

ห้องสมุดงานวิจัย สำนักงานคณะกรรมการวิจัยแห่งชาติ



249444

การแปลงการสอบถามเข้ารับการกำหนดจำนวนแอมป์แบบอิเล็กทรอนิกส์

ปวีรพรต ขวัญกล้าด

วิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมคอมพิวเตอร์

บัณฑิตวิทยาลัย
มหาวิทยาลัยเชียงใหม่
พฤษภาคม 2554

๒๐๐๒๕๔๕๕๔

ห้องสมุดงานวิจัย สำนักงานคณะกรรมการการวิจัยแห่งชาติ



249444

การแปลงการสอบถามสำหรับการคำนวณจำนวนแถวแบบเอ็กแซกท์



ปริวรรต ขวัญกลัด

วิทยานิพนธ์นี้เสนอต่อบัณฑิตวิทยาลัยเพื่อเป็นส่วนหนึ่ง
ของการศึกษาตามหลักสูตรปริญญา
วิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมคอมพิวเตอร์

บัณฑิตวิทยาลัย
มหาวิทยาลัยเชียงใหม่
พฤษภาคม 2554

กิตติกรรมประกาศ

วิทยานิพนธ์นี้สำเร็จลงได้ด้วยความกรุณาจาก ผู้ช่วยศาสตราจารย์ ดร. จักรพงษ์ นาทวิชัย อาจารย์ที่ปรึกษาวิทยานิพนธ์ผู้ซึ่งกรุณาให้ความรู้ คำแนะนำ คำปรึกษา และตรวจแก้ไขจน วิทยานิพนธ์เสร็จสมบูรณ์ ผู้เขียนขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

ขอกราบขอบพระคุณคณะกรรมการสอบทุกท่าน ที่รับเป็นกรรมการสอบวิทยานิพนธ์ และ ให้คำปรึกษาเป็นอย่างดี

ขอกราบขอบพระคุณอาจารย์ประจำภาควิชาวิศวกรรมคอมพิวเตอร์ที่กรุณาให้คำแนะนำ และขอกราบขอบพระคุณเจ้าหน้าที่ห้องธุรการประจำภาควิชาวิศวกรรมคอมพิวเตอร์ทุกท่านที่ให้ความช่วยเหลือคอยช่วยเหลือในด้านงานเอกสารจนวิทยานิพนธ์นี้เสร็จสมบูรณ์

ขอกราบขอบพระคุณครอบครัวของผู้เขียน ที่คอยสนับสนุนการศึกษาของผู้เขียน อีกทั้งยังเป็นกำลังใจให้กับผู้เขียนเมื่อเวลาที่มีปัญหา

ท้ายสุดนี้ หากมีสิ่งใดขาดตกบกพร่อง หรือมีข้อผิดพลาดประการใด ผู้เขียนขอภัยเป็นอย่างสูงในความผิดพลาดและข้อบกพร่องนั้น และผู้เขียนหวังว่าวิทยานิพนธ์นี้จะมีประโยชน์บ้าง ไม่มากก็น้อย สำหรับหน่วยงานที่เกี่ยวข้อง ตลอดจนผู้สนใจที่จะศึกษารายละเอียดเกี่ยวกับ วิทยานิพนธ์นี้ต่อไป

ปวีรรต ขวัญกล้า

ชื่อเรื่องวิทยานิพนธ์	การแปลงการสอบถามสำหรับการคำนวณจำนวนแถวแบบ เอ็กแซกท์
ผู้เขียน	นายปวิวรรต ขวัญกลัด
ปริญญา	วิศวกรรมศาสตรมหาบัณฑิต (วิศวกรรมคอมพิวเตอร์)
อาจารย์ที่ปรึกษาวิทยานิพนธ์	ผศ.ดร.จักรพงษ์ นาทวิชัย

บทคัดย่อ

249444

ในปัจจุบันฐานข้อมูลได้เข้ามามีบทบาทต่อองค์กรทั้งเล็กและใหญ่ ซึ่งเป็นตัวที่จะช่วยให้
องค์กรมีประสิทธิภาพในการทำงานมากขึ้น โดยฐานข้อมูลจำเป็นต้องมีตัวกลางที่คอยประมวลผล
ระหว่างผู้ใช้และฐานข้อมูล ซึ่งก็คือระบบจัดการฐานข้อมูล ในระบบจัดการฐานข้อมูลมีขั้นตอน
การทำงานมากมายหลายขั้นตอน ขั้นตอนที่สำคัญขั้นตอนหนึ่งคือ การเลือกแผนประมวลผล ซึ่งตัว
ที่ทำหน้าที่พิจารณาเลือกแผนประมวลผลคือ ออปติไมเซอร์

ออปติไมเซอร์มีหน้าที่ในการเลือกแผนประมวลผลที่ดีที่สุดเพื่อที่จะทำการประมวลผล
ผลลัพธ์ออกมา โดยสิ่งสำคัญที่ออปติไมเซอร์นำมาพิจารณาในการเลือกแผนประมวลผลคือ จำนวน
แถวของผลลัพธ์ ซึ่งเทคนิคการคำนวณจำนวนแถวแบบเอ็กแซกท์เป็นวิธีการหาจำนวนแถวที่มี
ความแม่นยำและมีประสิทธิภาพมากที่สุด

ในวิทยานิพนธ์ฉบับนี้ได้้นำการแปลงการสอบถามมาประยุกต์ใช้กับการคำนวณจำนวน
แถวแบบเอ็กแซกท์ โดยเป็นเทคนิคการแปลงการสอบถามที่มีชื่อว่า คอสต์เบส เพื่อที่จะ
ทำการศึกษาว่าประสิทธิภาพของการคำนวณจำนวนแถวเปลี่ยนแปลงไปอย่างไร ซึ่งขั้นตอนวิธีใน
งานวิจัยนี้สามารถบอกได้ว่าเทคนิคการแปลงการสอบถามรูปแบบใดที่เหมาะสมสำหรับฐานข้อมูล
ชนิดใด เพื่อที่จะทำให้การคำนวณจำนวนแถวแบบเอ็กแซกท์มีประสิทธิภาพมากขึ้น

Thesis Title Query Transformation for Exact Cardinality Computing
Author Mr. Prariwat Kwanglat
Degree Master of Engineering (Computer Engineering)
Thesis Advisor Asst. Prof. Dr. Juggapong Natwichai

ABSTRACT

249444

In general, query optimizers rely on a cost model to choose an appropriate query execution plan for the given queries. An important key parameter of the cost estimation is the cardinality of sub-expressions of the queries. Traditionally, the optimizers may use the estimation cardinality techniques, which can lead to the estimation errors, and hence the poor execution plans. The exact cardinality approach can be applied to resolve such problem, though its computational expense can be costly. A possible way to improve the efficiency is the query transformation since it can provide the alternation to the optimizers. In this paper, we focus on investigation at the effects of the query transformation to the exact cardinality computing processes. The query transformation techniques to be considered in our work are the traditional but widely applied techniques, i.e. subquery unnesting, group-by view merging, join factorization, and join predicate pushdown. The experiment results on the real-life datasets have been presented to validate such proposed work.

สารบัญ

หน้า

กิตติกรรมประกาศ	ค
บทคัดย่อภาษาไทย	ง
บทคัดย่อภาษาอังกฤษ	จ
สารบัญตาราง	ช
สารบัญภาพ	ฉ
บทที่ 1 บทนำ	1
1.1 ที่มาและปัญหา	1
1.2 แนวทางการแก้ปัญหา	11
1.3 สรุปสาระสำคัญจากเอกสารที่เกี่ยวข้อง	13
1.4 วัตถุประสงค์ของการวิจัย	13
1.5 ประโยชน์ที่ได้รับจากการศึกษา เชิงทฤษฎี และ/หรือเชิงประยุกต์	13
1.6 ขอบเขตการทำวิจัย	13
1.7 ระเบียบวิธีวิจัย	14
1.8 เครื่องมือในการพัฒนา	14
บทที่ 2 การคำนวณจำนวนแถวแบบเอ็กแซกท์	15
2.1 การคำนวณจำนวนแถวแบบเอ็กแซกท์	15
2.2 การปรับปรุงประสิทธิภาพการสอบถาม	22
บทที่ 3 การประยุกต์การแปลงการสอบถามสำหรับการคำนวณจำนวนแถวแบบเอ็กแซกท์	24
3.1 การประยุกต์การแปลงการสอบถามสำหรับการคำนวณจำนวนแถวแบบเอ็กแซกท์	24
3.2 การแปลงการสอบถามสำหรับการคำนวณจำนวนแถวแบบเอ็กแซกท์	29

บทที่ 4 การทดลองและผลการทดลอง	38
4.1 ข้อมูลที่ใช้ในการทดลอง	39
4.2 ผลการทดลอง	44
4.2.1 รูปแบบ Subquery Unnesting	44
4.2.2 รูปแบบ Group-by and Distinct View Merging	52
4.2.3 รูปแบบ Join Predicate Pushdown	55
4.2.4 รูปแบบ Join Factorization	61
4.3 สรุปผลการทดลอง	67
บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ	68
5.1 สรุปผลการวิจัย	68
5.2 ข้อเสนอแนะและงานวิจัยในอนาคต	70
เอกสารอ้างอิง	71
ประวัติผู้เขียน	73

สารบัญตาราง

ตาราง		หน้า
1.1	ข้อมูลของตาราง Student	3
1.2	ข้อมูลของตาราง Enrolment	4
1.3	ข้อมูลของตาราง Course	4
1.4	ลำดับการทำงานของออฟติไมเซอร์	9
4.1	ค่าประสิทธิภาพของอันเนตติงแบบไม่แปลงการสอบถาม	48
4.2	ค่าประสิทธิภาพของอันเนตติงแปลงการสอบถาม	49

สารบัญภาพ

รูป	หน้า	
1.1	สกีมาของฐานข้อมูลตัวหนึ่ง	2
1.2	การสอบถามข้อมูล	5
1.3	ลำดับการทำงานของ การสอบถาม	5
1.4	การประมาณค่าจำนวนแถว	6
1.5	สกีมา (Schema) ของการส่งสินค้า	7
1.6	การสอบถามข้อมูล	8
1.7	โครงสร้างการหาค่าจำนวนแถวแบบเอ็กแซกท์	11
1.8	การแปลงการสอบถาม	12
2.1	โครงสร้างการหาค่าจำนวนแถว	16
2.2	เงื่อนไขที่มีความคล้ายคลึงกัน	17
2.3	เคสคิวรี	17
2.4	ขั้นตอนการสร้างเคสคิวรี	18
2.5	การเลือกเงื่อนไขมาทำเคสคิวรี	19
2.6	ขั้นตอนการเลือกเคสคิวรี	20
2.7	ขั้นตอนวิธีของครอเวอร์ริงคิวรีออปติไมเซชัน	21
2.8	ขั้นตอนวิธีการคำนวณจำนวนแถวแบบเอ็กแซกท์	22
3.1	เวิร์คโหลด	26
3.2	การเลือกแคตติเคตเจเนเรชัน	27
3.3	เคสคิวรี	28
3.4	การแปลงการสอบถามจากเวิร์คโหลด	30
3.5	การสอบถามแบบ Unnesting	31
3.6	การแปลงการสอบถามแบบ Unnesting	32
3.7	การแปลงการสอบถามรูปแบบ Group-by and Distinct View Merging	33
3.8	การสอบถามแบบ Join Predicate Pushdown	34
3.9	การแปลงการสอบถามแบบ Join Predicate Pushdown	34
3.10	การสอบถามแบบ Join Factorization	35

3.11	การแปลงการสอบถามแบบ Join Factorization	36
3.12	แคณดิเคตเจเนเรชัน	36
4.1	สกีมาของฐานข้อมูล	39
4.2	การสอบถามแบบ Subquery Unnesting	45
4.3	การแปลงการสอบถามแบบ Subquery Unnesting	45
4.4	แผนประมวลผลแบบไม่แปลงการสอบถามของ Subquery Unnesting	46
4.5	แผนประมวลผลแบบแปลงการสอบถามของ Subquery Unnesting	47
4.6	ขั้นตอนแคณดิเคตเจเนเรชัน Subquery Unnesting	49
4.7	ขั้นตอนแคณดิเคตเจเนเรชัน Subquery Unnesting	50
4.8	กราฟผลการทดลอง Subquery Unnesting	51
4.9	การแปลงการสอบถามรูปแบบ Group-by and Distinct View Merging	52
4.10	แผนประมวลผลแบบแปลงการสอบถามของ Group-by and Distinct View Merging	53
4.11	ขั้นตอนแคณดิเคตเจเนเรชันของ Group-by and Distinct View Merging	54
4.12	กราฟผลการทดลอง Group-by and Distinct View Merging	54
4.13	การสอบถามแบบ Join Predicate Pushdown	56
4.14	การแปลงการสอบถามแบบ Join Predicate Pushdown	56
4.15	แผนประมวลผลแบบไม่แปลงการสอบถามของ Join Predicate Pushdown	57
4.16	แผนประมวลผลแบบแปลงการสอบถาม Join Predicate Pushdown	58
4.17	แคณดิเคตเจเนเรชัน Join Predicate Pushdown	59
4.18	แคณดิเคตเจเนเรชัน Join Predicate Pushdown	59
4.19	กราฟผลการทดลอง Join Predicate Pushdown	60
4.20	การสอบถามแบบ Join Factorization	62
4.21	การแปลงการสอบถามแบบ Join Factorization	62
4.22	แผนประมวลผลแบบไม่แปลงการสอบถามของ Join Factorization	63
4.23	แผนประมวลผลแบบแปลงการสอบถามของ Join Factorization	64
4.24	แคณดิเคตเจเนเรชัน Join Factorization	65
4.25	แคณดิเคตเจเนเรชัน Join Factorization	65
4.26	กราฟผลการทดลอง Join Factorization	66