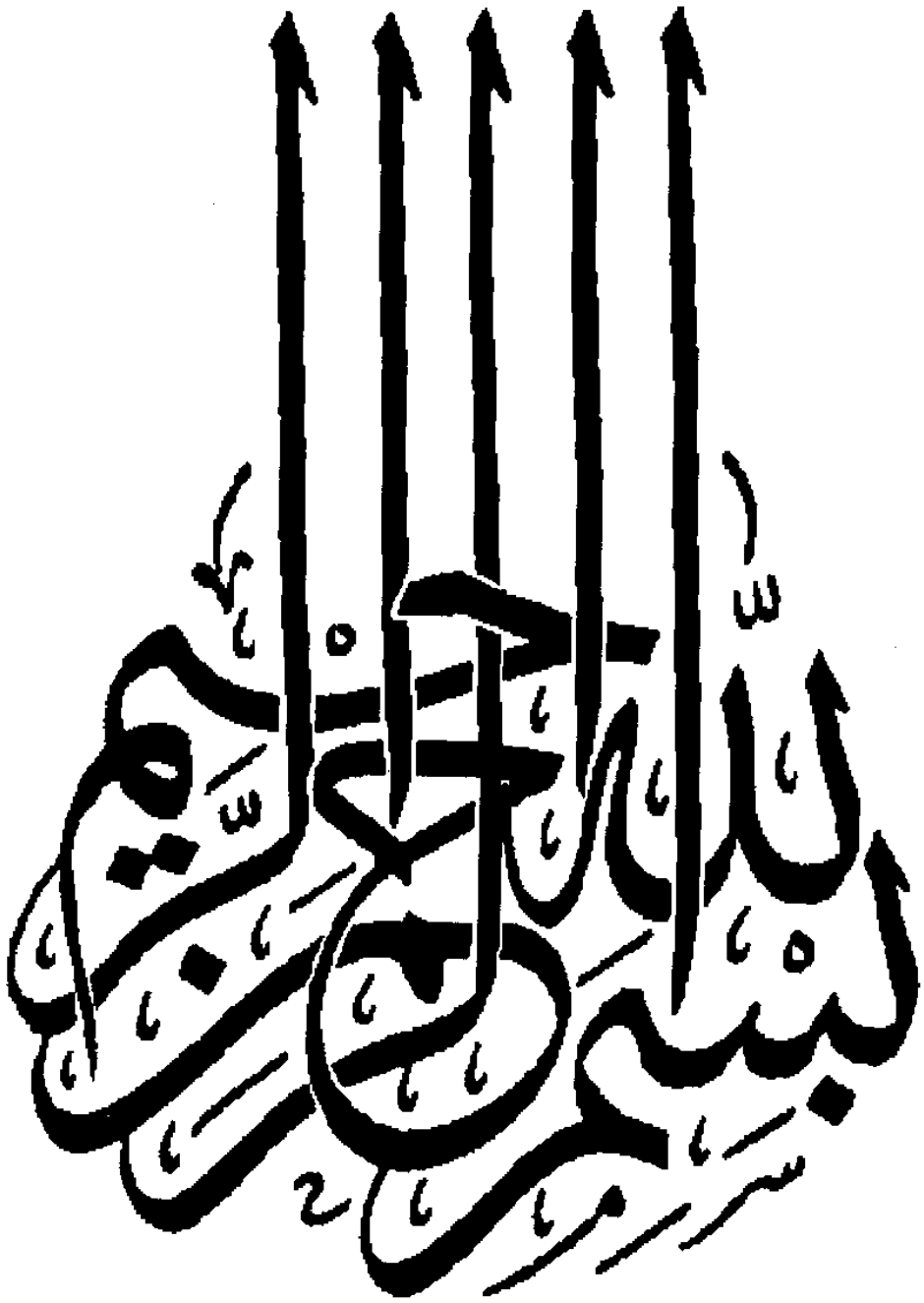


Central Library - Ain Shams University

Ain Shams University

Faculty of Engineering

Computer-Aided Design for Distributed Database Systems

By

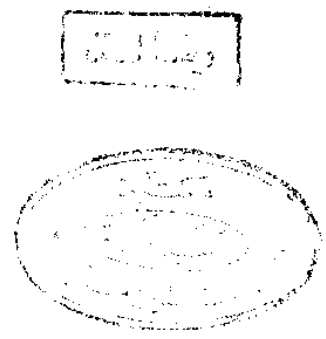
Hoda Korashy Mohamed

621.38195
K
11

A Thesis

Submitted in fulfillment of the requirements of the Degree of PH.D. in Electrical Engineering

(Electronics and Computer)



Supervised By

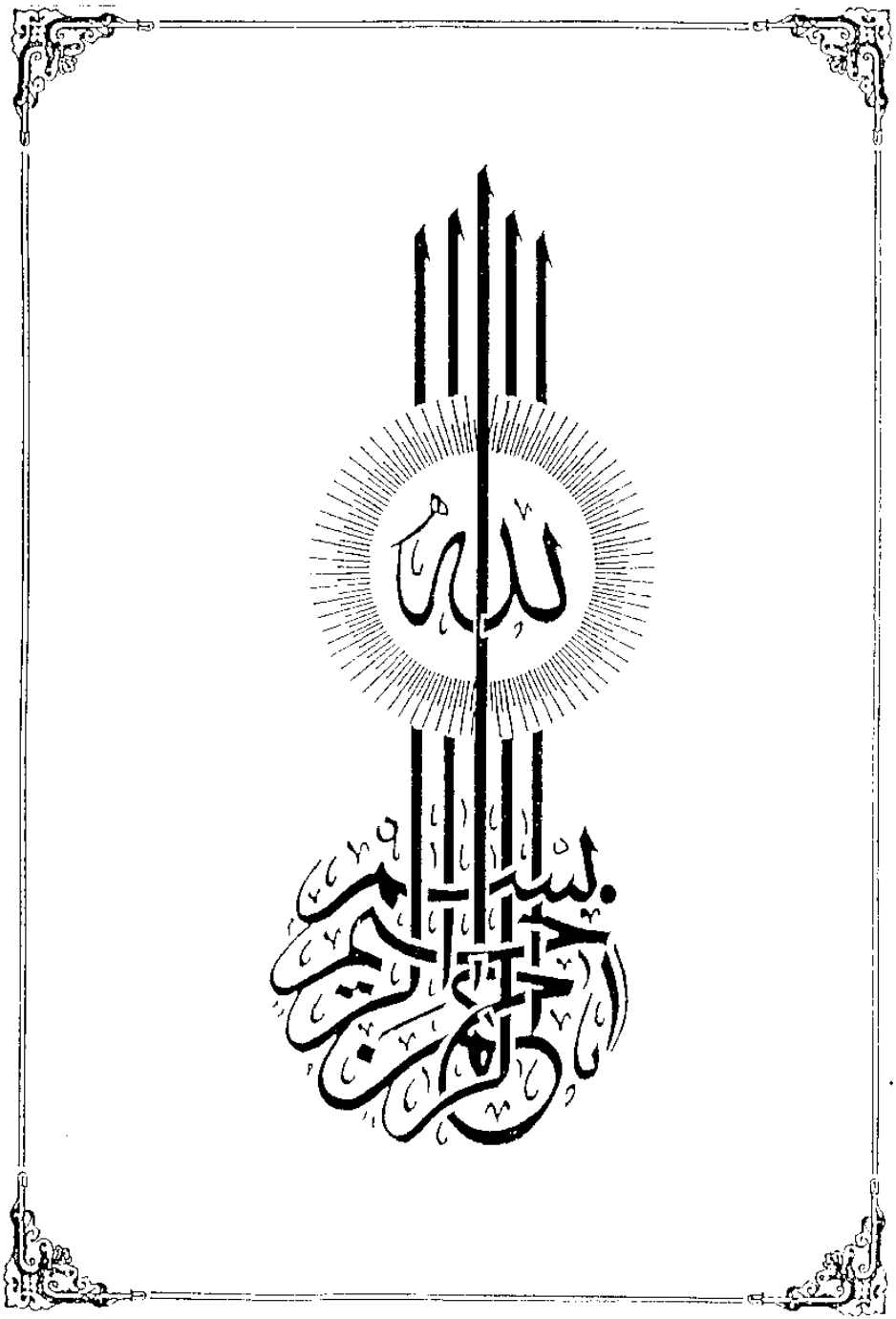
Prof. Dr. M. Adeeb R. Ghonaimy

Prof. Dr. Osman A. Badr

5.35

Cairo
(1992)

Central Library - Ain Shams University



Central Library - Ain Shams University

Examiners Committee

Name , Title and Affiliation	Signature
1- Prof. Dr. R.A. Amer Prof. at King Abdeel Aziz University, Saudi Arabia.	----------
2- Prof. Dr. F.H. Saleh Prof. at Cairo University, Egypt .	
3- Prof. Dr. M.A.R. Ghonaimy Prof. at Ain Shams University, Egypt.	M.A.R. Ghonaimy

Date: 10 / 3 / 1992

Central Library - Ain Shams University

STATEMENT

This dissertain is submitted to **Ain Shams University** for the degree of **PH.D.** in **Electrical** Engineering (**Electronics and computers**).

The work included in this thesis , was carried out by the author in the Department of **Faculty of Engineer**, Ain Shams University , from **2/1986** to **3/1992** .

No part of this thesis has been submitted for a degree or a qualification at any other university or institution .

Date : 3 /1992

Name : Hoda Korashy Mohamed

Signature : Hoda Korashy

Central Library - Ain Shams University

ACKNOWLEDGMENT

I would like to express my limitless gratitude to **Prof. Dr. M.A.Ghonaimy** for his invaliable advices, supervision and guidance throught the research.

I am also grateful to **Prof. Dr. O.A.Badr** for his patient guidance and continuing encouragement.

I would like also to thank all my collegeous, espicially Eng. Mona Fahmy, for their help and encouragement.

Finally, I would like to acknowledge the encouragement and support shown by my husband and my mother.

Central Library - Ain Shams University

ABSTRACT

A comparative study for the design of distributed database has been done. The design of distributed database, may be divided into four phases, each has its own techniques and tools.

A proposed system is applied to design the distributed database system. The proposed system (**DBDSG**) is a semiautomated methodology which covers all the design steps. It is divided into two parts.

DBDSG-1 system is a menu driven system to facilitate the interface with designers. Data are represented using *Extended-Entity Relationship model [EERM]*. View integration is applied if more than one view exists. The system transforms the EERM into a relational data model which is the data model required for the next phase of the design. Analysis of the important transactions is applied, to develop all possible partitioning of objects and the access path for each transaction .

In **DBDSG-2** system, a model is proposed for the nonreplicated allocation of data over the sites of the computer network. A modified model is presented to modify some of the constraints in the original model. Zero/one implicit enumeration algorithm is used to solve the allocation problem. For replicated allocation of data, a heuristic postoptimization of the optimal solution without replication is applied. An analysis of the results of the allocation model proposed with the effect of different parameters is also discussed . The parameters which affect the allocation model are the cost parameter such as transmission cost and the access cost of the computers at each site. The load of transactions at each site has a pronounced effect in the allocation problem.

Central Library - Ain Shams University

TABLE OF CONTENTS

Central Library - Ain Shams University

3 Data Distribution Design

Central Library - Ain Shams University

4 The Proposed Model (DBDSG)

5 Data Allocation Model (DBDSG-2)

Central Library - Ain Shams University

Central Library - Ain Shams University

Central Library - Ain Shams University

CHAPTER 1

INTRODUCTION

Central Library - Ain Shams University

CHAPTER 1

INTRODUCTION

1.1 What is meant by Distributed Database?

A typical definition of a distributed database is :

" A distributed database is a collection of data which are distributed over different computers forming a network. Each site of the network has autonomous processing capability and can perform local applications. Each site also participates in the execution of at least one global application, which requires accessing data at several sites using a communication subsystem" [CERI84].

The definition emphasizes two important aspects :

1- Distribution

Distribution means the fact that the data are not resident at the same site. That will distinguish a distributed database from a single centralized database.

2- Logical correlation

The fact that the data have some properties which tie them together is one of the characteristics of distributed databases , which are different from a set of local databases residing at different sites of a computer network.

Central Library - Ain Shams University

1.2 Features of Distributed versus Centralized Database

1- Centralized Control

In a **centralized database**, database administrator **DBA** is responsible for the safety of data. In a **distributed database**, there is *a global database administrator* who has the central responsibility of the whole database, and *local database administrators* who have the responsibility of their respective local databases. But **local DBA** may have a high degree of autonomy so the **global DBA** is completely missing and the intersite coordination is performed by the local administrators.

2- Data independence

Data independence means that *the actual organization of data is transparent to the users*. In **centralized databases**, this was achieved by conceptual schema. In **distributed databases**, we have **distribution transparency** which means that *users see the databases as if they were not distributed*.

3- Reduction of redundancy

In **centralized databases**, redundancy is reduced as far as possible by using *data sharing* that means allowing several applications to access the same files and records. In **distributed databases**, there are several reasons for considering data redundancy as a desired feature :

a- To increase site locality of applications.

b- The availability of the system can be increased.

4- Complex physical structures and efficient access

In **centralized databases**, complex physical structures are a major aspect for efficient access to the data. In **distributed**

Central Library - Ain Shams University