

AN ANALYTICAL MODEL FOR EVALUATING
DATABASE UPDATE SCHEMES

by

Kathryn C. Kinsley

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy
in the Department of Computer Science at
the University of Central Florida
Orlando, Florida

April 1983

Major Professor: Dr. Charles E. Hughes

Abstract

A methodology is presented for evaluating the performance of database update schemes. The methodology uses the M/Hr/l queueing model as a basis for this analysis and makes use of the history of how data is used in the database. Parameters have been introduced which can be set based on the characteristics of a specific system. These include update to retrieval ratio, average file size, overhead, block size and the expected number of items in the database.

The analysis is specifically directed toward the support of derived data within the relational model. Three support methods are analyzed. These are first examined in a central database system. The analysis is then extended in order to measure performance in a distributed system. Because concurrency is a major problem in a distributive system, the support of derived data is analyzed with respect to three distributive concurrency control techniques -- master/slave, distributed and synchronized.

In addition to its use as a performance predictor, the development of the methodology serves to demonstrate how queueing theory may be used to investigate other related database problems. This is an important benefit due to the lack of fundamental results in the area of using queueing theory to analyze database performance.

ACKNOWLEDGEMENTS

The author wishes to thank her adviser, Dr. Charles E. Hughes for his direction of this work. It has been a pleasure to work with him. His patience and insight is appreciated.

In addition, the author wishes to thank her research committee, namely, Drs. Mostafa Bassiouni, Ron Dutton, Ratan Guha and Brian Petrasko, each of whom contributed to the quality and direction of this dissertation. A special thanks is due to Dr. Terry J. Frederick, head of the Computer Science Department, who has provided support and encouragement throughout the author's graduate academic career.

And of course, the author expresses her appreciation to her husband for his understanding and encouragement throughout this endeavor.

TABLE OF CONTENTS

LIST OF FIGURES	vi
LIST OF SYMBOLS AND ABBREVIATIONS	ix
Chapter	
I. INTRODUCTION	1
A. Objective	1
B. Database Terminology	3
C. Queueing Theory Terminology	5
D. Organization of Presentation	9
II. SURVEY OF THE LITERATURE	11
A. Dynamic Derived Relations	11
B. Formal Analysis of Database Performance in a Centralized Environment	20
C. Distributive Databases	27
D. Formal Analysis of Database Performance in Distributive Database Systems	41
III. EVALUATING DDR SUPPORT METHODS IN A CENTRALIZED SYSTEM	47
A. Characteristics of the Three Update Methods	47
B. Using the History of Database Retrievals and Updates to Determine the Average Service Time	49
C. The M/G/1 Model	53
D. Using the M/Hr/1 Model to Represent the Immediate Actual Method	57
E. Using the M/Hr/1 Model to Represent the Deferred Actual Method	59
F. Using the M/Hr/1 Model to Represent the Potential Method	63
G. Incorporating Overhead into the Actual Deferred Method	65
H. Comparison of Performance of the Three ddr Support Methods	67

