

การออกแบบคลังข้อมูลสำหรับการไฟฟ้าส่วนภูมิภาค

น.ส. จุฑามาศ นันทธนะวานิช

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาวิทยาศาสตร์คอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2542

ISBN 974-334-770-4

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

I 19C 524 54

20 ก.พ. 2545

DESIGNING THE DATA WAREHOUSE FOR THE PROVINCIAL ELECTRICITY AUTHORITY

Miss Juthamard Nunthanawanit

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Computer Science

Department of Computer Engineering

Faculty of Engineering

Chulalongkorn University

Academic Year 1999

ISBN 974-334-770-4

หัวข้อวิทยานิพนธ์ การออกแบบคลังข้อมูลสำหรับการไฟฟ้าส่วนภูมิภาค
โดย น.ส. จุฑามาศ นันทน์ธนะวานิช
ภาควิชา วิศวกรรมคอมพิวเตอร์
อาจารย์ที่ปรึกษา อาจารย์ จารุมาตร ปิ่นทอง



คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้หัวข้อวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่ง
ของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรบัณฑิต

.....
(ศาสตราจารย์ ดร. สมศักดิ์ ปัญญาแก้ว) คณบดีคณะวิศวกรรมศาสตร์

คณะกรรมการสอบวิทยานิพนธ์

.....
(อาจารย์ วิวัฒน์ วัฒนาวุฒิ) ประธานกรรมการ

.....
(อาจารย์ จารุมาตร ปิ่นทอง) อาจารย์ที่ปรึกษา

.....
(อาจารย์ ดร. บุญเสริม กิจศิริกุล) กรรมการ

.....
(อาจารย์ ดร. ทวีติย์ เสนิงวงศ์ ณ อยุธยา) กรรมการ

จุฑามาศ นันทิธนะวานิช : การออกแบบคลังข้อมูลสำหรับการไฟฟ้าส่วนภูมิภาค
(DESIGNING THE DATA WAREHOUSE FOR PROVINCIAL ELECTRICITY

AUTHORITY) อ. ที่ปรึกษา : อ. จารุมาต ปิ่นทอง , 231 หน้า. ISBN 974-334-770-4.

วัตถุประสงค์ของวิทยานิพนธ์ฉบับนี้คือ เพื่อออกแบบคลังข้อมูลโดยใช้ข้อมูลของการไฟฟ้าส่วนภูมิภาคเป็นกรณีศึกษา และสร้างฐานข้อมูลเมตะดาตาเพื่อควบคุมการทำงานของคลังข้อมูลให้อยู่ในกรอบ โดยออกแบบให้รองรับการทำงานที่ต่อเนื่องกันเป็นชุดเพื่อให้โปรแกรมสามารถทำงานได้เองโดยใช้คนคุมให้น้อยที่สุด

การวิจัยครั้งนี้ใช้ข้อมูลที่เกี่ยวข้องกับรายได้และค่าใช้จ่ายของการไฟฟ้าส่วนภูมิภาคจากระบบงานหลัก 4 ระบบคือ ระบบบัญชีและการเงิน ระบบบริการผู้ใช้ไฟ ระบบผลิตดูแลคลัง และระบบบริหารบุคคลซึ่งจะใช้ข้อมูลจากระบบบุคลากร โดยที่ข้อมูลจากระบบงานเหล่านี้มีการจัดเก็บในรูปของแฟ้มข้อมูล ข้อมูลที่คัดเลือกเป็นข้อมูลนำเข้าคลังข้อมูลนี้ได้จากการสัมภาษณ์ การสำรวจจากรายงานที่เสนอผู้บริหารของการไฟฟ้าส่วนภูมิภาค และโครงสร้างแฟ้มข้อมูลจากระบบงานหลักทั้ง 4 ระบบนี้

การพัฒนาคลังข้อมูล ในที่นี้มี 4 ส่วน คือ การพัฒนาฐานข้อมูลเมตะดาตา การนำข้อมูลเข้าคลัง การพัฒนาฐานข้อมูลสำหรับคลังข้อมูล และการเรียกใช้ข้อมูลจากคลังข้อมูล ในที่นี้ใช้หลักการเชื่อมโยงแบบดาวในการออกแบบการเชื่อมโยงข้อมูลในคลังข้อมูลเพื่อรองรับการวิเคราะห์จำแนกตามมิติ การเรียกใช้ข้อมูลเป็นตัวอย่างโปรแกรมสำหรับทดสอบคลังข้อมูล โดยพัฒนาให้สามารถตอบคำถามเชิงสรุป คำถามเกี่ยวกับข้อมูลย้อนหลัง และคำถามเชิงเปรียบเทียบเพื่อการวิเคราะห์ได้เป็นอย่างดี

ภาควิชา วิศวกรรมคอมพิวเตอร์

สาขาวิชา วิทยาศาสตร์คอมพิวเตอร์

ปีการศึกษา 2542

ลายมือชื่อนิสิต จุฑามาศ นันทิธนะวานิช

ลายมือชื่ออาจารย์ที่ปรึกษา [ลายมือ]

ลายมือชื่ออาจารย์ที่ปรึกษาร่วม _____

3970311321 : MAJOR COMPUTER SCIENCE

KEY WORD : DATA WAREHOUSE / METADATA / LEGACY SYSTEM

JUTHAMARD NUNTHANAWANIT : DESIGNING THE DATA WAREHOUSE FOR

THE PROVINCIAL ELECTRICITY AUTHORITY. THESIS ADVISOR : CHARUMATR

PINTHONG. 231 pp. ISBN 974-334-770-4.

The purpose of this research is to design the database of data warehouse for the Provincial Electricity Authority and to create database of metadata to control all tasks in data warehouse. This design covers semiautomatic working.

This research collected data about income and expense from 4 subjects, namely Accounting and Finance, Customer Service, Inventory, and Personnel. The data in legacy systems was stored in conventional files. The selected data came from interviewing, periodic reports for manager and file description of the legacy systems.

This development was divided into metadata database, data extraction, database for data warehouse and data retrieve. The star schema is used for dimensional analysis. This data warehouse is able to answer the summary query, historical query and comparative query for business analysis at least.

ภาควิชา วิศวกรรมคอมพิวเตอร์
สาขาวิชา วิทยาศาสตร์คอมพิวเตอร์
ปีการศึกษา 2542

ลายมือชื่อผู้คิด จuthamad nunthanawanit
ลายมือชื่ออาจารย์ที่ปรึกษา [Signature]
ลายมือชื่ออาจารย์ที่ปรึกษาร่วม _____

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยความช่วยเหลืออย่างดีของอาจารย์ จารุมาตร ปิ่นทอง อาจารย์ที่ปรึกษาวิทยานิพนธ์ ซึ่งท่านได้ให้คำแนะนำและข้อคิดเห็นต่าง ๆ ในการวิจัยด้วยดีตลอด และขอขอบคุณบุคคลรายนามดังต่อไปนี้มา ณ ที่นี้ ที่ให้คำปรึกษาและให้ข้อมูลสำหรับงานวิจัยครั้งนี้

1. นาง รัตนา สิ้นธุประมา
2. นาย สมพงษ์ สมันเลาะ
3. นายณัฐชัยศ พิระปริชา
4. นางอภิญา ศรวิระศาสน์
5. นางไพฑูรย์ พุทธธานี
6. นายอนนท์ ศุภตระกูล
7. ทีมงานพัฒนาระบบบัญชีและการเงิน
8. ทีมงานพัฒนาระบบบริการผู้ใช้ไฟ
9. ทีมงานพัฒนาระบบพัสดุคงคลัง
10. ทีมงานพัฒนาระบบบริหารบุคคล
11. ทีมงานสารสนเทศสำหรับผู้บริหาร
12. ผู้ให้ข้อมูลทุกท่าน

ท้ายนี้ ขอกราบขอบพระคุณ บิดา-มารดา ซึ่งสนับสนุนในด้านการเรียนและให้กำลังใจเสมอจนสำเร็จการศึกษา

สารบัญ

บทคัดย่อภาษาไทย	ง
บทคัดย่อภาษาอังกฤษ	จ
กิตติกรรมประกาศ	ฉ
สารบัญ	ช
สารบัญตาราง	ญ
สารบัญภาพ	ฎ
บทที่ 1 บทนำ.....	1
ความเป็นมาและความสำคัญของปัญหา	1
วัตถุประสงค์ของวิทยานิพนธ์	2
ขอบเขตของวิทยานิพนธ์.....	2
ประโยชน์ที่คาดว่าจะได้รับ	3
บทที่ 2 แนวคิดและทฤษฎีที่เกี่ยวข้อง	4
ความหมายของคลังข้อมูล	4
องค์ประกอบของคลังข้อมูล.....	5
ความหมายของเมตาดาตา	7
การสร้างแบบจำลองข้อมูลของคลังข้อมูล ..	9
การเชื่อมโยงแบบดาว.....	10
ลักษณะของเทเบิลแฟคท์.....	10
ลักษณะของเทเบิลโดเมนชั้น	10
ขั้นตอนในการกำหนดแฟคท์และโดเมนชั้น.....	11
วิธีการออกแบบเทเบิลแฟคท์.....	11
วิธีการออกแบบเทเบิลโดเมนชั้น	11
การทำดินอร์มอลไลซ์.....	12
คุณภาพของข้อมูล.....	13
แนวคิดเพิ่มเติมที่ใช้ในการออกแบบคลังข้อมูล	14
ขั้นตอนการดำเนินการ	14

สารบัญ (ต่อ)

บทที่ 3 ภาพรวมการใช้ข้อมูล.....	16
ผู้ที่เกี่ยวข้องกับคลังข้อมูล.....	16
การสำรวจความต้องการใช้ข้อมูลของผู้ใช้ระบบ.....	16
ภาพรวมการเชื่อมโยงข้อมูลขององค์กร.....	18
ระบบงานหลักและขอบเขตของข้อมูล.....	25
ระบบงานที่เลือกเป็นกรณีศึกษาและเหตุผลที่เลือก.....	25
ระบบงานที่ไม่เลือกเป็นกรณีศึกษาและเหตุผลที่ไม่เลือก.....	26
วัตถุประสงค์ของระบบงานคอมพิวเตอร์.....	26
บทที่ 4 การออกแบบคลังข้อมูล.....	33
หลักเกณฑ์สำคัญสำหรับคลังข้อมูล กฟภ.....	33
การเก็บข้อมูลในคลังข้อมูล.....	33
การคัดเลือกข้อมูล.....	33
การจัดการข้อมูลที่ไม่เชื่อมโยงกัน.....	34
การเลือกใช้ข้อมูลที่ซ้ำซ้อนกันระหว่างระบบงาน.....	34
การจัดการข้อมูลที่ไม่ถูกต้องตรงกันระหว่างระบบงาน.....	36
การคัดเลือกรายการข้อมูลและเหตุผล.....	36
รอบเวลาของข้อมูล.....	44
ระบบบัญชีและการเงิน.....	44
ระบบบริการผู้ใช้ไฟ.....	45
ระบบพัสดุคลัง.....	46
ระบบบริหารบุคคล.....	47
สรุปเหตุการณ์ที่ทำให้ต้องส่งข้อมูลจากระบบต้นแบบเข้าคลังข้อมูล.....	48
การส่งข้อมูลระหว่างเหตุการณ์ที่ต้องสอดคล้องกัน.....	48
จำนวนตัวอักษรที่ใช้ในคลังข้อมูลโดยประมาณ.....	50
รูปแบบของข้อมูลนำเข้า.....	51
การแบ่งชุดข้อมูล.....	51
การจัดเทเบิล.....	51

สารบัญ (ต่อ)

การออกแบบเทเบิล	52
การออกแบบเทเบิลแพคท์สำหรับคลังข้อมูล	52
การออกแบบเทเบิลโดเมนชั้นสำหรับคลังข้อมูล	53
เทเบิลเวลา	53
เทเบิลหน่วยงาน	54
เทเบิลรายการข้อมูล	54
การจัดความสัมพันธ์ของเทเบิล	55
วัตถุประสงค์ของเทเบิลแพคท์	56
วัตถุประสงค์ของเทเบิลโดเมนชั้น	57
บทที่ 5 เมตะดาตา	62
แบบจำลองข้อมูลของเมตะดาตา	62
วัตถุประสงค์ของเทเบิลหลัก	68
วัตถุประสงค์ของเทเบิลอ้างอิง	70
การใช้ประโยชน์เมตะดาตาสำหรับคลังข้อมูล	71
การใช้ประโยชน์จากแอททริบิวต์ของเมตะดาตา	72
บทที่ 6 การพัฒนาค้างข้อมูล	83
การทำงานเกี่ยวกับคลังข้อมูลตามรอบเวลาของเหตุการณ์	83
การพัฒนาค้างข้อมูลและการไหลของข้อมูล	84
โปรแกรมที่พัฒนา	88
ชุดโปรแกรมจัดการเมตะดาตา	88
ชุดโปรแกรมจัดการคลังข้อมูล	90
ชุดโปรแกรมการเรียกใช้ข้อมูล	91
เครื่องมือที่ใช้ในการพัฒนา	91
บทที่ 7 การรักษาความปลอดภัยคลังข้อมูล	115
การแบ่งกลุ่มข้อมูล	115
การสำรองข้อมูล	115
การเรียกคืนข้อมูล	117

สารบัญ (ต่อ)

บทที่ 8 การทดสอบคลังข้อมูล	118
บทที่ 9 สรุปการวิจัย	126
สรุปหลักการที่ใช้ในการออกแบบ.....	126
ปัญหาที่พบจากการวิจัย	126
สรุปสิ่งที่ได้จากการวิจัยครั้งนี้	126
ข้อเสนอแนะ	127
รายการอ้างอิง.....	128
ภาคผนวก ก พจนานุกรมข้อมูลเกี่ยวกับกระบวนการทำงาน.....	130
ภาคผนวก ข พจนานุกรมข้อมูลเกี่ยวกับโครงสร้างเทเบิลของเมตาดาตาและคลังข้อมูล	159
โครงสร้างเทเบิลของเมตาดาตา.....	160
โครงสร้างเทเบิลของคลังข้อมูล	181
ภาคผนวก ค พจนานุกรมข้อมูลเกี่ยวกับโครงสร้างแฟ้มข้อมูลที่ระบบต้นแบบส่งเข้าคลังข้อมูล	187
ภาคผนวก ง พจนานุกรมข้อมูลเกี่ยวกับโครงสร้างแฟ้มข้อมูลของระบบต้นแบบ	200
ภาคผนวก จ ตัวอย่างจอภาพ.....	222
ประวัติผู้เขียน.....	231

สารบัญตาราง

ตารางที่ 4.1	ตัวอย่างการเลือกใช้ข้อมูลจากระบบบัญชีและการเงินและเหตุผล.....	37
ตารางที่ 4.2	ตัวอย่างการเลือกใช้ข้อมูลจากระบบบริการผู้ใช้ไฟและเหตุผล.....	38
ตารางที่ 4.3	ตัวอย่างการเลือกใช้ข้อมูลจากระบบพัสดุคงคลังและเหตุผล.....	43
ตารางที่ 4.4	ตัวอย่างการเลือกใช้ข้อมูลจากระบบบริหารบุคคลและเหตุผล.....	43
ตารางที่ 4.5	รอบเวลาของตัวอย่างข้อมูลจากระบบบัญชีและการเงิน.....	45
ตารางที่ 4.6	รอบเวลาของตัวอย่างข้อมูลจากระบบบริการผู้ใช้ไฟ.....	46
ตารางที่ 4.7	รอบเวลาของตัวอย่างข้อมูลจากระบบพัสดุคงคลัง.....	47
ตารางที่ 4.8	รอบเวลาของตัวอย่างข้อมูลจากระบบบุคลากร.....	47
ตารางที่ 4.9	ตัวอย่างเหตุการณ์ที่ทำให้ต้องส่งข้อมูลจากระบบต้นแบบเข้าคลังข้อมูล.....	48
ตารางที่ 4.10	ตัวอย่างการส่งข้อมูลระหว่างเหตุการณ์ที่ต้องสอดคล้องกัน.....	48
ตารางที่ 4.11	ตัวอย่างการส่งข้อมูลระหว่างเหตุการณ์ที่ต้องสอดคล้องกัน.....	49
ตารางที่ 4.12	ตัวอย่างการส่งข้อมูลระหว่างเหตุการณ์ที่ต้องสอดคล้องกัน.....	49
ตารางที่ 7.1	กลุ่มข้อมูล ชุดข้อมูล และที่เก็บ.....	115
ตารางที่ 7.2	วิธีการสำรองข้อมูลของคลังข้อมูล.....	116
ตารางที่ 7.3	วิธีการสำรองข้อมูลเมตะดาตา.....	116
ตารางที่ 7.4	วิธีการสำรองข้อมูลของไดเรกทอรีเก็บแฟ้มข้อมูลที่ได้รับจากระบบต้นแบบ.....	116

สารบัญภาพ

รูปที่ 2.1	องค์ประกอบของคลังข้อมูล.....	6
รูปที่ 2.2	มิติข้อมูลของคลังข้อมูล.....	9
รูปที่ 2.3	ตัวอย่างการเชื่อมโยงข้อมูลแบบดาวเพื่อสร้างข้อมูลหลายมิติ.....	9
รูปที่ 2.4	โครงสร้างแบบสตาร์เฟลค.....	13
รูปที่ 3.1	ลักษณะงานและสิ่งที่เกี่ยวข้อง.....	19
รูปที่ 3.2	ภาพรวมระบบงานและการเชื่อมโยงข้อมูล.....	20
รูปที่ 3.3	การเชื่อมโยงข้อมูล (ระบบบริการผู้ใช้ไฟ).....	21
รูปที่ 3.4	การเชื่อมโยงข้อมูล (ระบบงานบัญชีและการเงิน).....	22
รูปที่ 3.5	การเชื่อมโยงข้อมูล (ระบบพัสดุคลัง).....	23
รูปที่ 3.6	การเชื่อมโยงข้อมูล (ระบบบริหารบุคคล).....	24
รูปที่ 4.1	ประเภทย่อยของข้อมูลรายการบัญชี.....	54
รูปที่ 4.2	สัญลักษณ์ที่ใช้ในการเชื่อมโยงแบบดาว.....	55
รูปที่ 4.3	แบบจำลองการเชื่อมโยงข้อมูลแบบดาว สำหรับการวิเคราะห์รายการบัญชีและการเงิน.....	57
รูปที่ 4.4	แบบจำลองการเชื่อมโยงข้อมูลแบบดาว สำหรับการวิเคราะห์การบริการผู้ใช้ไฟ.....	58
รูปที่ 4.5	แบบจำลองการเชื่อมโยงข้อมูลแบบดาว สำหรับการวิเคราะห์การพัสดุคลัง.....	59
รูปที่ 4.6	แบบจำลองการเชื่อมโยงข้อมูลแบบดาว สำหรับการวิเคราะห์การบริหารบุคคล.....	60
รูปที่ 4.7	การใช้เทเบิลไดเมนชันร่วมกันเพื่อวิเคราะห์เจาะลึกระหว่างเทเบิลแฟคท์.....	61
รูปที่ 5.1	เมตาดาตา แสดงสัญลักษณ์ที่ใช้ในรูปที่ 5.2 – 5.6.....	62
รูปที่ 5.2	เมตาดาตา แสดงสิ่งที่มีอยู่ในคลังข้อมูล.....	63
รูปที่ 5.3	เมตาดาตา แสดงข้อมูลเข้าและข้อมูลออก.....	64
รูปที่ 5.4	เมตาดาตา แสดงการกำหนดเวลานำข้อมูลเข้าคลัง.....	65
รูปที่ 5.5	เมตาดาตาสำหรับการวัดคุณภาพของข้อมูล.....	66
รูปที่ 5.6	แสดงเมตาดาตาสำหรับการตรวจความสอดคล้องของข้อมูล.....	67
รูปที่ 6.1	ขั้นตอนการจัดการข้อมูลสำหรับคลังข้อมูล.....	92
รูปที่ 6.2	สัญลักษณ์ที่ใช้สำหรับแผนผังการไหลของข้อมูล ตั้งแต่รูปที่ 6.3 – 6.23.....	93
รูปที่ 6.3	คอนเทกซ์ไดอะแกรมของการจัดทำคลังข้อมูล.....	94
รูปที่ 6.4	แผนภาพการไหลของข้อมูล ระดับที่ 1.....	95
รูปที่ 6.5	แผนภาพการไหลของข้อมูล ระดับที่ 1.....	96

สารบัญญภาพ (ต่อ)

รูปที่ 6.7	แผนภาพการไหลของข้อมูล ระดับที่ 1	98
รูปที่ 6.8	แผนภาพการไหลของข้อมูล ระดับที่ 1	99
รูปที่ 6.9	แผนภาพการไหลของข้อมูล ระดับที่ 1	100
รูปที่ 6.10	แผนภาพการไหลของข้อมูล ระดับที่ 1	101
รูปที่ 6.11	แผนภาพการไหลของข้อมูล ระดับที่ 2	102
รูปที่ 6.12	แผนภาพการไหลของข้อมูล ระดับที่ 2	103
รูปที่ 6.13	แผนภาพการไหลของข้อมูล ระดับที่ 2	104
รูปที่ 6.14	แผนภาพการไหลของข้อมูล ระดับที่ 2	105
รูปที่ 6.15	แผนภาพการไหลของข้อมูล ระดับที่ 2	106
รูปที่ 6.16	แผนภาพการไหลของข้อมูล ระดับที่ 2	107
รูปที่ 6.17	แผนภาพการไหลของข้อมูล ระดับที่ 2	108
รูปที่ 6.18	แผนภาพการไหลของข้อมูล ระดับที่ 2	109
รูปที่ 6.19	แผนภาพการไหลของข้อมูล ระดับที่ 2	110
รูปที่ 6.20	แผนภาพการไหลของข้อมูล ระดับที่ 3	111
รูปที่ 6.21	แผนภาพการไหลของข้อมูล ระดับที่ 3	112
รูปที่ 6.22	แผนภาพการไหลของข้อมูล ระดับที่ 3	113
รูปที่ 6.23	แผนภาพการไหลของข้อมูล ระดับที่ 3	114
รูปที่ 8.1	ระบบงานหลัก	120
รูปที่ 8.2	ประเภทหลักของข้อมูล ประเภทย่อย และรายการข้อมูลระดับสุดท้าย	121
รูปที่ 8.3	ผลลัพธ์แบบสรุปจากการสร้างคำถามสรุปและดริลดาวน์	122
รูปที่ 8.4	ผลลัพธ์ที่เป็นรายละเอียดจากการสร้างคำถามสรุปและดริลดาวน์	123
รูปที่ 8.5	ผลลัพธ์ของคำถามสถิติ	124
รูปที่ 8.6	ผลลัพธ์จากคำถามเปรียบเทียบ	125