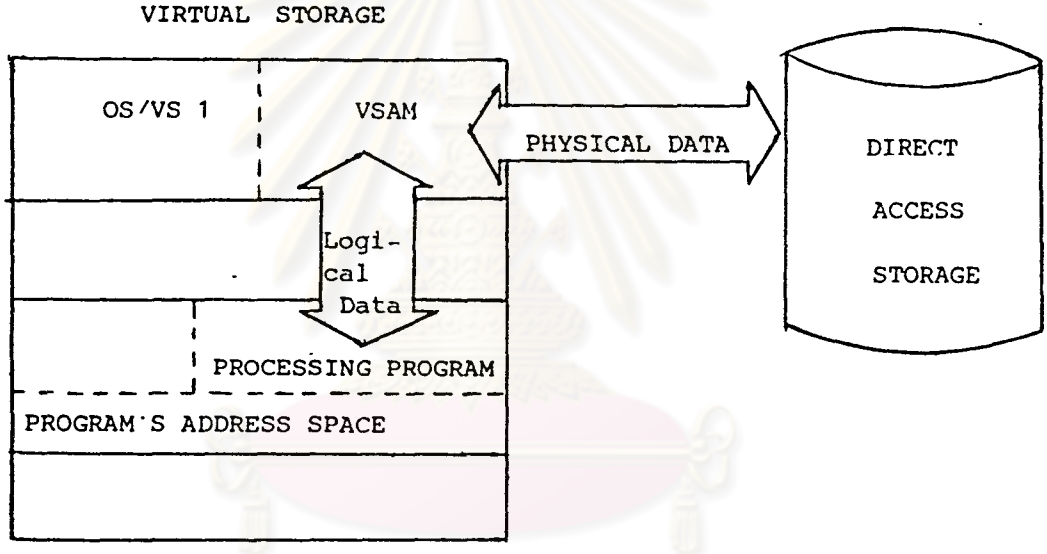
---

[1] บัตร DD ใน เจซีแอล ใช้บอกข่าวสารที่ใช้ควบคุม ดาต้า เซ็ท เพื่อทำให้ ดีซีบี ได้ข่าวสารสมบูรณ์

013238

2.5 วีแซม

เป็นวิธีการเข้าถึงข้อมูล(access method) อย่างหนึ่ง ซึ่งวิ่งภายใต้โปรแกรมควบคุมระบบ ที่ใช้หน่วยความจำเทียม[1] (Virtual Storage Operating System) เช่น คอส/วีเอส หรือ โอเอส/วีเอส 1 เป็นต้น โปรแกรมของ วีแซม จะอยู่ในหน่วยความจำเทียม เฉพาะส่วนที่ใช้งานในช่วงเวลาหนึ่งเท่านั้นที่ใช้หน่วยความจำจริง(real storage) ตำแหน่งของ วีแซม และ โอเอส/วีเอส 1 แสดงในรูปที่ 2.6




รูปที่ 2.6 ตำแหน่งของ วีแซม และ โอเอส/วีเอส 1

[1] IBM, VSAM - Using Access Method Service (OS/VS, OS/MVS), ISP, (Bangkok : IBM , co., ltd. (Thailand)), P. 2 - 1

2.5.1 ซีไอ และ ซีเอ (CI and CA) วีแซม ดาต้า เซ็ท มีหน่วยของข้อมูลเรียกว่า ซีไอ (CI ย่อมาจาก Control Interval) เป็นหน่วยซึ่ง วีแซม ใช้ในการอ่านหรือบันทึกข้อมูล 1 ครั้ง ขนาดของ ซีไอ ผู้ใช้เป็นผู้กำหนด ซึ่งมีความยาวอยู่ระหว่าง .5 ถึง 32 เค ไบท์ (K bytes) โดยเริ่มจาก 512 ไบต์ และเพิ่มขึ้นเป็นทวีคูณของ 512 หรือ 2048 ไบต์ ข้อมูลที่บันทึกใน ซี ไอ แบ่งออกเป็น 2 ส่วน คือ

ส่วนที่ 1 ใช้บันทึกระเบียนข้อมูลของ ดาต้า เซ็ท ซึ่งบันทึกจากซ้ายไปขวา

ส่วนที่ 2 จะบันทึกจากขวาของ ซีไอ ไปทางซ้าย ซึ่งเป็นข้อมูลที่บรรยายถึงระเบียนข้อมูลแต่ละระเบียเข้อมูล ที่บันทึกใน ซีไอ นั้น โดยทางขวาสุดจะใช้บันทึกตำแหน่งของที่ว่าง(freespace) ใน ซีไอ เรียกว่า ซีไอดีเอฟ (CIDF หรือ Control Interval Difinition Field) ถัดจาก ซีไอดีเอฟ มาทางซ้ายเป็น อาดีเอฟ (RDF หรือ Record Difinition Field) ซึ่งเป็นข้อมูลเกี่ยวกับความยาว และตำแหน่งของระเบียนข้อมูลตรรก ที่อยู่ใน ซีไอ นั้น โดยปกติจำนวน อาดีเอฟ จะเท่ากับจำนวนของระเบียนข้อมูลตรรกที่อยู่ใน ซีไอ นั้น ยกเว้นในกรณีที่มีระเบียนข้อมูลที่มีความยาวเท่ากันตั้งแต่ 2 ระเบียนขึ้นไป อยู่ติดกัน กลุ่มของระเบียนข้อมูลดังกล่าว ต้องการ อาดีเอฟ เพียง 2 อาดีเอฟ เท่านั้น

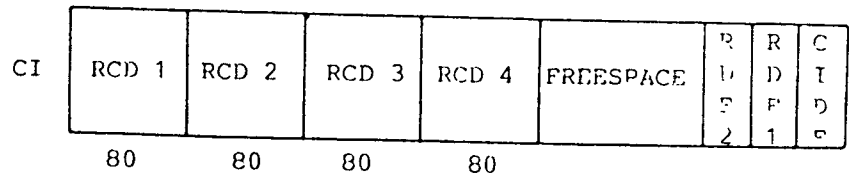
ซีไอ ตั้งแต่ 1 ซีไอ ขึ้นไปรวมกันเป็น ซีเอ (CA หรือ Control Area) ขนาดของ ซีเอ ในดาต้า เซ็ท จะเท่ากันทุก ซีไอ และมีขนาดได้ไม่เกิน 1 ไซเลนเดอร์

2.5.2 รูปแบบระเบียนข้อมูล(record format)[1] แบ่งออกเป็น 3 รูปแบบดังนี้คือ

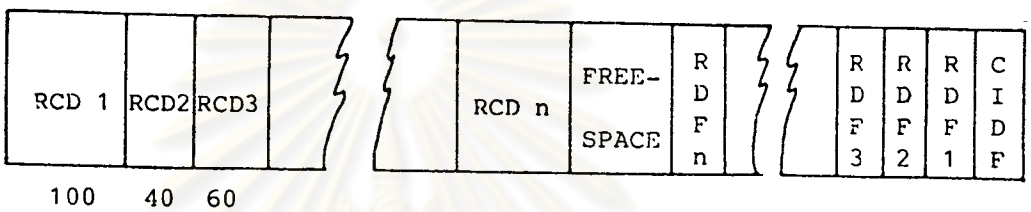
2.5.2.1 ระเบียนข้อมูลที่มีความยาวคงที่(fixed length format) การบันทึกระเบียนข้อมูลชนิดนี้ลงใน ซีไอ จะใช้ อาดีเอฟ เพียง 2 ตัวเท่านั้น โดยจะบอกรายละเอียดเกี่ยวกับความยาวของระเบียนข้อมูล และจำนวนของระเบียนข้อมูลใน ซีไอ ดังแสดงในรูปที่ 2.7 ก.

---

[1] Ibid., PP. 3 - 5 - 3 - 6




ก) ระเบียนข้อมูลที่มีความยาวคงที่

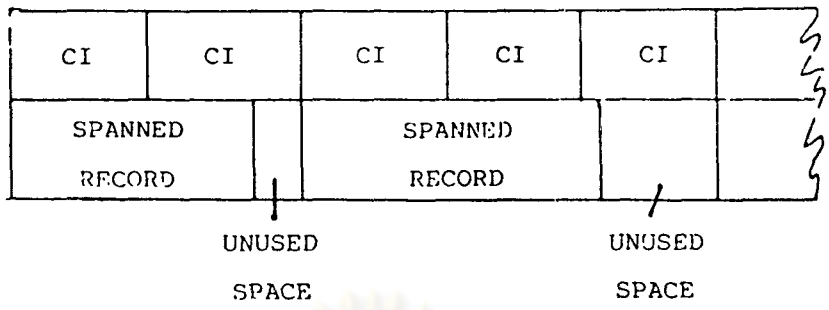



ข) ระเบียนข้อมูลที่มีความยาวไม่คงที่

รูปที่ 2.7 รูปแบบของระเบียนข้อมูลที่บันทึกใน วีแซม ดาต้า เซ็ท

2.5.2.2 ระเบียนข้อมูลที่มีความยาวไม่คงที่(Variable length format) ระเบียนข้อมูลจะบันทึกใน ซีไอ ดังแสดงในรูปที่ 2.7 ข. ซึ่งมี อาดีเอฟ เท่ากับจำนวนระเบียนข้อมูลที่บันทึกใน ซีไอ ยกเว้นในกรณีที่มีระเบียนข้อมูลความยาวเท่ากัน ตั้งแต่ 3 ระเบียนขึ้นไปอยู่ติดกัน จะใช้เพียง 2 อาดีเอฟ เท่านั้น

2.5.2.3 ระเบียนข้อมูลขยาย(Spanned record format) เป็นรูปแบบข้อมูลที่ระเบียนข้อมูลตรรกยาวกว่า ซีไอ ระเบียนข้อมูลประเภทนี้มีความยาวได้ไม่เกินขนาดของ ซีเอ ลบด้วยเขตควบคุม(RDF+CIDF) และคีย์ของระเบียนข้อมูลจะต้องอยู่ใน ซีไอ แรก ระเบียนข้อมูลใหม่จะเริ่มต้นที่จุดเริ่มต้นของ ซีไอ เสมอ พื้นที่ที่เหลือใน ซีไอ สุดท้ายของระเบียนข้อมูลจะทิ้งว่างไว้เฉย ๆ ดังแสดงในรูปที่ 2.8




รูปที่ 2.8 รูปแบบของระเบียนข้อมูลขยาย

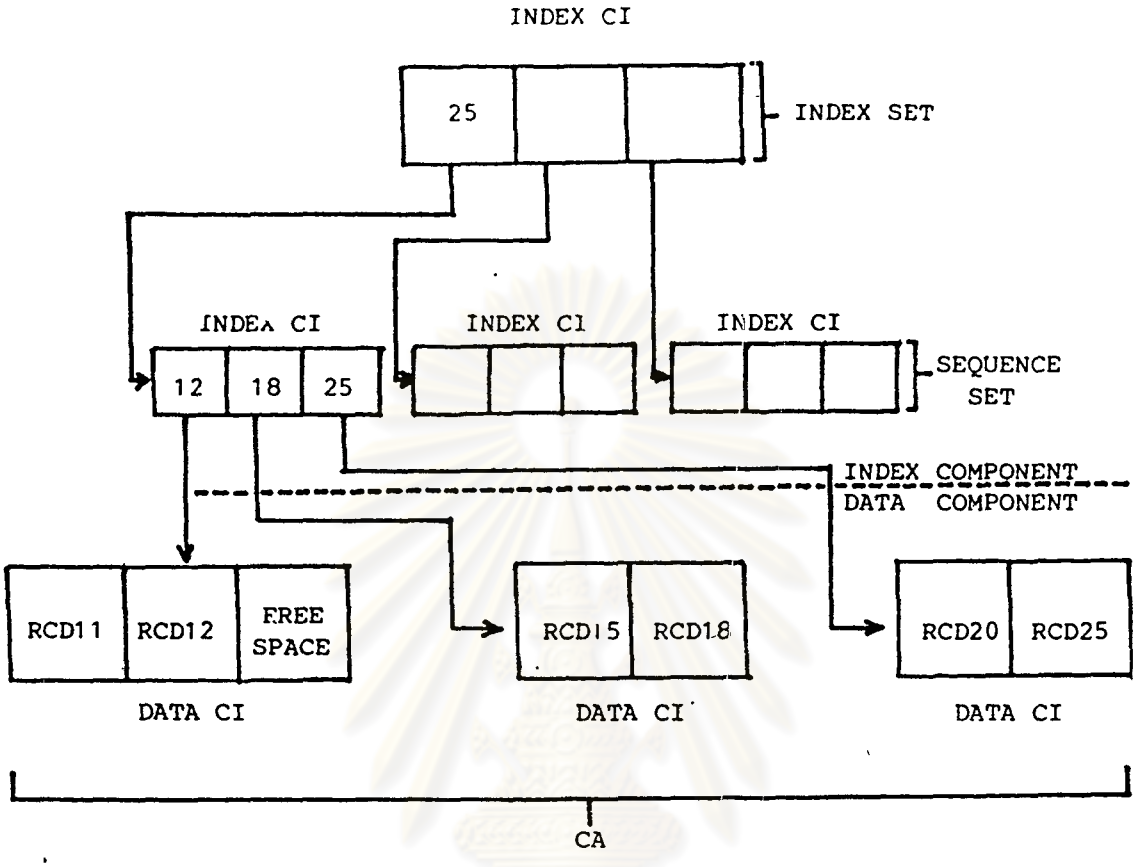
2.5.3 การจัดองค์การดาต้า เซ็ท (Data Set Organization)[1] วีแซม ดาต้า เซ็ท แบ่งการจัดองค์การออกเป็น 3 ชนิด คือ

2.5.3.1 เคเอสดีเอส (KSDS ย่อมาจาก Key Sequenced Data Set) ข้อมูลจะบันทึกลงใน ดาต้า เซ็ท เรียงตามลำดับคีย์ โดยแยกการบันทึก ดาต้า เซ็ท ออกเป็น 2 ส่วน คือ ส่วนข้อมูลและส่วนดัชนี(Index) โดยใน เอนทรี(entry) ของดัชนี ประกอบด้วยค่าสูงสุดของคีย์ใน ซีไอ ดาต้า เซ็ท และตัวชี้(pointer) ไปยังจุดเริ่มต้นของข้อมูลใน ซี ไอ นั้น ใน เอนทรีของดัชนีจะชี้ไปยังข้อมูล 1 ซีไอ และ 1 ซีไอ ของดัชนีจะชี้ไปยังข้อมูล 1 ซีเอ ในส่วนดัชนีเป็นดัชนีหลายระดับ(multi-level indixes) เพื่อลดเวลาในการค้นหาข้อมูล ระดับต่ำสุดของดัชนี เรียกว่า ซี เควน เซ็ท(Sequence Set) และดัชนีในระดับสูงขึ้นไป เรียกว่า อินเดค เซ็ท(Index Set) ดังรูปที่ 2.9

ในส่วนของข้อมูลเราสามารถกำหนดพื้นที่ว่าง(freespace) ใน ซีไอ หรือ ซีเอ ได้ โดยกำหนดให้เป็นพื้นที่ว่าง เป็นร้อยละของ ซีไอ หรือ ซีเอ ทั้งนี้เพื่อประโยชน์ในการขยายขนาดของระเบียนข้อมูลหรือแทรกระเบียนข้อมูลเข้าไปใน ดาต้า เซ็ท ได้ ในภายหลัง ในดัชนีจะมี เอนทรี ชี้มายัง ซีไอ ที่เป็นพื้นที่ว่างด้วย

---

[1] Ibid., PP. 2-3 - 2-5




รูปที่ 2.9 ส่วนประกอบของ เค เอส ดี เอส

ในการบันทึกข้อมูลลงใน ดาต้า เซ็ท ครั้งแรก ระเบียนข้อมูลจะเรียงลำดับในสื่อบันทึกข้อมูลตามคีย์ ต่อมาในภายหลัง เมื่อมีการแทรกข้อมูลลงใน ซีไอ ซึ่งมีที่ว่างไม่พอ จะเกิดการแบ่ง ซีไอ (CI split)[1] ขึ้น ทำให้ลำดับของระเบียนข้อมูลไม่เป็นไปตามคีย์ แต่เอนทรีในดัชนียังคงเรียงลำดับตามคีย์อยู่เหมือนเดิม โดย วีแซม จะทำดัชนีให้ทันสมัย(update) ถ้าหากใน ซีเอ ที่ต้องการจะแทรกระเบียนข้อมูล ไม่มี ซีไอ ว่างเหลืออยู่ วีแซม จะแบ่ง ซีเอ (CA split)[2]

---

[1] Ibid., P. 3-14

[2] Ibid., P. 3-17

โดยนำเอาครึ่งหนึ่งของ ซีเอ ไปบันทึกใน ซีเอ ที่ว่าง(CA freespace) ทำให้ ซีเอ เดิม มีที่ว่าง สามารถนำระเบียนข้อมูลใหม่แทรกเข้าไปได้

เคเอสดีเอส . สามารถเข้าถึงข้อมูลได้ทั้งแบบโดยตรง(direct access) โดยการใช้ อาบีเอ (RBA ย่อมาจาก Relative Byte Address)[1] หรือโดยการใช้คีย์ และเข้าถึงแบบเรียงลำดับตามคีย์ได้

2.5.3.2 อีเอสดีเอส (ESDS ย่อมาจาก Entry-Sequenced Data Set) ในการจัดองค์การข้อมูลแบบนี้ ระเบียนข้อมูลจะบันทึกตามลำดับที่ระเบียนข้อมูลถูกบันทึกเข้าไปใน ดาต้า เซ็ท ระเบียนข้อมูลใหม่จะบันทึกต่อท้ายระเบียนข้อมูลเดิมที่มีอยู่ การแทรกข้อมูลไม่สามารถทำได้ การประมวลผลทำได้แบบเรียงลำดับ และแบบโดยตรง โดยใช้ อาบีเอ

2.5.3.3 อาอาร์ดีเอส . (RRDS ย่อมาจาก Relative-Record Data Set) วีแซม จะแบ่ง ดาต้า เซ็ท ออกเป็นช่อง(Slot) มีขนาดเท่ากับระเบียนข้อมูลตรรก แต่ละช่องมีหมายเลขประจำอยู่ เริ่มจากหมายเลขหนึ่ง ถึงหมายเลขสูงสุดของระเบียนข้อมูลที่จะบันทึกในดาต้า เซ็ท ข้อมูลจะบันทึกลงในดาต้า เซ็ท ตามหมายเลขที่กำหนดให้ ความยาวของระเบียนข้อมูลเปลี่ยนแปลงไม่ได้ แต่สามารถแทรกระเบียนข้อมูลเข้าไปในช่องว่างของดาต้า เซ็ทได้ การเข้าถึงข้อมูลทำได้ทั้งแบบเรียงลำดับ และแบบโดยตรง โดยใช้หมายเลขช่องหรือหมายเลขระเบียนข้อมูล

2.5.4 วีแซม แคทตาล็อค (VSAM catalog) ดาต้า เซ็ท ที่บันทึกใน แคส วอลลุม ต้องใช้พื้นที่ในวอลลุมนั้น หน่วยของพื้นที่ซึ่ง วีแซม ควบคุม สำหรับการบันทึกข้อมูล เรียกว่า ดาต้า สเพซ(data space)[2] และดาต้า สเพซ ของ วีแซม จะใช้โดย วีแซม เท่านั้น วีแซม ดาต้า เซ็ท จะต้องบันทึกใน วีแซม ดาต้า สเพซ เสมอ วีแซม จะเก็บข่าวสาร ซึ่งควบคุมการใช้ ดาต้า สเพซ และดาต้า เซ็ท ของวอลลุม ไว้ที่ศูนย์กลางการควบคุม เรียกว่า วีแซม แคทตาล็อค ซึ่งใน วีแซม แคทตาล็อค จะบันทึก เอนทรี ของทุก วีแซม ดาต้า เซ็ท และ ดาต้า สเพซ ไว้ การกำหนด

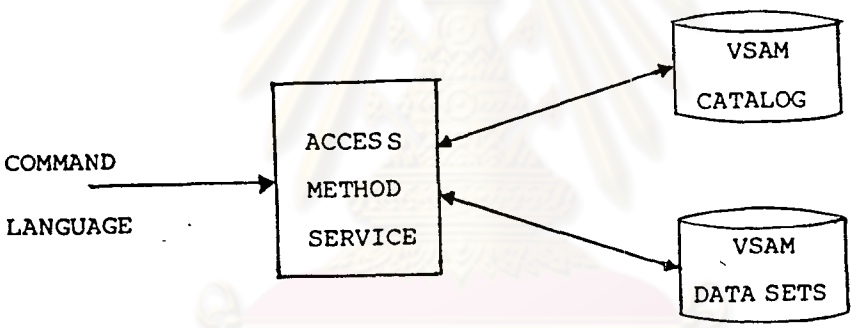
---

[1] Ibid., P. 2-6

[2] Ibid., P. 3-37

คาต้า สเพช จะใช้คำสั่ง DEFINE SPACE [1] ซึ่งเป็นคำสั่งหนึ่งของ เอเอ็มเอส (AMS ย่อมาจาก Access Method Services) เอเอ็มเอส จะไปค้นหาสเพช ใน วีทอค เมื่อพบแล้ววีแซมจะนำเอา ดี เอส ซี บี ไปไว้ใน วีทอค และบันทึกว่า สเพช นี้เป็นของ วีแซม วีแซม จะสร้าง เอนทรี สำหรับ คาต้า สเพช ใน วีแซม แคทตาล็อค และเมื่อต้องการสร้าง วีแซม คาต้า เซ็ท วีแซม จะไปดูสเพช ใน วีแซม แคทตาล็อค

2.5.5 เอเอ็มเอส [2] เป็นโปรแกรมอำนวยประโยชน์ ซึ่งเรียกใช้ภายใต้ ชื่อ "IDCAMS" ใช้ในการสร้าง วีแซม แคทตาลอค กำหนด คาต้า สเพช และ วีแซม คาต้า เซ็ท รวมทั้งการโหลด คาต้า เซ็ท การเรียกใช้ เอเอ็มเอส ทำได้โดยการใช้คำสั่ง(command) ความสัมพันธ์ระหว่าง เอเอ็มเอส และ วีแซม คาต้า เซ็ท ดังแสดงในรูปที่ 2.10




รูปที่ 2.10 ความสัมพันธ์ระหว่าง เอเอ็มเอส และ วีแซม คาต้า เซ็ท

---

[1] Ibid., P. 3-42

[2] Ibid., P. 4-1

ในการสร้างระบบการจัดการข้อมูล จะใช้แฟ้มข้อมูลแบบ วีแซม ทั้งนี้เพราะ วีแซม สามารถใช้ระเบียนข้อมูลชนิดที่มีความยาวไม่คงที่ได้ และการเข้าถึงข้อมูลสามารถทำได้ทั้งแบบโดยตรง และแบบเรียงลำดับ นอกจากนี้ วีแซม ไม่มีพื้นที่ โอเวอร์ โฟล (overflow area) ในการแทรกข้อมูล หากมีที่ว่างไม่พอใน ซีไอ และ ซีเอ วีแซม สามารถแบ่ง ซีไอ และ ซีเอ. ได้ ดังกล่าวมาแล้วข้างต้น ทำให้การจัดองค์การข้อมูลใหม่ทำได้สะดวกขึ้น

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย