บทที่ 3

ระบบฐานข้อมูลแบบกระจาย

ระบบฐานข้อมูลแบบกระจาย(Distributed Database System)

ฐานข้อมูลแบบกระจาย คือ ระบบฐานข้อมูลที่มีฐานข้อมูลเก็บอยู่ในระบบคอมพิวเตอร์หลาย ๆ เครื่อง ที่ติดตั้งอยู่ตามที่ต่าง ๆ โดยระบบคอมพิวเตอร์เหล่านี้สามารถสื่อสารถึงกันได้ซึ่งผู้ใช้สามารถเรียกใช้ข้อมูลจากฐานข้อมูลที่อยู่ในเครื่องใดก็ได้ โดยที่ผู้ใช้ไม่จำเป็นต้องรับทราบว่าข้อมูลที่ตนต้องการนั้นถูกจัดเก็บอยู่บนเครื่องใด กล่าวคือในทางกายภาพแล้วข้อมูลจะถูกเก็บกระจายอยู่ตามที่ต่างๆในระบบเครือข่ายคอมพิวเตอร์ แต่ในแง่มุมของผู้ใช้จะเสมือนว่ากำลังใช้ฐานข้อมูลแบบรวมศูนย์อยู่ตามปกติ นอกจากนี้แล้วแต่ละสถานียังมีความสามารถดำเนินงานและควบคุมระบบงานของตนเองได้อย่างอิสระ (ดวงแก้ว สวามิภักดิ์,2521)

การออกแบบฐานข้อมูลแบบกระจาย

1. การออกแบบจากบนลงล่าง (Top-down approach)

วิธีนี้จะเริ่ มการออกแบบระดับเชิงมโนภาพรวมก่อน แล้วจึงออกแบบลักษณะการกระจายของข้อมูลสำหรับการกำหนดระดับเชิงมโนภาพย่อยอีกทอดหนึ่ง และต่อมาจึงทำการออกแบบฐานข้อมูลเชิงกายภาพซึ่งเป็นการแปลงระดับเชิงมในภาพย่อยไปยังอุปกรณ์ที่ใช้เก็บข้อมูลจริง ซึ่งวิธีนี้เหมาะกับการสร้างฐานข้อมูลใหม่ โดยที่ยังไม่เคยมีระบบฐานข้อมูลใด ๆ มาก่อนเลย

2. การออกแบบจากล่างขึ้นบน (Bottom-up approach)

วิธีนี้เหมาะสำหรับการพัฒนาฐานข้อมูลที่มีอยู่แล้วให้เป็นฐานข้อมูลแบบกระจาย ซึ่งต้องรวบรวมระบบฐานข้อมูลหลายๆระบบที่มีอยู่ให้เข้าเป็นระบบเดียวกัน กล่าวคือเป็นการรวบรวมระดับมโนภาพย่อยให้เป็นระดับมโนภาพรวมนั่นเอง

การออกแบบการแตกกระจายของฐานข้อมูล (Database Fragmentation Design)

การออกแบบการแตกกระจายของฐานข้อมูล คือการแยกข้อมูลออกเก็บตามสถานีประมวลผลต่าง ๆ เพื่อให้ระบบมีผลลัพธ์ของประสิทธิภาพสูงสุด ซึ่งโดยทั่ว ๆ ไปก็จะยึดหลักของการเก็บข้อมูลไว้ในระบบที่จะมีการเรียกใช้บ่อยที่สุดเพราะจะได้ไม่เสียเวลาไปกับการส่งถ่ายข้อมูลไปมาระหว่างระบบต่าง ๆ ในเครือข่ายคอมพิวเตอร์นั่นเอง และวิธีการแบ่งแยกข้อมูลมีอยู่สามวิธีดังต่อไปนี้

1. การแตกกระจายตามแนวนอน (Horizontal Fragmentation)

เป็นการแบ่งตารางความสัมพันธ์ออกเป็นส่วนๆ โดยที่ตารางความสัมพันธ์ที่ได้ใหม่และของเดิมจะมีโครงสร้างเหมือนกัน แต่ระเบียนข้อมูลที่เก็บจะแตกต่างกันตามเงื่อนไขของการแบ่งแยก เช่นดังในรูปที่ 3.1 เป็นตัวอย่างความสัมพันธ์เดิม เราสามารถแบ่งออกเป็นสองตารางความสัมพันธ์ได้โดยกำหนดให้ตารงความสัมพันธ์หนึ่งเก็บข้อมูลเฉพาะของกรุงเทพฯ และอีกตารงความสัมพันธ์จะเก็บเฉพาะข้อมูลของต่างจังหวัด ซึ่งก็จะได้ตารางความสัมพันธ์ที่แตกกระจายตามแนวนอนสองตัวดัง รูปที่ 3.2

Salesman code	Salesman name	Amount	Area
11111	Somchai	125460	Bangkok
22222	Adisak	524123	Upcountry
33333	Boonsri	248321	Bangkok
44444	Somkid	846521	Bangkok
55555	Paiboon	114256	Upcountry
66666	Wanpen	421535	Upcountry

รูปที่ 3.1 ตัวอย่างความสัมพันธ์

Salesman code	Salesman name	Amount	Area
11111	Somchai	125460	Bangkok
33333	Boonsri	248321	Bangkok
44444	Somkid	84652	Bangkok

Salesman code	Salesman name	Amount	Area
22222	Adisak	524123	Upcountry
55555	Paiboon	114256	Upcountry
66666	Wanpen	421535	Upcountry

รูปที่ 3.2 ผลของการแบ่งตารางความสัมพันธ์ตามแนวนอน

2. การแตกกระจายตามแนวตั้ง (Vertical Fragmentation)

เป็นการแบ่งตารางความสัมพันธ์ออกเป็นส่วนๆ โดยที่โครงสร้างของตารางความสัมพันธ์ใหม่รวมกันจึงจะเท่ากับโครงสร้างของตารางความสัมพันธ์เดิม ซึ่งตารางความสัมพันธ์ใหม่ทั้งหมดจะต้องมีกุญแจนอกที่จะใช้เป็นความสัมพันธ์กับตารางอื่นๆ เพื่อที่จะให้ได้มาซึ่งมโนภาพเชิงรวมเดิมนั่นเอง ตัวอย่างดังแสดงใน รูปที่ 3.3

Salesman code	Salesman name
11111	Somchai
22222	Adisak
33333	Boonsri
44444	Somkid
55555	Paiboon
66666	Wanpen

Salesman code	Amount	Area
11111	125460	Bangkok
22222	524123	Upcountry
33333	248321	Bangkok
44444	846521	Bangkok
55555	114256	Upcountry
66666	421535	Upcountry

รูปที่ 3.3 ผลของการแตกกระจายตามแนวตั้ง

3. การแตกกระจายแบบผสม (Hybrid Fragmentation)

เป็นการแตกกระจายทั้งในแนวดิ่งและแนวนอนร่วมกัน จากตัวอย่างข้างต้น หลังจากที่เราทำการแตกกระจายตามแนวตั้งแล้วตาม รูปที่ 3.3 เราก็สามารถมาทำการแตกกระจายตามแนวนอนได้อีกโดยใช้กฎเกณฑ์เดียวกับตัวอย่างใน รูปที่ 3.2 ซึ่งจะได้เป็นตารางความสัมพันธ์ดัง รูปที่ 3.4

Salesman code	Salesman name
11111	Somchai
22222	Adisak
33333	Boonsri
44444	Somkid
55555	Paiboon
66666	Wanpen

Salesman code	Amount	Area
11111	125460	Bangkok
33333	248321	Bangkok
44444	846521	Bangkok

Salesman code	Amount	Area
22222	524123	Upcountry
55555	114256	Upcountry
66666	421535	Upcountry

รูปที่ 3.4 การแตกกระจายแบบผสม

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

<u>การกำหนดสถานที่เก็บข้อมูล</u> (The Allocation of Fragmentation)

ภายหลังจากการออกแบบลักษณะการแตกกระจายของข้อมูลแล้ว ขั้นตอนต่อไปก็คือการกระจายข้อมูลออกไปเก็บตามสถานีประมวลผลต่าง ๆ โดยเลือกเก็บส่วนต่าง ๆ ของฐานข้อมูลเหล่านั้นไว้ในสถานีประมวลผลที่จะมีการใช้ข้อมูลบ่อยที่สุดเป็นหลักสำคัญ เพื่อประโยชน์ในด้านประสิทธิภาพของการทำงานของระบบในแง่ต่าง ๆ เช่นของการลดปริมานข้อมูลที่จะต้องส่งไปมาระหว่างระบบต่าง ๆ ในเครือข่ายซึ่งจะทำให้เวลาในการตอบสนองน้อยลง(Response time) และผลลัพธ์ของระบบงาน(Througthput)ที่มากที่สุดนั่นเอง ซึ่งการจัดเก็บฐานข้อมูลแบบกระจายนี้มีอยู่ 4 แบบดังนี้

1. ฐานข้อมูลแบบรวมศูนย์ (Centralized Database)

วิธีนี้เป็นวิธีที่จัดเก็บฐานข้อมูลทั้งหมดไว้ที่คอมพิวเตอร์ศูนย์กลางเท่านั้น โดยที่ผู้ใช้ที่อยู่ที่ระบบประมวลผลอื่นในระบบเครือข่ายจะต้องจะต้องติดต่อเข้ามาที่ศูนย์นี้เสมอเมื่อต้องการใช้ข้อมูล

2. ฐานข้อมูลแบบเก็บบางส่วน (Partitioned Database)

วิธีนี้จะเป็นการกระจายฐานข้อมูลเป็นส่วน ๆ ไปเก็บยังสถานีประมวลผลต่าง ๆ กัน โดยที่แต่ละส่วนที่แบ่งออกมาจะไม่มีข้อมูลที่ซ้ำกันเลย ซึ่งการเก็บแบบบางส่วนนี้ จะต้องอาศัยการออกแบบการแตกกระจายของข้อมูลเป็นสำคัญ ไม่ว่าจะเป็นแบบแตกกระจายตามแนวนอน แนวตั้ง หรือแบบผสมก็ตาม สำหรับการแบ่งข้อมูลระบบฐานข้อมูลแบบกระจายอาจจะแบ่งข้อมูลตามการใช้งานของข้อมูล หรือแบ่งตามสภาพทางภูมิศาสตร์ก็ได้

3. ฐานข้อมูลแบบเก็บซ้ำ (Replicated Database)

วิธีนี้ฐานข้อมูลจะถูกสำเนาไปเก็บไว้ที่สถานีประมวลผลตั้งแต่สองสถานีขึ้นไป ซึ่งปกติฐานข้อมูลที่จะถูกเก็บซ้ำนี้จะเป็นฐานข้อมูลที่ทุก ๆ สถานีประมวลผลมีโอกาสที่จะต้องใช้บ่อย ๆ เป็นประจำเท่า ๆ กัน ตัวอย่างเช่นฐานข้อมูลราคาสินค้าจะถูกเรียกใช้ทุกครั้งที่มีการเกิดรายการขายสินค้าเกิดขึ้น เป็นต้น

4. ฐานข้อมูลแบบผสม (Mixed Database)

วิธีนี้แต่ละสถานีประมวลผลจะเก็บส่วนของฐานข้อมูลที่แตกต่างกันออกไป ทั้งนี้ขึ้นอยู่กับรูปแบบของการใช้ฐานข้อมูลนั้น ๆ เช่นถ้าข้อมูลส่วนใดที่ถูกใช้เฉพาะสถานีนั้น ๆ ก็จะถูกแยกเก็บแบบบางส่วน แต่ถ้าเป็นข้อมูลที่ต้องใช้ร่วมกันกับสถานีอื่นก็จะถูกจัดเก็บแบบเก็บซ้ำ เป็นต้น

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

การจัดการรายการแบบกระจาย (The Management of Distributed Transactions)

ในการจัดการรายการของฐานข้อมูลแบบกระจาย จะต้องประสพกับปัญหาต่าง ๆ ในด้านข้อจำกัดของระบบสื่อสารของเครือข่าย เช่น ความน่าเชื่อถือของระบบ การควบคุมภาวะพร้อมกันของระบบ และประสิทธิภาพโดยรวมของระบบ ผลงานวิจัยต่าง ๆ ส่วนใหญ่จะพยายามที่จะแก้ปัญหาต่าง ๆ เหล่านั้น โดยเฉพาะอย่างยิ่งการฟื้นสภาพและการควบคุมภาวะพร้อมกันของระบบฐานข้อมูลแบบกระจาย ซึ่งทั้งสองสิ่งนี้มีความสัมพันธ์เกี่ยวข้องกันอยู่มาก

1. คุณสมบัติพื้นฐานของรายการข้อมูล (Properties of Transactions)

1.1 ความเป็นหนึ่งเดียวของรายการ (Atomicity)

ในขณะที่กำลังทำรายการอยู่ เช่นกำลังปรับปรุงรายการอยู่เกิดหยุดการทำงานลงอย่างกระทันหัน จะต้องไม่ส่งผลกระทบทำให้บางส่วนของรายการปรับปรุงเสร็จสิ้นแล้วแต่บางส่วนยังไม่ได้ปรับปรุง ซึ่งการจะทำให้เกิดลักษณะเช่นนี้ได้โดยการฟื้นสภาพรายการ (Transaction recovery)นั่นเอง กล่าวคือการกระทำต่อรายการจะเสมือนมีเพียงขั้นตอนเดียวซึ่งสั้นมาก และผลลัพธ์ที่ได้ก็จะเป็นเพียงสำเร็จหรือไม่สำเร็จเท่านั้น

1.2 ความยั่งยืนของรายการ (Durability)

หลังจากได้รับการยืนยันว่าการทำรายการสำเร็จ เราก็สามารถที่การันตีได้ว่ารายการที่ทำไปจะไม่สูญหาย ไม่ว่าจะเกิดอะไรขึ้นก็ตาม โดยการทำงานที่เรียกว่า การฟื้นสภาพฐานข้อมูล(Database recovery)

1.3 ผลลัพธ์ลำดับตามการกระทำ (Serializability)

ในกรณีที่รายการถูกใช้โดยผู้ใช้พร้อม ๆ กันหลาย ๆ คน ผลลัพธ์ที่กระทำบนรายการนั้นจะต้องเป็นไปตามกฎก่อนหลัง เช่นมีผู้ใช้สองคนเรียกใช้ข้อมูลรายการเดียวกันเพื่อการปรับปรุง ผลของการปรับปรุงจะเป็นผลของการปรับปรุงครั้งสุดท้ายซึ่งการทำงานนี้จะเรียกว่า การควบคุมภาวะพร้อมกัน(Concurrent control)

1.4 ไม่มีผลกระทบซึ่งกันและกัน (Isolation)

ในกรณีที่บางรายถูกดำเนินการไม่สำเร็จจะต้องไม่ส่งผลกระทบต่อการดำเนินการของรายการอื่น ๆ ซึ่งอาจจะทำให้มีผลกระทบให้ระบบหยุดการดำเนินรายการแบบลูกโซ่ได้

2. เป้าหมายของการจัดการรายการ

ในการปฏิบัติการของผู้จัดการรายการจะมุ่งประเด็นไปที่สามประการหลักที่เกียวข้องกับการทำรายการ คือ ประสิทธิภาพ ความน่าเชื่อถือ และภาวะพร้อมกัน ซึ่งทั้งสามสิ่งนี้มีความสันพันธ์กันอย่างมากโดยที่มีองค์ประกอบย่อยต่าง ๆ คือ การใช้หน่วยประมวลผลและหน่วยความจำกลางอย่างมีประสิทธิภาพ การใช้ปริมาณข่าวสารที่ใช้ในการควบคุมน้อยที่สุด การใช้เวลาในการตอบสนองสั้นที่สุด และท้ายสุดคือต้องมีความน่าเชื่อถือสูง

<u>สภาวะการมองผ่าน</u> (Transparency)

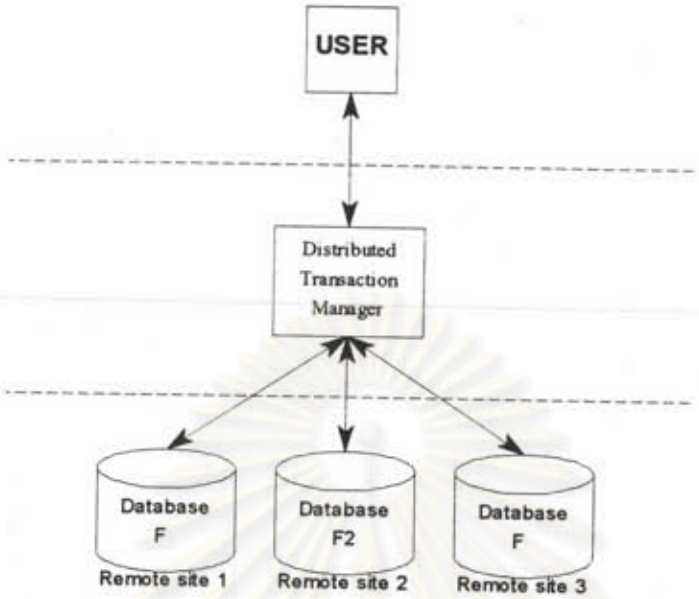
จากที่กล่าวไปแล้วข้างต้นว่าสำหรับฐานข้อมูลแบบกระจายผู้ใช้ไม่จำเป็นต้องรู้ว่ารายการข้อมูลที่ต้องการใช้นั้นถูกเก็บอยู่ที่ใดและเก็บอย่างไร เพราะคุณสมบัติสำคัญของจัดการรายการข้อมูลแบบกระจายที่เรียกว่า "สภาวะการมองผ่าน" นั่นเอง โดยที่ระบบฐานข้อมูลแบบกระจายจะซ่อนความยุ่งยากนี้ไว้ในระบบจัดการรายการข้อมูลแบบกระจาย ซึ่งสภาวะการมองผ่านจะมีอยู่สองลักษณะดังนี้

1. การมองผ่านการแตกกระจาย (Fragmentation Transparency)

จากผลของการออกแบบการแตกกระจายของข้อมูล ซึ่งจะทำให้มีการแยกตารางความสัมพันธ์หนึ่ง ๆออกไปเป็นหลาย ๆตารางความสัมพันธ์ ดังนั้นเพื่อที่จะให้ได้ผลในทางตรงกันข้ามกับการแบ่งแยกก็คือต้องทำการรวมตารางความสัมพันธ์ทั้งหมดที่ถูกแบ่งแยกออกให้เข้ามาเป็นตารางความสัมพันธ์เดิม ทำให้มุมมองจากผู้ใช้เหมือนไม่ได้มีการแบ่งแยกตารางความสัมพันธ์ออกเป็นส่วน ๆและจะเรียกสภาวะนี้ว่า "การมองผ่านการแตกกระจาย" ดังจากตัวอย่างดัง รูปที่ 3.2 เป็นตารางความสัมพันธ์สองตารางที่เกิดจากการแบ่งแยกตามแนวนอนของตารางใน รูปที่ 3.1 แต่เมื่อผ่านกระบวนการมองผ่านการแตกกระจายจะทำให้เสมือนมีเพียงตารางเดียวคือตารางใน รูปที่ 3.1 เท่านั้น

2. การมองผ่านสถานที่ตั้ง (Location Transparency)

หลังจากได้แบ่งแยกตารางความสัมพันธ์เป็นหลาย ๆ ตารางและถูกจัดเก็บในที่ต่าง ๆ กันจะทำให้การเรียกใช้ข้อมูลจำเป็นต้องรู้ว่ารายการข้อมูลที่ต้องการถูกจัดเก็บอยู่ที่ใด แต่ผู้จัดการรายการข้อมูลแบบกระจายจะเป็นตัวจัดการรายการข้อมูลที่กระจายอยู่ที่ระบบต่าง ๆ โดยที่ผู้ที่เรียกใช้ไม่ต้องรับรู้ว่าถูกเก็บไว้ที่ใดในเครือข่ายเสมือนกำลังใช้ข้อมูลท้องถิ่นนั่นเอง เพียงแต่จะใช้เวลามากกว่าเท่านั้นดังแสดงใน รูปที่ 3.5



รูปที่ 3.5 การมองผ่านสถานที่ตั้ง

พิธีการยืนยัน 2 ระยะ (The 2-Phase Commitment Protocol)

หลักการพื้นฐานของการทำงานการยืนยันสองระยะก็ คือจะต้องมีระบบอยู่หนึ่งระบบในเครือข่ายที่ทำหน้าที่พิเศษเป็นผู้ประสานงาน(Coordinator) และระบบอื่น ๆ ที่เหลืออยู่จะมีบทบาทเป็นผู้ร่วมงาน(Participant) โดยที่ระบบที่เป็นผู้ประสานงานจะมีหน้าที่สำคัญที่จะเป็นผู้ตัดสินใจว่าการทำงานในแต่ละครั้งที่กระทำกับรายการข้อมูลจะเป็นการยืนยันความสำเร็จ (Commit) หรือยกเลิกการทำงานนั้น(Abort) ส่วนระบบผู้ร่วมงานที่เหลือจะมีหน้าที่ในการกระทำตามผลของการตัดสินใจของระบบผู้ประสานงาน กล่าวคือการตัดสินใจกระทำต่อรายการข้อมูลใด ๆ จะต้องมีเพียง "ยืนยันการทำงาน" หรือ "ยกเลิกการทำงาน" เพียงหนึ่งเดียวเท่านั้นในระบบ(unique) ซึ่งการทำงานจะแบ่งเป็นสองระยะ โดยที่ระยะที่หนึ่งจะเป็นการทำงานเพื่อให้ได้มาซึ่งผลการตัดสินใจร่วมของระบบ(common decision) ส่วนระยะที่สองเป็นการทำงานตามผลการตัดสินใจในขั้นที่หนึ่งนั่นเอง ซึ่งการทำงานโดยดังแสดงใน รูปที่ 3.5

```
Coordinator: Write "prepare" record in the log;
             Send PREPARE message and activate timeout;

Participant: Wait for PREPARE message;
             If the participant is willing to commit then
               begin
                 Write subtransaction's records in the log;
                 Write "ready" record in the log;
                 Send READY answer message to coordinator;
               end
             else
               begin
                 Write "abort" record in the log;
                 Send ABORT answer message to coordinator;
               end

Coordinator: Wait for ANSWER message (READY or ABORT) from all participants or timeout;
             If time out expired or some answer message is ABORT then
               begin
                 Write "global_abort" record in the log;
                 Send ABORT command message to all participants;
               end
             else (* all answers arrived and were READY *)
               begin
                 Write "global_commit" record in the log;
                 Send COMMIT command message to all participants;
               end

Participant: Wait for command message;
             Write "abort' or "commit" record in the log;
             Send the ACK message to coordinator;
             Execute command;

Coordinator: Wait for ACK message from all participants;
             Write "complete" record in the log;
```

รูปที่ 3.6 ขั้นตอนพื้นฐานของพิธีการยืนยันสองระยะ

ระยะที่ 1 (Phase 1)

ในระยะที่หนึ่งนี้ ระบบผู้ประสานงานจะสั่งให้ระบบผู้ร่วมงานเตรียมพร้อมที่จะปฏิบัติต่อรายการข้อมูลพร้อมทั้งตรวจสอบผลการเตรียมงานเพื่อนำไปใช้ในการตัดสินใจสำหรับการปฏิบัติการในระยะที่สองต่อไป โดยรายละเอียดขั้นตอนการทำงานมีดังนี้

ที่ระบบผู้ประสานงาน (Coordinator)

ขั้นที่ 1 ระบบผู้ประสานงานทำการบันทึกรายการที่ระบุว่า "เตรียมพร้อม" (prepare) ลงในแฟ้มปูม(log file) เป็นการเริ่มต้นการปฏิบัติการ

ขั้นที่ 2 ระบบผู้ประสานงานส่งคำสั่ง "เตรียมพร้อม" ไปที่ระบบผู้ร่วมงาน พร้อมทั้งกระตุ้นการจำกัดเวลาโดยการนับถอยหลังให้ทำงาน เพื่อระบบจะได้รู้ว่าไม่มีการตอบสนองจากระบบปลายทาง

ขั้นที่ 3 รอรับคำตอบจากระบบปลายทางว่าเป็น "เรียบร้อย" หรือ "ยกเลิก" (READY or ABORT)

ที่ระบบผู้ร่วมงาน (Participants)

ขั้นที่ 1 รับคำสั่งเตรียมการ(prepare)จากระบบผู้ควบคุม ถ้าพร้อมที่จะปฏิบัติการต่อรายการข้อมูลให้ไปทำต่อขั้นที่2 แต่ถ้าไม่พร้อมให้ไปทำขั้นที่3

ขั้นที่ 2 บันทึกรายการข้อมูลและรายการควบคุมที่แสดงสถานะภาพว่า "พร้อม" ลงในแฟ้มปูม พร้อมกันนั้นให้ส่งคำตอบว่า "พร้อม" กลับไปให้ระบบผู้ประสานงาน และ กระตุ้นการจำกัดเวลาโดยการนับถอยหลังให้ทำงาน เพื่อระบบจะได้รู้ว่าไม่มีการตอบสนองจากระบบผู้ประสานงานกลับมา แล้วไปทำงานต่อในระยะที่2

ขั้นที่ 3 บันทึกรายการควบคุมที่แสดงสถานะภาพว่า "ยกเลิก" ลงในแฟ้มปูม พร้อมกันนั้นให้ส่งคำตอบว่า "ยกเลิก" กลับไปให้ระบบผู้ประสานงาน แล้วหยุดการทำงาน

<u>ระยะที่ 2</u> (Phase 2)

ในระยะที่ 2 นี้จะเป็นการตัดสินใจรวมว่าจะยืนยันการทำงานหรือจะทำการยกเลิกปฏิบัติการทั้งหมด ซึ่งรายละเอียดขั้นตอนการทำงานดังนี้

ที่ระบบผู้ประสานงาน (Coordinator)

ขั้นที่ 1 ถ้าคำตอบเป็น "ยกเลิก" หรือหมดเวลาที่รอคอยจากบางระบบให้ไปทำบันทึกรายการ "ยกเลิกการทำงานทั้งหมด" (global_abort) ลงในแฟ้มปูม และส่งคำสั่ง "ยกเลิกการทำงาน" ไปให้กับทุกๆระบบผู้ร่วมงานที่ได้รับคำตอบว่าเรียบร้อยมาก่อนหน้านี้ และหยุดการทำงาน

ขั้นที่ 2 ถ้าหลังจากได้รับคำตอบจากทุกๆระบบว่า "เรียบร้อย" ให้ทำการบันทึกรายการลงในแฟ้มปูมว่า "ยืนยันการทำงานรวม" พร้อมทั้งส่งคำสั่ง "ยืนยันการทำงาน" ไปให้กับทุกๆระบบเพื่อที่จะปฏิบัติการกับรายการข้อมูลจริง

ขั้นที่ 3 รอผลการทำงาน ถ้าผลการทำงานของทุกๆระบบเสร็จสมบูรณ์ให้บันทึกรายการลงในแฟ้มปูมว่า "เสร็จสมบูรณ์" (complete)

ที่ระบบผู้ร่วมงาน (Participants)

ขั้นที่ 1 รอรับคำสั่งต่อมา ถ้าคำสั่งเป็น "ยกเลิก" หรือหมดเวลาที่รอคอยให้ทำขั้นที่2 แต่ถ้าคำสั่งเป็น "ยืนยัน" ให้ไปทำขั้นที่3

ขั้นที่ 2 บันทึกรายการควบคุมที่แสดงสถานะภาพว่า "ยกเลิก" ลงในแฟ้มปูม และส่งข่าวสารเพื่อให้รู้ว่าปฏิบัติการเสร็จสิ้น(ACK message) แล้วหยุดการทำงาน

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ขั้นที่ 3 บันทึกรายการควบคุมที่แสดงสถานะภาพว่า "ยืนยัน" ลงในแฟ้มปูม และส่งข่าวสารเพื่อให้รู้ว่าปฏิบัติการเสร็จลุล่วง(ACK message) แล้วเริ่มปฏิบัติการต่อรายการข้อมูลจนเสร็จสิ้น แล้วหยุดการทำงาน

<u>การรับมือกับภาวะล้มเหลว</u> (Failures)

ในระหว่างที่กำลังดำเนินการของ พิธีการสองระยะ จะเห็นได้ว่ามีขั้นตอนต่าง ๆ มากมายพอสมควร ดังนั้นการเกิดภาวะล้มเหลวในขั้นตอนใดขั้นตอนหนึ่งหรือระหว่างขั้นตอนต่าง ๆ จึงมีโอกาสเกิดขึ้นได้เสมอ ซึ่งภาวะล้มเหลวต่าง ๆ อาจเกิดได้ดังต่อไปนี้

1. สถานีประมวลผลล้มเหลว (Site failures)

1.1 ที่ระบบผู้ร่วมงานล้มเหลว ก่อนที่จะบันทึกรายการแสดงสถานะ "พร้อม" ลงในแฟ้มปูม ซึ่งในกรณีนี้ที่ระบบผู้ประสานงานก็จะไม่ได้รับคำตอบจากระบบผู้ร่วมงานจนทำให้เกิดภาวะหมดเวลาที่รอคอย แล้วระบบผู้ประสานงานก็จะดำเนินการในขั้นต่อมา คือจะทำการสรุปการตัดสินใจรวมเป็น <u>ยกเลิกการทำงานทั้งหมด</u> แล้วสั่งให้ระบบผู้ร่วมงานที่ผ่านการเตรียมการมาแล้วทำการยกเลิกการทำงานกับรายการข้อมูลด้วย ส่วนที่ระบบผู้ร่วมงานที่ล้มเหลวภายหลังจากได้ฟื้นสภาพขึ้นมาแล้วจะไม่มีการดำเนินการใด ๆ ทั้งสิ้นเพราะไม่มีการดำเนินการใดก่อนที่ระบบจะล้มเหลว

1.2 ที่ระบบผู้ร่วมงานล้มเหลว <u>หลัง</u> จากที่ได้บันทึกรายการแสดงสถานะ "พร้อม" ลงในแฟ้มปูม ในกรณีนี้ที่ระบบผู้ประสานงานจะทำงานเช่นเดียวกับกรณีที่ 1 แต่ที่ระบบผู้ร่วมงานที่ล้มเหลวเอง ภายหลังจากระบบกลับมาใช้งานได้ดังเดิมจะต้องดำเนินการยกเลิกการปฏิบัติการกับรายการที่ทำค้างไว้ก่อนที่ระบบจะล้มเหลว

1.3 ที่ระบบผู้ประสานงานล้มเหลว หลังจากที่ได้บันทึกรายการแสดงสถานะ "เตรียมพร้อม" แต่ก่อนที่จะบันทึกรายการแสดงสถานะ "ยืนยันรวม" หรือ "ยกเลิกรวม" ลงในแฟ้มปูม ในกรณีนี้ระบบผู้ร่วมงานทั้งหมดที่ได้ผ่านขั้นตอนการเตรียมพร้อมมาแล้วจะต้องรอให้ระบบผู้ประสานงานฟื้นสภาพขึ้นมา แล้วหลังจากนั้นระบบผู้ประสานงานก็จะเริ่มดำเนินการที่ค้างอยู่จากจุดเริ่มต้นใหม่อีกครั้งหนึ่ง โดยที่การทำงานซ้ำในครั้งนี้อาจเป็นการ "ยืนยันรวม" หรือ "ยกเลิกรวม" ก็ได้แต่ต้องรอจนหมดเวลาที่รอคอยแล้วให้ทำการยกเลิกรายการเอง

I16575660

1.4 ที่ระบบผู้ประสานงานล้มเหลว หลังจากที่ได้บันทึกรายการแสดงสถานะ "ยืนยันรวม" หรือ "ยกเลิกรวม" แต่ก่อนที่จะบันทึกรายการแสดงสถานะ "สมบูรณ์" ลงในแฟ้มปูม ในกรณีนี้ระบบผู้ร่วมงานทั้งหมดจะต้องรอให้ระบบผู้ประสานงานฟื้นสภาพขึ้นมาแล้วหลังจากนั้นระบบผู้ประสานงานจะเริ่มส่งคำสั่ง "ยืนยันรวม" หรือ "ยกเลิกรวม" ไปให้ระบบผู้ร่วมงานทั้งหมดอีกครั้งหนึ่ง เพื่อทำให้การปฏิบัติการครบถ้วนสมบูรณ์ แต่ต้องรอจนหมดเวลาที่รอคอยแล้วให้ทำการยกเลิกรายการเอง

2. ข่าวสารควบคุมสูญหาย (Lost message)

2.1 ไม่มีคำตอบใด ๆ จากระบบผู้ร่วมงาน ในกรณีเช่นนี้จะทำให้เกิดภาวะหมดเวลาในการรอคอย ซึ่งผลจะทำให้ระบบโดยรวมเกิดภาวะยกเลิกการทำงาน จะสังเกตุได้ว่ากรณีนีจะเป็นการปฏิบัติการที่มองเฉพาะที่ระบบผู้ประสานงานเท่านั้น โดยจะมีผลเท่ากับการที่ระบบผู้ร่วมงานล้มเหลวด้วย แต่ถ้ามองจากมุมของระบบผู้ร่วมงานซึ่งระบบไม่ได้ล้มเหลวจริงจะมองเห็นเสมือนระบบผู้ประสานงานล้มเหลวเสียเอง ดังนั้นระบบผู้ร่วมงานนี้ก็จะทำการยกเลิกการปฏิบัติการต่อรายการข้อมูลดังกล่าวด้วยเช่นกัน

2.2 คำสั่งเตรียมพร้อมสูญหาย ในกรณีนี้จะเข้าข่ายเดียวกันกันกรณีแรกคือจะไม่ได้รับคำตอบกลับมาจากระบบผู้ประสานงานนั่นเอง

2.3 คำสั่ง "ยืนยันรวม" หรือ "ยกเลิกรวม" สูญหาย ลักษณะเช่นนี้จะทำให้ระบบผู้ร่วมงานเกิดภาวะหมดเวลาในการรอคอย ซึ่งระบบผู้ร่วมงานอาจทำการส่งคำขออีกจนกระทั่งครบจำนวนที่กำหนดไว้หรือจนกระทั่งได้รับคำสั่งมา โดยที่ถ้ากระทำการขอซ้ำจนครบจำนวนที่กำหนดไว้ แล้วยังไม่มีคำตอบจากระบบผู้ประสานงานก็ให้ทำการยกเลิกการปฏิบัติการต่อรายการข้อมูลนั้นได้เลย แต่ถ้าได้รับคำตอบมาก็ให้ปฏิบัติการต่อไปตามคำสั่งที่ได้

2.4 การตอบสนองด้วย "ACK" จากระบบผู้ร่วมงานสูญหาย ในกรณีนี้ที่ระบบผู้ประสานงานจะเกิดภาวะหมดเวลาในการรอคอย ซึ่งระบบผู้ประสานงานอาจทำการส่งคำสั่ง "ยืนยันรวม" หรือ "ยกเลิกรวม" ไปอีก จนกระทั่งครบจำนวนที่กำหนดไว้ หรือจนกระทั่งได้รับการตอบสนอง โดยที่ถ้ากระทำการส่งคำสั่งซ้ำจนครบจำนวนที่กำหนดไว้แล้วยังไม่มีการตอบสนองจากระบบผู้ร่วมงานก็ให้ทำการ "ยกเลิกรวม" ต่อรายการข้อมูลนั้นกับระบบผู้ร่วมงานที่เหลือได้เลย แต่ถ้าได้รับการตอบสนองมาก็ให้ปฏิบัติการต่อไปตามปกติ

การควบคุมภาวะพร้อมกัน (Concurrency control for Distributed Transactions)

หลักการพื้นฐานที่จะใช้ในการควบคุมภาวะพร้อมกันก็คือการล็อค(Lock)ต่อสิ่งที่ต้องการใช้ ซึ่งจะใช้พิธีการล็อคสองระยะ(2-phase-locking) โดยจะพิจารณาจาก พิธีการล็อคสองระยะ ของระบบฐานข้อมูลแบบรวมศูนย์ก่อน แล้วจะขยายผลไปยังพิธีการล็อคสองระยะของระบบฐานข้อมูลแบบกระจายต่อไป

1. การควบคุมภาวะพร้อมกันของฐานข้อมูลแบบรวมศูนย์

แนวความคิดพื้นฐานของการล็อค คือ เมื่อต้องการใช้รายการข้อมูลใดต้องล็อคข้อมูลนั้นแต่ถ้ารายการข้อมูลนั้นถูกล็อคโดยผู้ใช้รายอื่นแล้วจะต้องรอจนกว่ารายการข้อมูลนั้นได้รับการปลดปล่อย(unlock)เป็นอิสระเสียก่อน ปกติการล็อคสามารถจะมีได้หลาย ๆ ลักษณะ(mode)ตามความต้องการในการใช้ข้อมูล เช่น แบบใช้ร่วมได้(shared) สำหรับเมื่อต้องการอ่านอย่างเดียว(read only) หรือแบบใช้ผู้เดียว(exclusive) สำหรับเมื่อต้องการปรับปรุงรายการ(update) ดังนั้นกฎเกณฑ์ของการล็อคในกรณีที่ล็อคแบบใช้ร่วมสามารถ ล็อคได้หลายผู้ใช้ แต่ถ้าเป็นการล็อคแบบใช้ผู้เดียวจะต้องไม่มีผู้ใดล็อคอยู่และเมื่อล็อคแล้วจะไม่มีใครจะใช้ได้

ในพิธีการล็อคสองระยะนี้จะสอดคล้องกับกฎของการเป็นหนึ่งเดียว ของของคุณสมบัติรายการข้อมูล กล่าวคือเมื่อเริ่มปฏิบัติการจะต้องล็อครายการข้อมูลแบบใช้ผู้เดียวไปจนกระทั่งสิ้นสุดการปฏิบัติการ ดังแสดงใน รูปที่ 3.6

```
(Begin application)
Begin transaction
Acquire locks before reading or writing
Other application operation
...
...
...
Commit
Release lock(unlock)
(End application)
```

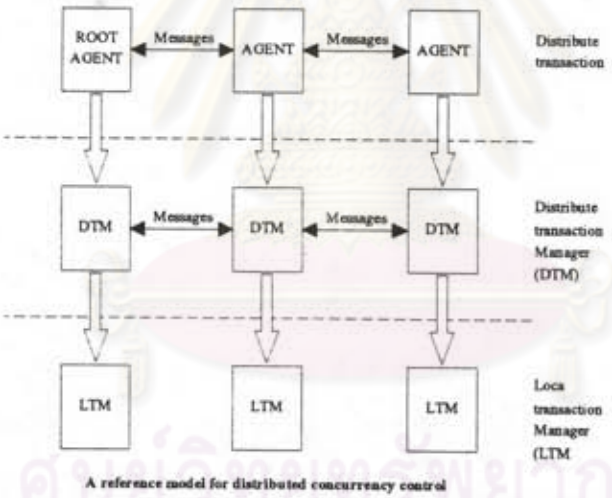
รูปที่ 3.7 แสดงพิธีการล็อคสองระยะ

2. การควบคุมภาวะพร้อมกันของฐานข้อมูลแบบกระจาย

จากความสามารถในการควบคุมภาวะพร้อมกันของระบบฐานข้อมูลแบบรวมศูนย์ เราสามารถนำมาใช้ในการทำงานสำหรับฐานข้อมูลแบบกระจายได้ด้วย โดยส่วนที่ทำงานของระบบรวมศูนย์จะถือว่าเป็นการทำงานในระดับท้องถิ่นเรียกว่า ผู้จัดการรายการ

ท้องถิ่น(Local Transaction Manager LTM) โดยหน้าที่ส่วนหนึ่งก็คือจะต้องทำงานในการล็อคและปลดปล่อยรายการข้อมูลในระดับท้องถิ่นนั่นเอง โดยจะเป็นการรับคำสั่งมาจากผู้จัดการรายการกระจาย(Distributed Transaction Manager DTM) ซึ่งจะทำงานในระดับรวม(Global) ดังแสดงใน รูปที่ 3.7 และประการสำคัญก็คือการจัดการรายการข้อมูลแบบกระจายจะต้องอยู่บนกฎเกณฑ์เดียวกันกับแบบรวมศูนย์ กล่าวคือในการล็อคแบบใช้ร่วมผู้ใช้สามารถล็อคได้หลาย ๆ คน แต่ถ้าเป็นการล็อคแบบใช้ผู้เดียวจะต้องไม่มีผู้ใดล็อคอยู่และเมื่อล็อคแล้วจะไม่มีใครสามารถใช้ได้นั่นเอง ดังนั้นพิธีการล็อคสองระยะของแบบกระจายก็จะเหมือนกับแบบรวมศูนย์นั่นเอง

ในการทำงานกับรายการข้อมูลที่มีการเก็บซ้ำกันหลาย ๆ ที่นี้ กลไกในการล็อครายการข้อมูลจะทำประหนึ่งเหมือนกับมีรายการข้อมูลอยู่เพียงชุดเดียว แต่ผู้จัดการรายการท้องถิ่นจะไม่รับรู้ถึงการมีรายการข้อมูลที่เหมือนกันซึ่งได้เก็บไว้ในระบบอื่น ๆ ดังนั้นจึงเป็นหน้าที่ของผู้จัดการรายการระดับรวมในแต่ละระบบจะต้องประสานงานกันในการทำงานตามพิธีการล็อคสองระยะ เพื่อให้ได้มาซึ่งผลของการล็อคที่เป็นหนึ่งเดียว



รูปที่ 3.8 การควบคุมภาวะพร้อมกันแบบกระจาย

3. การตรวจจับภาวะล็อคค้าง (Deadlock detection)

การเกิดล็อคค้างในระบบฐานข้อมูลแบบกระจาย จะเป็นสถานะการณ์ที่เรียกว่า "การรอเป็นวงกลม"(Circular-waiting) ซึ่งเหตุการณ์เช่นนี้จะเกิดได้กับการทำรายการหลาย ๆ รายในขณะเดียวกัน ไม่ใช่เพียงคู่ใดคู่หนึ่งเท่านั้น ลักษณะพื้นฐานของการเกิดล็อคค้างก็คือมีผู้ทำรายการหนึ่งได้ทำการล็อครายการข้อมูลหนึ่งไว้ และในขณะเดียวกันก็ต้องการที่จะใช้อีกรายการข้อมูลซึ่งถูกผู้ใช้อีกรายหนึ่งล็อคไว้แล้ว ทำให้ผู้ใช้รายนี้จะต้องรอจนกระทั่งผู้ใช้ที่ล็อครายการนี้ปลดปล่อยเป็นอิสระ