บทที่ 2

การประมวลผลข้อมูล

จุดประสงค์ของบทนี้คือ การศึกษาถึงการนำระบบคอมพิวเตอร์มาใช้ในวงการธุรกิจ เอกชน องค์การ พร้อมทั้งแนวคิดการประมวลผล การพัฒนาการประมวลผล และปัญหา

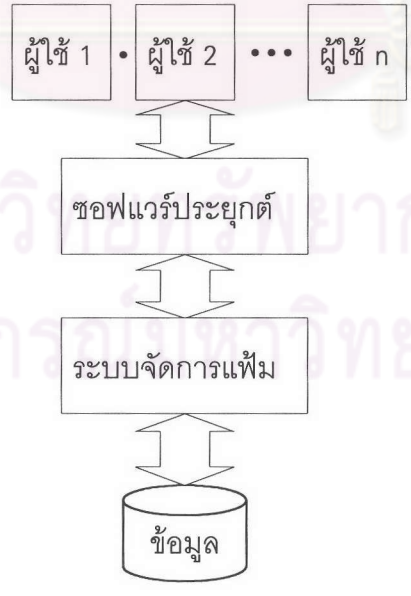
จากบทนี้จะได้ทราบถึงวิวัฒนาการของการประมวลผลข้อมูลด้วยคอมพิวเตอร์อันได้แก่

1. องค์ประกอบของการประมวลผลข้อมูล
2. ข้อมูลดำเนินการและสายงาน
3. รูปแบบของระบบการประมวลผล

2.1 องค์ประกอบของการประมวลผลข้อมูล (Components of Data Processing) (Date, 1981)

โดยพื้นฐานแล้วการประมวลผลจะต้องมีองค์ประกอบ ที่ต้องเกี่ยวข้องกับการที่ทำให้ธุรกิจดำเนินไปได้รวดเร็ว และถูกต้อง

1. ข้อมูล (Data)
2. ฮาร์ดแวร์ (Hardware)
3. ซอฟแวร์ (Software)
4 ผู้ใช้ (User)



รูปที่ 2.1 ความสัมพันธ์ระหว่างองค์ประกอบต่างๆ ของการประมวลผล

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

2.1.1 ข้อมูล (Data)

ในส่วนของข้อมูลจะถูกเก็บในแต่ละส่วนรวมกันทั้งหมดไว้ในระบบ และใช้ข้อมูลร่วมกัน ข้อมูลที่รวมกันไว้จะเป็นหนึ่งเดียว เป็นข้อมูลที่ชัดเจน ตัดส่วนที่ซ้ำซ้อนออก เช่นข้อมูลพนักงานจะเก็บชื่อ, ที่อยู่, แผนก, เงินเดือน อื่นๆ ส่วนข้อมูลในการฝึกอบรมเป็นตัวแสดงว่าพนักงานได้ไปเข้าฝึกอบรม สมมติว่ามีคำสั่งให้พนักงานเข้าอบรม ที่ศูนย์ฝึกจะต้องรู้ชื่อแผนกที่ทำงานของพนักงานแต่ละคน ข่าวสารดังกล่าวนี้ในข้อมูลของฝ่ายฝึกอบรมก็ไม่จำเป็นต้องบันทึกซ้ำอีก เนื่องจากอ้างอิงถึงข้อมูลพนักงานได้อยู่แล้วเป็นต้น

การใช้ข้อมูลร่วมกันหมายถึง แต่ละส่วนของข้อมูลในแฟ้มข้อมูลอาจจะให้ผู้ใช้ที่ที่อยู่ต่างหน่วยงานกันใช้ข้อมูลร่วมกัน ซึ่งวัตถุประสงค์การเข้าใช้อาจต่างกัน (ดังเช่นตัวอย่างในข้อมูลพนักงานกับการฝึกอบรมข้างต้น โดยปกติผู้ใช้จะเกี่ยวข้องกับข้อมูลเพียงบางส่วนเท่านั้น

2.1.2 ฮาร์ดแวร์ (Hardware)

ฮาร์ดแวร์ของเครื่องคอมพิวเตอร์ ประกอบด้วยหน่วยความจำสำรองเช่น ดิสค์ และอื่นๆ ไว้ทำหน้าที่เก็บ และใช้ค้นหาข้อมูล โดยมีหน่วยประมวลผลกลาง, หน่วยความจำ และอื่นๆ ที่จัดการให้อุปกรณ์ทำงานร่วนกันได้

2.1.3 ซอฟแวร์ (Software)

การควบคุมอุปกรณ์ให้ทำงานได้สัมพันธ์กัน รองรับผู้ใช้ได้พร้อมกันหลายคนเป็นหน้าที่ของระบบปฏิบัติการ ในส่วนของระหว่าง ผู้ใช้ กับข้อมูลทางกายภาพ (ข้อมูลจริงที่ถูกเก็บอยู่บนหน่วยความจำสำรอง) จะมีโปรแกรมประยุกต์ทำงานประสานกับระบบปฏิบัติการ เพื่อให้ได้ข้อมูลตามที่ผู้ใช้ต้องการ โดยเป็นโปรแกรมที่เขียนจากภาษาที่คนเข้าใจได้ง่าย แล้วใช้ตัวแปลคอมไพเลอร์ แปลเป็นภาษาเครื่อง เรียกว่าภาษาโปรแกรมมิ่ง เช่น โคบอล (COBOL) จะเป็นตัวช่วยให้สั่งฮาร์ดแวร์ให้ทำงานระดับเครื่องได้เช่น การดึงข้อมูลจากแฟ้มพนักงานเป็นต้น

2.1.4 ผู้ใช้ (User)

ในส่วนของผู้ใช้แบ่งได้เป็น 3 ส่วน

1. ผู้เขียนโปรแกรม จะต้องเขียนโปรแกรมให้สามารถเข้าใช้ข้อมูลโดยใช้ภาษาโปรแกรม เช่น โคบอล เป็นต้น จะได้เป็นโปรแกรมประยุกต์ ซึ่งจะทำงานกับข้อมูลทุกๆ อย่างตั้งแต่ขอดูข่าวสาร, เพิ่ม, เปลี่ยนแปลง ลบข้อมูล โดยโปรแกรมจะสนับสนุนช่วยเหลือการทำงานให้กับผู้ใช้

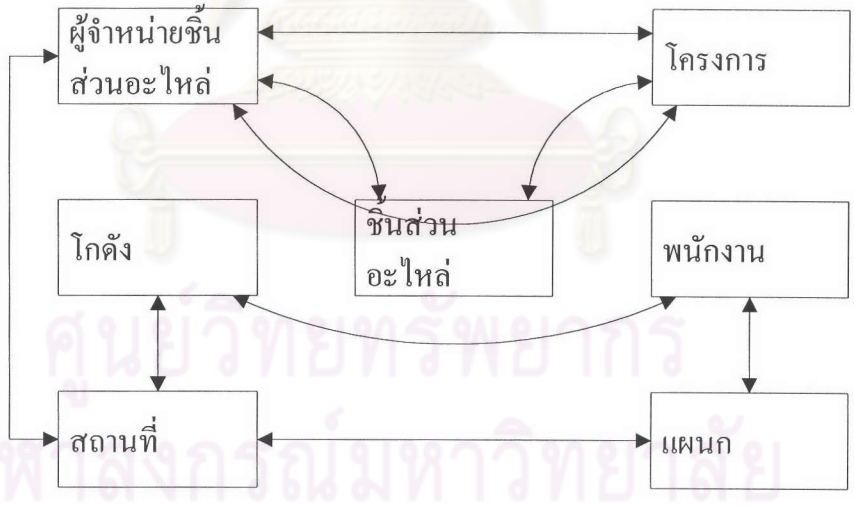
2. ผู้ใช้ ที่เป็นผู้ปฏิบัติงานทางธุรกิจ จะเข้าถึงข้อมูลผ่านทางเทอร์มินัล หรืออุปกรณ์รับเข้า หรือแสดงผลข้อมูลโดยจะเป็นหน้าที่ของโปรแกรมประยุกต์

3. ผู้ดูแลระบบจะต้องดูแลซอฟแวร์ และฮาร์ดแวร์ให้งานของหน่วยงานเดินไปได้

2.2 <u>ข้อมูลดำเนินการ และสายงานของข้อมูล</u> (Operational data and Information Flow)

ข้อมูลทุกชนิดจะถูกนำข้อมูลเข้า และนำข้อมูลออก โดยข้อมูลที่ถูกรวบรวมเก็บไว้เหล่านั้นจะเป็นข้อมูลที่ใช้ในการดำเนินการ ตามที่โปรแกรมประยุกต์ที่พัฒนาไว้ใช้สำหรับองค์กรโดยเฉพาะ องค์กรอาจจะเป็นหน่วยงานทางธุรกิจ, วิทยาศาสตร์, เทคนิค หรือองค์กรอื่นๆ ตัวอย่างเช่น โรงงานผลิต, ธนาคาร, โรงพยาบาล, มหาวิทยาลัย องค์กรของรัฐบาล ดังนั้นองค์กรเหล่านี้จำเป็นต้องเก็บและดูแลข้อมูลที่ใช้ในการทำงานเป็นจำนวนมาก ซึ่งข้อมูลเหล่านี้ก็คือข้อมูลในการทำงาน เช่น ข้อมูลผลผลิต, ข้อมูลบัญชี, ข้อมูลคนไข้, ข้อมูลนักศึกษา, ข้อมูลการวางแผน เป็นต้น

ข้อมูลที่ใช้ดำเนินการไม่ใช่มีแค่ข้อมูลนำเข้าที่เกี่ยวเนื่องกับเทอมินัลหรือเทปแม่เหล็ก หรือข้อมูลนำออกคือ ข่าวสาร และรายงานจากระบบ อาจจะแสดงที่จอเทอมินัลหรือเครื่องพิมพ์ก็ได้ รวมถึงเอนทิตีพื้นฐานด้วย (Basic entity)



รูปที่ 2.2 ตัวอย่างของข้อมูลที่ทำให้เกิดการดำเนินงาน การไหลไปของข้อมูล

แนวคิดที่แสดงให้เห็นการทำงานของข้อมูล พิจารณากรณีตัวอย่างที่เป็นโรงงานผลิตต้องคอยควบคุมโครงการที่อยู่ในความดูแล เช่นชิ้นส่วนที่ต้องใช้ในโครงการ ผู้จัดหาชิ้นส่วน

อะไหล่, โกดังเก็บของ พนักงานที่ทำงานกับโครงการ นี่คือ เอนทิตีพื้นฐานที่ข้อมูลจะต้องเก็บไว้ในระบบ ทั้งหมดนั้นจะมีความสัมพันธ์กัน ซึ่งต้องเชื่อมเข้าหากันให้รวมกันให้ได้ ตามรูปจะแทนการเชื่อมต่อกันระหว่างเอนทิตี ความสัมพันธ์ระหว่างผู้จำหน่ายชิ้นส่วนอะไหล่ กับชิ้นส่วน แต่ละผู้จำหน่ายจะมีอะไหล่ที่จะต้องส่งไม่เหมือนกัน ในทำนองเดียวกันชิ้นส่วนที่ใช้ในโครงการ, โครงการใช้ชิ้นส่วนอะไร, ชิ้นส่วนเก็บไว้ในโกดังที่ไหน และโกดังเก็บชิ้นส่วนอะไร แสดงให้เห็นความสัมพันธ์ในสองทิศทาง จากตัวอย่างความสัมพันธ์ระหว่างพนักงานกับแผนก อาจจะให้คำตอบได้ทั้งสองด้าน

(ก). ในแผนกมีผู้ใดเป็นพนักงานบ้าง

(ข). พนักงานอยู่แผนกใด

2.3 รูปแบบของระบบการประมวลผล (Processing System Configuration)

รูปแบบวิธีการที่จะนำข้อมูลมาประมวลผลแบ่งได้ 3 รูปแบบ

1. การประมวลผลข้อมูลแยกศูนย์ (Decentralized Data Processing)
2. การประมวลผลข้อมูลแบบรวม (Centralized Data Processing)
3. การประมวลผลข้อมูลแบบกระจาย (Distributed Data Processing)

2.3.1 การประมวลผลข้อมูลแบบแยกศูนย์

รูปแบบแยกศูนย์นี้ เป็นการประมวลผลที่ทั้งข้อมูล และโปรแกรมประยุกต์ของแต่ละหน่วยงาน ซึ่งเป็นหน่วยงานภายในองค์กรเดียวจะแยกกัน เป็นการประมวลผลที่แต่ละหน่วยงานต่างฝ่ายต่างประมวลผลไปได้เองโดยมีแฟ้มข้อมูลเป็นของหน่วยงาน ข้อมูลต่างๆนั้นไม่สามารถใช้ร่วมกันได้ แต่ละหน่วยจะดูแลรักษาข้อมูลของตนเอง ลักษณะงานและการประมวลผลจะเป็นดังนี้

1. แต่ละหน่วยภายในองค์กรเดียวกันต่างแยกกันประมวลผล
2. แต่ละหน่วยจะเป็นเจ้าของข้อมูล
3. โปรแกรมประยุกต์เป็นแบบเชิงประมวลผล ไม่ใช่แบบเชิงข้อมูล คือคำนึงถึงหน่วยงานเป็นหลัก โดยหน่วยงานจะพิจารณาเฉพาะผลลัพธ์ที่จะได้ และความรวดเร็ว

ปัญหาของการประมวลผลแบบแยกศูนย์คือ

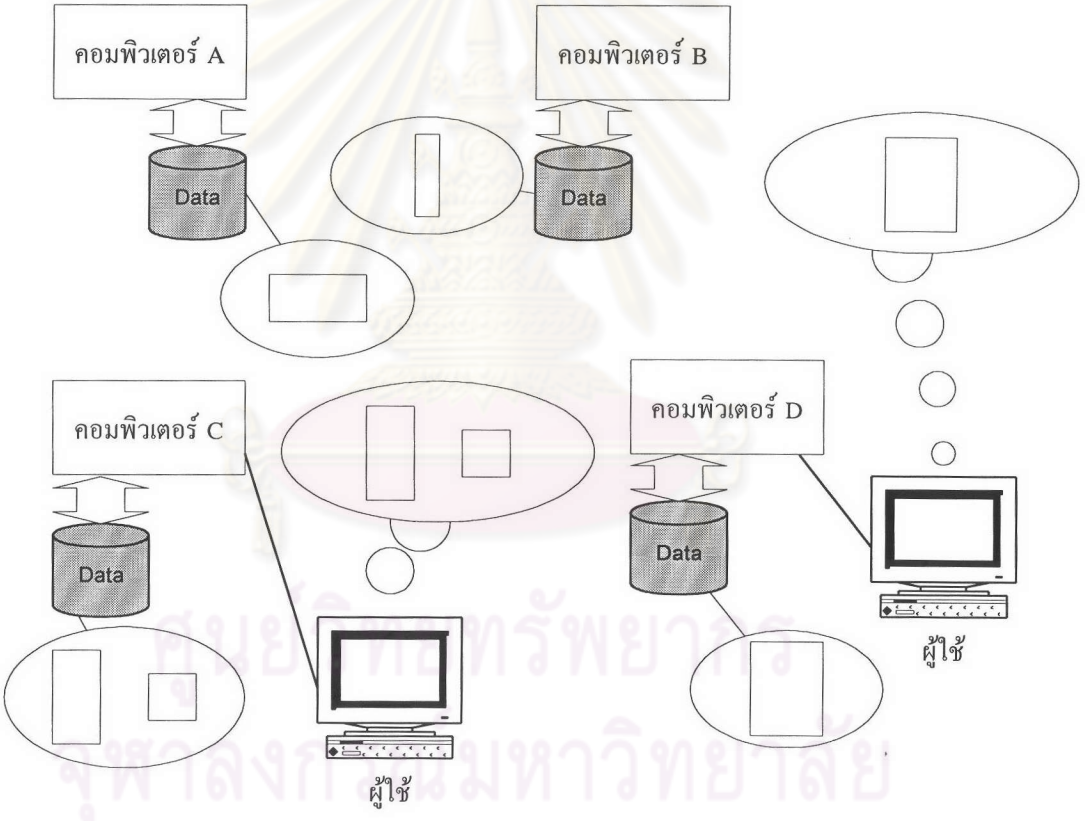
1. ข้อมูลแต่ละหน่วยงานจะเก็บข้อมูลอย่างเดียวกัน ถูกเก็บไว้หลายๆ แห่ง ทำให้เกิดการซ้ำกันของข้อมูลในองค์กรเดียวกัน (redundancy)

2. ข้อมูลที่เก็บไว้ซ้ำกันหลายๆ ที่นั้น เมื่อเกิดการที่ข้อมูลต้องถูกแก้ไข จะทำไม่ตรงกัน อาจเนื่องจากเหลื่อมกันทางด้านเวลา หรือทำไม่เหมือนกันทุกแห่ง (inconsistency)

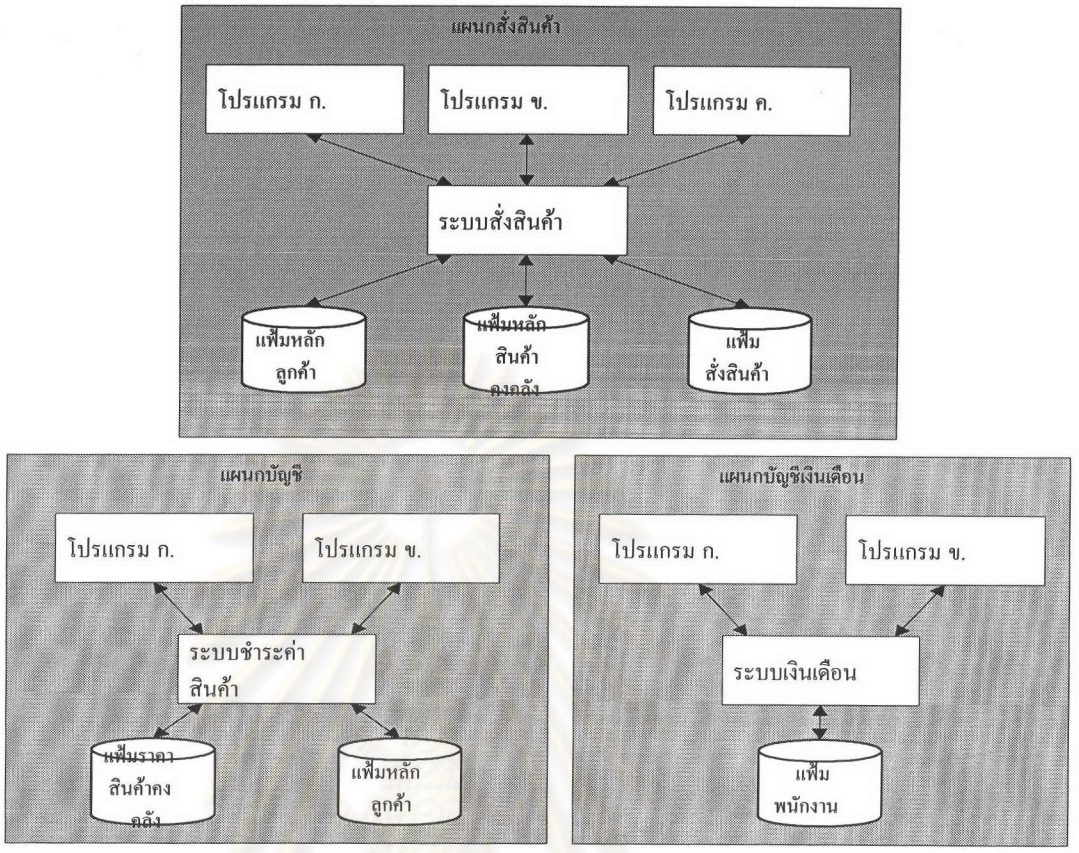
3. มีข้อจำกัดในการนำข้อมูลมาใช้ร่วมกันได้ (shared data) ร่วมทั้งโปรแกรมประยุกต์ก็ต่างแยกกันพัฒนาและแยกกันใช้งาน

4. เนื่องจากต่างแยกกันพัฒนา ทำให้ไม่เกิดมาตราฐานของข้อมูลในองค์กร (standards) เนื่องจากควบคุมการกำหนดให้ใช้ไม่ได้

5. ค่าใช้จ่ายในการบำรุงรักษาอุปกรณ์, และโปรแกรมประยุกต์สูง เนื่องจากแยกกันพัฒนา และดูแล



รูปที่ 2.3 การประมวลผลแบบแยกศูนย์



รูปที่ 2.4 แสดงแฟ้มข้อมูล และการประมวลผลแยกกันในแต่ละหน่วยงาน

การปรับปรุงข้อมูลในรูปแบบแยกศูนย์นี้เพื่อไม่ให้ข้อมูลขัดแย้งกัน (inconsistency) โดยให้แต่ละหน่วยงานส่งแฟ้มข้อมูล ที่หน่วยงานอื่นต้องใช้ แต่ในเงื่อนไขต้องแปลงรูปแบบแฟ้มให้ใช้กันได้ ผ่านอุปกรณ์สื่อกลางสำหรับส่งแฟ้มข้อมูล และต้องคำนึงถึงความปลอดภัยของข้อมูลด้วย

จากรูปที่ 2.4 เป็นรูปแบบของระบบงานแยกศูนย์ จะเห็นได้ว่าแผนกสั่งสินค้า กับแผนกบัญชีมีความเกี่ยวเนื่องกันในด้านข้อมูล แผนกสั่งสินค้าเมื่อได้รับใบเบิกอะไหล่ ต้องตรวจดูคงคลัง ถ้าอะไหล่หมดจะต้องทำการสั่งซื้อ ข้อมูลการสั่งซื้อต้องแจ้งให้แผนกบัญชีส่วนแผนกบัญชีจะต้องมีข้อมูลไว้เตรียมที่จะเบิกเงินเพื่อชำระค่าอะไหล่เมื่อผู้จำหน่ายได้ส่งใบแจ้งหนี้มาเรียกเก็บเงิน แผนกบัญชีต้องตรวจการรับของจากเอกสารการส่งสินค้า ทั้งหมดต้องทำงานสัมพันธ์โดยผ่านทางข้อมูล ซึ่งข้อมูลจะเดินทางจากแผนกหนึ่งไปอีกแผนกหนึ่ง อาจใช้สื่อกลางเช่นเอกสาร เป็นต้น

2.3.2 <u>การประมวลผลข้อมูลแบบรวมศูนย์</u>

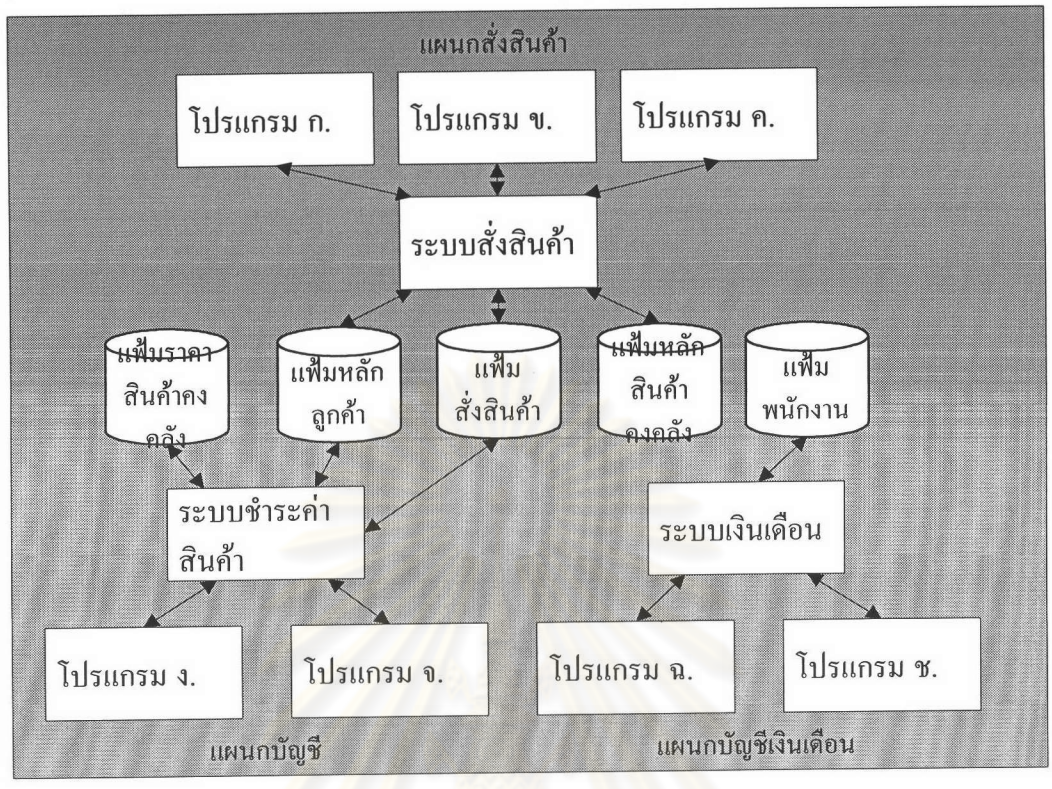
การประมวลผลที่ที่นำมาข้อมูลเดิมแต่ละหน่วยงานที่อยู่ภายในองค์กร มาจัดรวมกันอยู่ที่เดียวหมายถึง แฟ้มข้อมูลบันทึกอยู่ในเครื่องคอมพิวเตอร์เครื่องเดียวกัน แล้วให้ผู้ใช้ที่อยู่ ณ.ที่หน่วยงานต่าง ๆ เข้ามาประมวลผล โดยใช้ข้อมูลร่วมกัน ที่ศูนย์กลางแล้วรับผลกลับไป จากการที่นำข้อมูลมาอยู่ที่เดียวแล้วใช้ข้อมูลร่วมกัน เป็นการช่วยแก้ปัญหาของการประมวลผลแบบแยกศูนย์ได้หลายอย่างคือ

1. แต่ละหน่วยงานจะใช้ข้อมูลเดียวกัน ลดการเก็บข้อมูลเดียวกันซ้ำซ้อน
2. ข้อมูลถูกต้องตรงกับความเป็นจริง เมื่อต้องการแก้ไขข้อมูลจะกระทำครั้งเดียวทุกหน่วยงานจะเห็นข้อมูลชุดนั้นเหมือนกันหมด
3. ใช้ข้อมูลร่วมกันได้
4. สามารถบังคับให้เป็นมาตราฐานได้
5. ง่ายต่อการพัฒนาโปรแกรม
6. ความปลอดภัย และความเป็นอันหนึ่งอันเดียวกัน
7. สามารถเข้าถึงข้อมูลที่ต้องการทันที ไม่ต้องรอขอข้อมูลข้ามหน่วยงาน
8. ลดการดูแลและปรับปรุงโปรแกรม เนื่องการพัฒนาโปรแกรมประยุกต์แล้วใช้ร่วมกัน

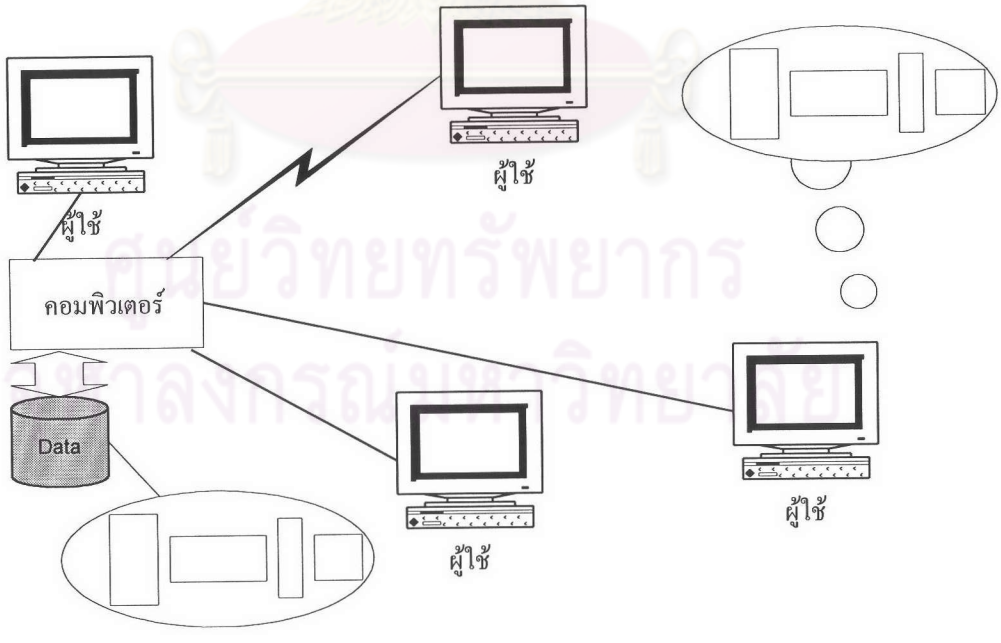
งานหลักๆ ที่มีความสำคัญต่อองค์กร เช่นงานบัญชี ควบคุมพัสดุคงคลัง ระบบการรับการสั่งซื้อ ระบบจัดซื้อ และอื่นๆ อีกมาก ซึ่งระบบเหล่านี้ต้องประสานกันทั้งหมด เพื่อให้เป็นไปตามข้างต้น จึงต้องนำข้อมูลมารวมกัน แล้วส่งผ่านข้อมูลข้ามถึงกันโดยอัตโนมัติ หรือเป็นข้อมูลชุดเดียวกันที่ทุกระบบงานดึงไปใช้งานได้ตลอด

2.3.2.1 <u>โครงสร้างของระบบรวมศูนย์</u>

เมื่อระบบปฏิบัติการสามารถประมวลผลงานได้หลายๆ งานพร้อมกันในเวลาเดียวกันเรียกว่า ระบบหลายงาน (Multitasking System) และมีอุปกรณ์ฮาร์ดแวร์สำหรับติดต่อกับผู้ใช้ได้หลายคนผ่านเทอร์มินัล หรืออุปกรณ์รับ และแสดงผล ระบบปฏิบัติการก็สามารถรองรับผู้ใช้ได้มากกว่าหนึ่ง จะเป็นระบบแบบหลายผู้ใช้ หลายงาน (Multi-user Multitasking System)



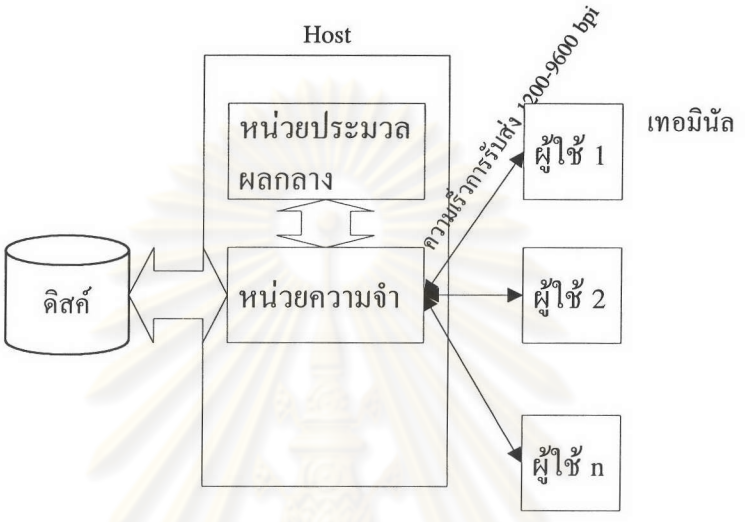
รูปที่ 2.5 แสดงแฟ้มข้อมูล และการประมวลผลร่วมกันของแต่ละหน่วยงานในส่วนอุปกรณ์เดียวกัน



รูปที่ 2.6 การประมวลผลแบบรวมศูนย์

2.3.2.2 ปัญหาของการประมวลผลแบบรวมศูนย์

เนื่องจากงานทั้งหมดในองค์กรจะถูกส่งมาประมวลผลที่เครื่องคอมพิวเตอร์ศูนย์กลาง ทำให้เครื่องรับภาระงานมาก ถ้าองค์กรมีการขยายงาน อาจจะทำให้เกิดรับภาระงานเกินกำลังได้ วิธีแก้ไขคือเปลี่ยนเครื่องคอมพิวเตอร์ใหม่ ที่สมรรถนะที่รองรับงานทั้งหมดได้

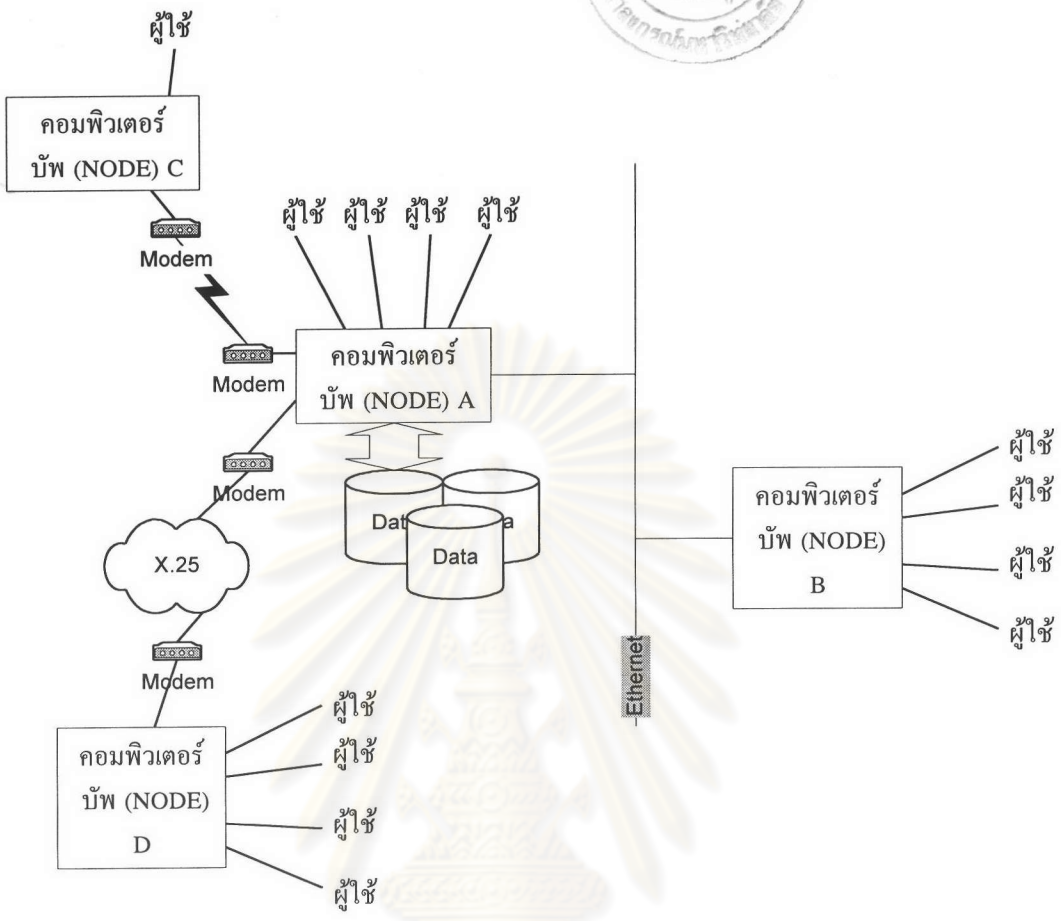


รูปที่ 2.7 ผังแสดงระบบฮาร์ดแวร์ที่ประมวลผลแบบหลายผู้ใช้ หลายงานในเวลาเดียวกัน

2.3.3 การประมวลผลข้อมูลแบบการกระจาย (Distributed Data Processing)

2.3.3.1 แนวความคิด และเป้าหมายของการกระจาย

องค์ประกอบที่เกี่ยวข้องในการประมวลผลข้อที่ 5 ที่เพิ่มเข้ามาในปัจจุบันก็คือ การต่อคอมพิวเตอร์ ผ่านสายสัญญาณ หรือระบบการสื่อสาร มีซอฟแวร์ควบคุมการรับส่งระหว่างเครื่องคอมพิวเตอร์ระหว่างสองเครื่องหรือมากกว่า ให้สามารถแลกเปลี่ยนข้อมูล ข่าวสารกันได้ เมื่อต่อคอมพิวเตอร์หลายเครื่อง ก็จะเป็นเครือข่ายคอมพิวเตอร์ โดยผู้ใช้สามารถเรียกใช้ซอฟแวร์นี้กับเครื่องที่อยู่ระยะไกลได้โดยเสมือนเป็นเครื่องเดียวกันกับที่ผู้ใช้ ปฏิบัติงานอยู่



รูปที่ 2.8 รูปแบบของเครือข่ายคอมพิวเตอร์ (Computer Network Configuration)

รูปแบบการดำเนินงานขององค์กรธุรกิจในการจัดองคาพยพ (Organization of Business) มักจะแยกกันบริหารเพื่อความคล่องตัวแยกเป็นแผนก หรือตามความเชี่ยวชาญของบุคลากร หรือถ้าธุรกิจต้องขยายสาขาไปตามที่ต่าง ๆ ที่ต้องห่างไกลจากศูนย์กลางเดิม ทำให้แต่ละสาขาต้องแยกตัวออกไป ลักษณะการทำงานก็จะเป็นการทำงานเฉพาะสาขาของตนเอง การที่นำข้อมูล และโปรแกรมประยุกต์มารวมอยู่ที่ศูนย์กลางทำให้เครื่องประมวลผลส่วนกลางต้องทำงานหนัก เพราะต้องรับหน้าที่เกี่ยวกับการจัดเก็บ และการประมวลทั้งหมดของทุกสาขา และของศูนย์กลางเองด้วย หากระบบส่วนกลางเกิดชำรุด การทำงานทั้งหมดก็จะหยุดชะงักทันที

ส่วนรูปแบบการกระจายข้อมูลนั้นในเชิงโปรแกรมประยุกต์ ได้แก่การกระจายแฟ้มข้อมูลเฉพาะที่หน่วยงานนั้นเป็นเจ้าของเช่น ถ้าเป็นแผนกบุคคล ก็เป็นเจ้าของแฟ้มพนักงาน เป็นต้น ถ้าเป็นสาขาก็จะเก็บแฟ้มข้อมูลของสาขา และข้อมูลลูกค้าของสาขาตัวเองโดยข้อมูล และโปรแกรม

ประยุกต์จะเก็บไว้ที่เครื่องคอมพิวเตอร์ของแผนก หรือของสาขา การประมวลผลใดๆ ก็จะทำบนเครื่องของแผนก หรือของสาขาเท่านั้น ถ้าการประมวลผลที่แผนก (ท้องถิ่น หรือ local) ต้องการข้อมูลข้ามแผนก (ระยะไกล หรือ remote) หรือข้ามสาขา ซอฟแวร์เครือข่ายจะทำหน้าที่ไปนำข้อมูลจากเครื่องคอมพิวเตอร์จากแผนกที่ตั้งอยู่ระยะไกลแล้วส่งกลับมาให้แผนกท้องถิ่นประมวลผลต่อ เทคนิคนี้เป็นการลดภาระงานของการประมวลผลแบบรวมศูนย์ และเมื่อเครื่องคอมพิวเตอร์แผนกใดชำรุด เครื่องคอมพิวเตอร์ของแผนกอื่นๆ ก็ยังทำงานต่อไปได้ เนื่องจากเครื่องคอมพิวเตอร์แต่ละแผนกมีความเป็นอิสระต่อกัน โดยต่อกันอยู่แบบหลวมๆ (Loosely Coupled)

ดังที่ได้กล่าวมาแล้วว่าแทนที่จะใช้เครื่องคอมพิวเตอร์ประมวลผลเพียงเครื่องเดียวที่ศูนย์กลาง แต่ให้แต่ละสาขามีระบบเป็นของตัวเอง ถ้าสาขาใดมีปริมาณงานมากก็ให้ความเร็ว และความจุในการเก็บมากๆ ถ้าแผนก หรือสาขาใดปริมาณงานไม่มากก็ใช้ระบบเล็กก็ได้ ส่วนถ้าเกิดงานเพิ่ม หรือขยายก็สามารถเปลี่ยนเครื่องให้ดีขึ้นเฉพาะสาขาที่จำเป็น ไม่ต้องเปลี่ยนทั้งระบบ เป็นการชะลอ หรือลดการลงทุนได้ ส่วนเครื่องเก่าก็สามารถนำไปติดตั้งใช้ที่แผนกที่ยังไม่มี หรือมีแล้วแต่ปริมาณงานเพิ่มขึ้นเป็นผลดีข้อหนึ่ง

2.3.3.2 แนวความคิด

แนวความคิดที่ต้องการทำให้เกิดรูปแบบการกระจายคือ

1. ธุรกิจส่วนใหญ่มักจะขยายหน่วยงานไปอยู่หลายที่
2. โครงสร้างของระบบธุรกิจ เพื่อง่ายต่อการบริหารงาน และจัดการภายในจะแบ่งแยกเป็นหน่วยงานย่อย
3. แบ่งภาระการทำงานของการประมวลผลศูนย์
4. ติดตั้งเครื่องที่เหมาะสมกับภาระงานของแต่ละหน่วย

สรุประบบแบบกระจาย (Distributed System) คือระบบที่เก็บแฟ้มข้อมูลขององค์กรไว้ในเครื่องคอมพิวเตอร์ ที่ติดตั้งอยู่ตามแผนก หรือสาขาต่าง ๆ หลายเครื่องโดยที่แต่ละเครื่องมีการติดต่อสื่อสารถึงกันทางข้อมูล ซึ่งผู้ใช้สามารถใช้งานได้โดยไม่จำเป็นต้องรู้ที่ตั้งของแฟ้มข้อมูล แต่จะเป็นหน้าที่ของโปรแกรมประยุกต์ และซอฟแวร์เครือข่าย

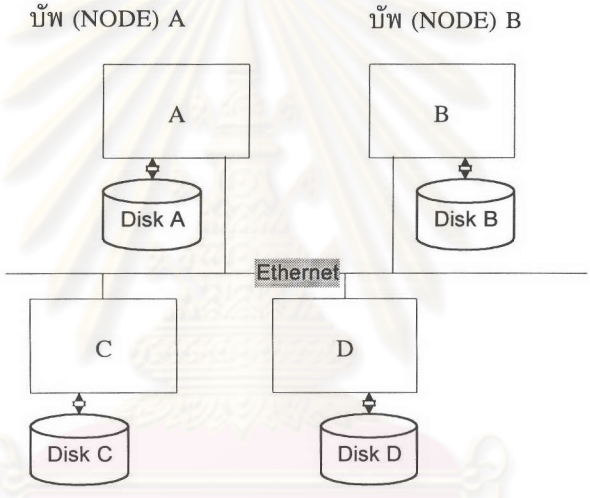
2.3.3.3 ระบบจะกระจายได้ก็ต่อเมื่อ มีเงื่อนไขดังนี้

1. อรรถประโยชน์ของคอมพิวเตอร์ทำงานเป็นเครือข่าย สามารถทำงานแบบกระจายการประมวลผล (Distributed Processing) หรือกระจายข้อมูล (Distributed Data) ได้

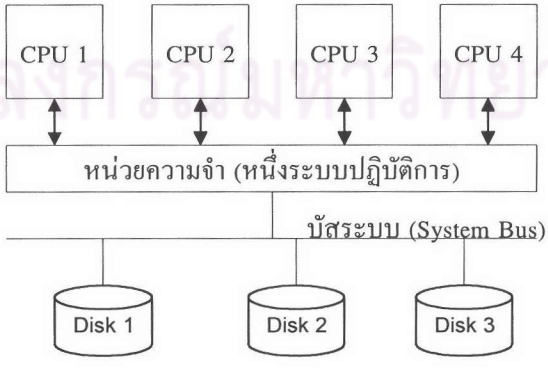
2. คอมพิวเตอร์แต่ละบัพ (node) ประมวลผลงานได้อย่างเป็นอิสระจากตัวอื่น (ส่งถ่ายข้อมูลกันอย่างหลวม ๆ) ตามรูปที่ 2.9 (Loosely Coupled System) ตัวประมวลผลแบบส่งถ่ายอย่างหลวมนั้นไม่ใช่เป็นการใช้หน่วยความจำร่วมกัน หรือใช้จานบันทึกร่วมกัน แต่เป็นการใช้ข้อมูลร่วมกัน โดยการส่งข่าวสารถึงกันข้ามไปมาทางเครือข่าย ส่วนระบบตามรูป 2.10 ไม่สามารถกระจายแฟ้มข้อมูลได้

3. ระบบที่ดำเนินงานแบบกระจายแฟ้มได้สำเร็จ ผู้ใช้ไม่จำเป็นต้องรู้ถึงการกระจาย โดยเป็นหน้าที่โปรแกรมประยุกต์ และซอฟแวร์เครือข่าย

4. การออกแบบระบบแฟ้มข้อมูล และโปรแกรมประยุกต์ต้องใช้วิธีการของ เอนทิตี รีเลชั่นโมเดล (Entity-Relation Model) จึงจะสามารถกระจายแฟ้มข้อมูลได้



รูปที่ 2.9 เครือข่ายคอมพิวเตอร์ เป็นระบบที่ส่งถ่ายกันแบบหลวมๆ (Loosely Coupled System)



รูปที่ 2.10 สถาปัตยกรรมคอมพิวเตอร์ที่เป็นแบบ ระบบส่งถ่ายกันแบบแน่นหนา (Tightly Coupled System)

2.4 การมองผ่าน (Transparency) ของผู้ใช้ต่อแฟ้มข้อมูล (Pratt and Adamski, 1994)

เมื่อเกิดการกระจายแล้วการประมวลผลงานยังต้องเหมือนกันกับที่วัตถุประสงค์ของระบบแบบรวมศูนย์คือ ผู้ใช้ยังมีความรู้สึกต่อการทำงานเหมือนเดิมที่เรียกว่า การมองผ่าน (Transparency)

การมองผ่าน (Transparency) หมายถึงมุมมองของผู้ใช้ที่มีต่อข้อมูล มีหลายแบบ แต่จุดประสงค์หลักที่ผู้ใช้ยอมรับได้ และโปรแกรมประยุกต์ทำงานได้โดยไม่ต้องคำนึงถึงตำแหน่งที่ตั้งของแฟ้มข้อมูล

ชนิดของการมองผ่านประกอบด้วย

1. การมองผ่านสถานที่ตั้ง (Location Transparency) ผู้ใช้ไม่ต้องคำนึงถึงที่ตั้ง ลักษณะการประมวลผลแล้ว ผู้ใช้มีความรู้สึกว่ากำลังใช้ข้อมูลที่เครื่องเฉพาะที่ของเขาเอง (Local site) โดยที่จริงข้อมูลที่ผู้ใช้ๆอยู่เป็นข้อมูลจากเครื่องระยะไกล (Remote Site)

2. การมองผ่านเส้นทาง (Routing Transparency) ผู้ใช้ไม่ต้องรู้ถึงเส้นทางที่ไปยังที่ตั้งของแฟ้มข้อมูล เนื่องจากการต่อคอมพิวเตอร์เป็นเครือข่ายจะมีหลายเส้นทางที่จะไปเรียกใช้ข้อมูลจากเครื่องที่ตั้งอยู่ระยะไกลได้

3. การมองผ่านถึงการเก็บซ้ำ (Replication Transparency) ผู้ใช้ไม่ต้องรับรู้ถึงการเชื่อมโยงระหว่างบัพไม่ได้ ต้องจัดให้มีทางเลือก โดยที่ผู้ใช้ยังสามารถเรียกใช้แฟ้มข้อมูลได้แม้ว่าการเชื่อมโยงขาดจากกัน นั้นคือต้องเก็บข้อมูลซ้ำไว้หลายเครื่อง หรือถ้าผู้ใช้ต้องการความเร็วในการประมวลผลก็เก็บซ้ำไว้ที่เครื่องเฉพาะที่ (local) แต่ต้องคอยปรับปรุงแฟ้ม และการประมวลผลนั้นจะต้องไม่ใช่เป็นการประมวลผลข้อมูล ที่ข้อมูลต้องทันกาลตลอด

4. การมองผ่านการแตกกระจาย (Fragmentation Transparency) คือการแตกกระจายแฟ้ม แยกข้อมูลออกมาเป็นแฟ้มใหม่ แล้วไปติดตั้ง บัพที่ผู้ใช้ใช้งานแบบปรับปรุงข้อมูล เนื่องจากการออกแบบเดิมแฟ้มจะเก็บข้อมูลที่มีความสัมพันธ์กันไว้ในแฟ้มเดียว เป็นการลดการเก็บข้อมูลซ้ำซ้อน แต่เมื่อพัฒนาการกระจายแฟ้มแล้ว การแยกข้อมูลออกมาเป็นแฟ้ม เพื่อประโยชน์กรณีที่บางแฟ้มบางหน่วยงานใช้ข้อมูลในแฟ้มเพียงบางส่วน และหน่วยอื่นก็ใช้แฟ้มเดียวกันแต่เป็นข้อมูลอีกส่วนหนึ่งเป็นต้น ถ้าพิจารณาแยกข้อมูลออกมาเป็นแฟ้มใหม่ก็จะทำให้ลดการเข้าถึงแฟ้มระยะไกล

2.5 ชนิดของระบบแบบกระจาย (Oracle Corporation Ltd, 1990)

1. การกระจายการประมวลผล (Distributed Processing)
2. การกระจายแฟ้มข้อมูล (Distributed Data Files)

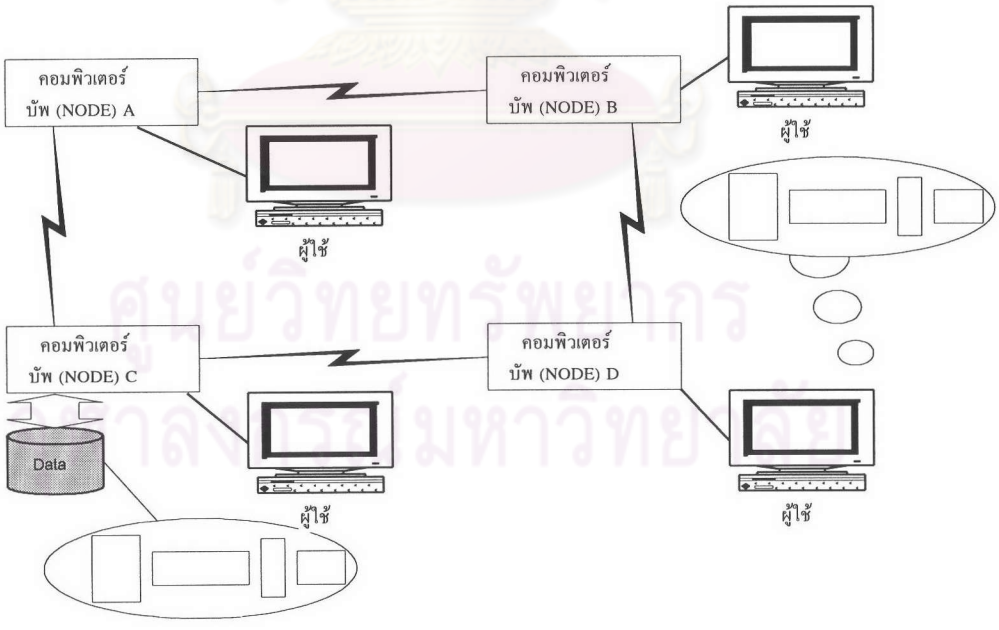
2.5.1 การกระจายการประมวลผล

ลักษณะการประมวลผลคือ โปรแกรมประยุกต์จะดำเนินงานบนบัพหนึ่ง แล้วไปเรียกใช้ข้อมูลที่เก็บไว้อีกบัพหนึ่ง โดยโปรแกรมประยุกต์ทุกกระบวนการจะเข้าไปเรียกใช้ข้อมูลที่ บัพ (node) เดียวกัน ที่ถูกกำหนดเป็นที่เก็บข้อมูลเท่านั้น โดยที่โปรแกรมประยุกต์จะดำเนินงาน อยู่ที่บัพเดียวกัน หรือต่างบัพกันได้

กระจายการประมวลผลไปที่บัพต่างๆ แต่ข้อมูลตั้งอยู่ที่บัพเดียว ตัวอย่างของกระบวนการแบบนี้เป็นรูปแบบที่เรียกว่า ไคล์นเอนท์/เซิฟเวอร์ (Client/Server)

ข้อดีของการกระจายการประมวลผล

1. สามารถช่วยลดงานที่เครื่องศูนย์กลางได้ โดยตัดงานการประมวลผลบางอย่างออกไป
2. คอมพิวเตอร์ต่างหน่วยงานกันเข้าใช้ฐานข้อมูลกลาง ทุก ๆ หน่วยงานจะเห็นข้อมูลเป็นปัจจุบัน



รูปที่ 2.11 การกระจายการประมวลผล โดยข้อมูลอยู่ที่เดียวที่กำหนด

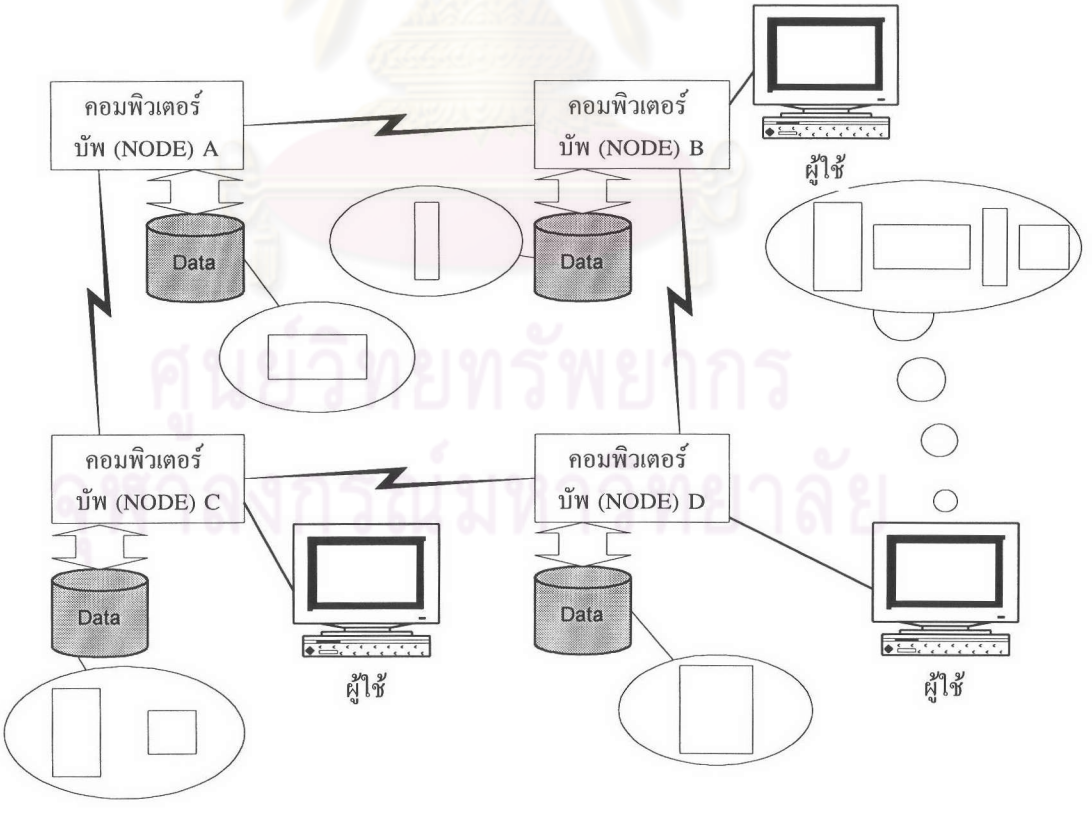
2.5.2 การกระจายแฟ้มข้อมูล

แฟ้มข้อมูลทางกายภาพจะมีมาก และตั้งอยู่หลายที่ อยู่ต่างบัพกัน แต่จะปรากฏเป็นข้อมูลทางตรรกะชุดเดียว

ข้อดีของการกระจายข้อมูล คือการมองผ่านที่ตั้ง (Location Transparency) ผู้ใช้มองเห็นเป็นข้อมูลชุดเดิมอยู่ทำให้ง่ายต่อการเติบโตคือ

1. เพิ่มคอมพิวเตอร์เข้าไปในเครือข่ายได้มากขึ้นอยู่กับความต้องการขององค์กร
2. เพิ่มเนื้อที่ในการเก็บข้อมูลได้มากขึ้น เนื่องจากข้อมูลถูกกระจายออกไปทำให้มีเนื้อที่ว่างสำหรับเก็บข้อมูลที่จะเก็บเพิ่มขึ้นอีก
3. เพิ่มจำนวนผู้ใช้ เนื่องจากบัพที่ต่ออยู่กับเครือข่ายสามารถรองรับได้ตามการเพิ่มขึ้นของบัพ

จะเห็นว่าการกระจายข้อมูลยังเป็นการกระจายการประมวลผลไปด้วยในตัว เนื่องจากการประมวลผลจะกระทำได้ทุกบัพที่ต้องการ บัพทุกตัวจะเป็นได้ทั้งไคลน์เอนท์ และเซิฟเวอร์ในตัวเดียวกัน



รูปที่ 2.12 การกระจายแฟ้มข้อมูลไปตามบัพในเครือข่าย โดยผู้ใช้ยังเห็นเป็นข้อมูลเดียว