บทที่ 5
สรุปผลการวิจัย และข้อเสนอแนะ

5.1 สรุปผลการวิจัย

เมื่อมีการสอบถาม (Query) ที่ต้องการข้อมูลจากฐานข้อมูล (Database) สิ่งที่เป็นตัวกลางที่คอยทำหน้าที่ให้ได้ผลลัพธ์คือ ระบบจัดการฐานข้อมูล (DBMS) ซึ่งขั้นตอนที่ระบบจัดการฐานข้อมูลมีหลายขั้นตอนด้วยกัน โดยมีขั้นตอนที่สำคัญขั้นตอนหนึ่งนั่นคือ ขั้นตอนการทำงานของออพติไมเซอร์ (Optimizer) ซึ่งหน้าที่ของออพติไมเซอร์คือการนำแผนประมวลผล (Execution Plan) หลายๆแผน แล้วทำการเลือกแผนประมวลผลที่มีประสิทธิภาพที่ดีที่สุดเพื่อที่จะนำไปประมวลผล (Execute) เพื่อที่จะนำผลลัพธ์ออกมาจากฐานข้อมูล สิ่งหนึ่งที่ออพติไมเซอร์นำมาพิจารณาในการเลือกแผนประมวลผลที่ดีที่สุดคือ จำนวนแถว (Cardinality) ซึ่งวิธีการคำนวณจำนวนแถวแบบเอ็กแซกท์ (Exact Cardinality Computing) ก็เป็นอีกเทคนิคหนึ่งที่ออพติไมเซอร์เลือกใช้แต่การคำนวณจำนวนแถวใช้เวลาในการวิเคราะห์นานกว่าเทคนิคอื่นๆ แต่ก็มีประสิทธิภาพมากกว่าเทคนิคอื่นๆ เช่นกัน จึงมีการคิดค้นวิธีต่างๆที่จะช่วยให้เวลาในการคำนวณจำนวนแถวแบบเอ็กแซกท์ใช้เวลาน้อยที่สุดและมีประสิทธิภาพมากที่สุดเท่าที่จะเป็นไปได้

การแปลงการสอบถาม (Transformation) ก็เป็นอีกวิธีหนึ่งที่จะช่วยให้ประสิทธิภาพของการคำนวณจำนวนแถวแบบเอ็กแซกท์ดีขึ้นแต่ทั้งนี้ต้องขึ้นอยู่กับปัจจัยและตัวแปรต่างๆ ของฐานข้อมูลด้วยเช่นกัน

เทคนิคการแปลงการสอบถามที่ใช้ในงานวิทยานิพนธ์ฉบับนี้ใช้เทคนิคการแปลงการสอบถามที่ชื่อว่าคอสต์เบส (Cost-Based Transformation) ซึ่งได้นำเทคนิคย่อยของการแปลงการสอบถามแบบคอสต์เบสมีด้วยกัน 4 เทคนิคด้วยกันคือ

1. Subquery Unnesting
2. Group-By and Distinct View Merging
3. Join Predicate Pushdown

4. Join Factorization

จากการนำเทคนิคการแปลงการสอบถามทั้ง 4 รูปแบบมาช่วยในการคำนวณจำนวนแถวแบบเอ็กแซกท์พบว่าผลการทดลองที่เป็นการเปรียบเทียบระหว่างเวลาและเซ็ตข้อมูล (Dataset) ที่มีทั้งเซ็ตข้อมูลที่มีขนาดเพิ่มขึ้นและเซ็ตข้อมูลที่มีขนาดลดลงนั้นทำให้ทราบว่าแต่ละรูปแบบมีลักษณะที่แตกต่างกันออกไป

การแปลงการสอบถามในรูปแบบ Subqueruy Unnesting ต้องดูการสบถามย่อยที่มีความสัมพันธ์กับการสอบถามย่อยที่ซ้อนกันอยู่ในการสอบถามหลัก ซึ่งการแปลงการสอบถามจะมีประสิทธิภาพที่ดีกว่าแบบไม่แปลงการสอบถามก็ต่อเมื่อการสอบถามย่อยที่นำมาไว้ในวิว (View) มีผลลัพธ์ของข้อมูลมากๆ

การแปลงการสอบถามแบบ Group-by and Distinct View Merging นั้นจากผลการทดลองจะเห็นได้ว่าเหมาะสำหรับฐานข้อมูลที่มีขนาดใหญ่ เพราะที่เซ็ตข้อมูลที่ 6000% แบบแปลงการสอบถามทำงานได้ดีกว่าแบบไม่แปลงการสอบถาม

การแปลงการสอบถามรูปแบบ Join Predicate Pushdown เหมาะสำหรับฐานข้อมูลที่มีขนาดไม่ใหญ่ โดยจากผลการทดลองที่เซ็ตข้อมูล 12% และ 50% แบบแปลงการสอบถามทำงานได้ดีกว่าแบบไม่แปลงการสอบถามแต่ตั้งแต่ 100% ขึ้นไปแบบไม่แปลงการสอบถามกลับทำงานดีกว่า

การแปลงการสอบถามรูปแบบ Join Factorization การแปลงการสอบถามของรูปแบบนี้เหมาะสำหรับฐานข้อมูลขนาดใหญ่ โดยจากการทดลองเซ็ตข้อมูลที่ 100% ขึ้นไปแบบแปลงการสอบถามทำงานได้ดีกว่าแบบไม่แปลงการสอบถามอย่างชัดเจน

ในวิทยานิพนธ์ฉบับนี้ได้มุ่งเน้นในการศึกษาผลกระทบของการแปลงการสอบถามของอินพุตที่มีชื่อว่า เวิร์คโหลด (Workload) สำหรับการคำนวณจำนวนแถวแบบเอ็กแซกท์ ซึ่งผลกระทบของการแปลงการสอบถามของแต่ละรูปแบบก็ได้มีลักษณะที่แตกต่างกันออกไปดังที่ได้กล่าวมาแล้วข้างต้น ซึ่งแต่ละรูปแบบก็เหมาะสมสำหรับลักษณะของฐานข้อมูลที่แตกต่างกันไปดังที่ได้แสดงในผลการทดลองในบทที่ 4 ซึ่งสามารถที่จะช่วยให้ออพติไมเซอร์สามารถทำงานในการคำนวณจำนวนแถวได้อย่างมีประสิทธิภาพมากยิ่งขึ้น

5.2 ข้อเสนอแนะ

จากงานวิจัยได้พบวิธีที่จะช่วยให้พัฒนาการทำงานของออพติไมเซอร์สำหรับการคำนวณจำนวนแถวแบบเอ็กแซกท์ คือ การศึกษาผลกระทบของเวิร์คโหลดแบบแบทช์ (Batch-workload) หรือทำการศึกษาการดีเลย์ของการประมวลผลการสอบถาม ซึ่งการศึกษาทั้งสองชนิดที่กล่าวมาเป็นตัวช่วยในการเพิ่มประสิทธิภาพให้แก่การแปลงการสอบถาม