# DB2:
## Concepts, Design, and Programming

# DB2:
## Concepts, Design, and Programming


## JAMES MARTIN

with
Kathleen Kavanagh Chapman
Joe Leben



PRENTICE HALL Englewood Cliffs, New Jersey 07632

**Library of Congress Cataloging-in-Publication Data**

Martin, James
DB2, concepts, design, and programming.

"A James Martin book."
Includes index.
1. Data base management. 2. IBM Database 2
(Computer system) I. Chapman, Kathleen Kavanagh.
II. Leben, Joe. III. Title.
QA76.9.D3M363 1989 005.75'65 88-30727
ISBN 0-13-198581-7

Editorial/production supervision: *Kathryn Gollin Marshak and Karen Skrable Fortgang*
Jacket design: *Bruce Kenselaar*
Manufacturing buyer: *Mary Ann Gloriande*



Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

ISBN 0-13-198581-7

PRENTICE-HALL INTERNATIONAL (UK) LIMITED, *London*
PRENTICE-HALL OF AUSTRALIA PTY. LIMITED, *Sydney*
PRENTICE-HALL CANADA INC., *Toronto*
PRENTICE-HALL HISPANOAMERICANA, S.A., *Mexico*
PRENTICE-HALL OF INDIA PRIVATE LIMITED, *New Delhi*
PRENTICE-HALL OF JAPAN, INC., *Tokyo*
SIMON & SCHUSTER ASIA PTE. LTD., *Singapore*
EDITORA PRENTICE-HALL DO BRASIL, LTDA., *Rio de Janeiro*

# PREFACE

Databases are playing an increasingly important role in supporting an enterprise's information processing requirements. The development of corporate databases will continue to be an extremely important data processing activity for many years to come. Data will be regarded as a vital corporate resource that must be managed so as to maximize its value. Both the quantity of data stored and the complexity of its organization are increasing by leaps and bounds.

The pointer-structured databases of an earlier era are giving way to *relational* database technology because relational databases are more flexible, are based on formal mathematics, and support a wide diversity of user languages. SQL is now an international standard, and many different user tools produce SQL code that can execute with DB2 or other relational databases.

DB2 is a *relational* database management system (DBMS) and is IBM's primary strategic product in the database area. Much other IBM software is being designed to link to DB2, and this represents a major part of IBM's strategic thrust in software. DB2, along with associated products such as Query Management Facility (QMF) and Data Extract (DXT), provides a wide range of capabilities, suitable for both ad hoc access by end users and the development of production applications by professional programmers.

**ACKNOWLEDGMENTS** The authors would like to thank Peter Tierney, Vice President of Marketing for Oracle Corporation, who supplied us with a copy of the very fine Oracle relational DBMS. We were able to use Oracle, running on an IBM Personal Computer AT, to initially develop the SQL examples and to validate the content of tables in Chapter 2 and in Appendix I. Two people at IBM contributed to the writing of this book: John Gridley made it possible for us to develop the QBE examples on an IBM main-

frame in Chicago, and Bill Kellow provided the technical information needed to write the parts of the book that deal with the referential integrity features of DB2.

*James Martin*
*Kathleen K. Chapman*
*Joe Leben*

# CONTENTS

# PART I INTRODUCTION

Part I of this book examines key concepts underlying DB2, IBM's large-system relational database management system. Chapter 1 examines the characteristics generally associated with a *database management system* (DBMS) and the functions and facilities provided by the DB2 program product and related products such as *Query Management Facility* (QMF) and *Data Extract* (DXT). Chapter 2 then introduces the *relational model*, examines key concepts associated with it, and shows how DB2 implements relational DBMS facilities. Chapter 3 looks at the three views of data that are used with database management systems and examines how these views are implemented by DB2.

# 1 THE DB2 ENVIRONMENT

## THE DB2 PRODUCT FAMILY

DATABASE 2, or DB2 as it is more commonly called, is one of several database products offered by IBM. DB2 is a relational database management system designed for the large-system (MVS) environment.* IBM offers another relational database management system, called SQL/DS, for midrange systems running DOS/VS or VM. The first database management system offered by IBM was Information Management System (IMS), which also runs in the MVS environment. A product called DL/I is the counterpart to IMS for midrange systems running the DOS/VS operating system.

In addition to DB2 itself, there are two other important program products that are widely used in the DB2 environment to extend the capabilities of DB2. These two products are closely associated with DB2, and we will be making numerous references to them throughout the book:

- **Query Management Facility (QMF).** QMF is a product that serves as an easy-to-use front end to DB2. It allows end users to retrieve data interactively, to format data into reports, and to display the data graphically.
- **Data Extract (DXT).** DXT is used to extract data from IMS, SQL/DS, or DB2 databases, from VSAM files, or from physical sequential files in order to make the data available to DB2.

*Many terms used in data processing literature begin as two or more words, such as *real-time* or *offline*. It is quite common in the computer industry for a widely accepted, often-used term to evolve into a new single word. The term *database* is now widely used and generally accepted as one word. In this book, we use the term *database* except when we are referring to an IBM product or manual name that uses the older two-word form. It is interesting to note that at the time of this writing, IBM uses the one-word form in naming the DB2 product—DATABASE 2—but continues to use the two-word form in the text of the DB2 reference manuals.

**BOX 1.1 IBM products that can be used with DB2.**

- **DB2 Performance Monitor.** This product is used to collect and format information about DB2 system and application activity.
- **Data Base Relational Application Directory.** This product provides directory support for application development in the DB2 environment.
- **Host Data Base View.** This product can be used to extract DB2 data and make it available to an IBM Personal Computer.
- **Data Base Migration Aid Utility.** This product helps migrate data and catalog information from one DB2 system to another.
- **Data Base Edit Facility.** This product provides a simple-to-use interface for DB2 data manipulation.
- **Cross System Product (CSP).** This product is a fourth-generation language (4GL), primarily intended for professional application developers, that can be used to generate applications that access DB2 databases.
- **Application Development Facility (ADF).** This is another 4GL, also primarily intended to be used by information systems professionals, to generate IMS and DB2 applications.
- **Application System (AS).** This product is a 4GL, primarily intended for end users, that can be used to access DB2 databases.

We will not discuss the many other IBM program products that can be used with DB2. Some of these are briefly described in Box 1.1. Also, many products from other vendors can also be used to access DB2 databases.

## THE DB2 OPERATING ENVIRONMENT

Describing the operating system and telecommunications environment in which the DB2 software runs requires generating a little IBM "alphabet soup." If you are not well acquainted with the software that is used in the IBM large-system environment, this section can safely be skipped over and the alphabet soup ignored. Having said that, DB2 runs under any of the currently supported versions of the MVS operating systems and can be used in conjunction with either IMS, CICS, or TSO. When run with TSO, DB2 data can be accessed from a TSO terminal running in foreground or by a batch job running in background. When using DB2 with TSO, DB2 provides the database services and TSO provides data communication services. When DB2 is run in the IMS environment, an application program can access data in both DB2 and IMS databases; IMS-DC provides data communication services in the IMS en-

vironment. When DB2 is run in the CICS environment, CICS provides the data communication services. CICS also provides an interface to IMS databases so that a single CICS application program can access both DB2 and IMS data. When an application program uses the data communication services of IMS or CICS, DB2 provides the facilities needed to be able to access both DB2 and IMS databases with full data integrity and recoverability; the user need not know the physical locations of the data in either environment.

IBM has stated that DB2 is an important part of its *Systems Application Architecture* (SAA). SAA is IBM's long-range plan for integrating its diverse computing system lines and providing common user and programming interfaces for all of its large, medium, and small systems.

DB2 provides services that are used to access data either interactively using a terminal or via an application program. DB2 application programs can be written in a variety of programming languages, including COBOL, PL/I, FORTRAN, BASIC, APL2, Assembler, C, and many fourth-generation languages.

## THE SQL DATA LANGUAGE

The primary language used by both application programmers and end users in accessing a DB2 database is the *Structured Query Language* (SQL). SQL (often pronounced "sequel") is the standard language interface to DB2 and supports facilities that allow it to be used as both a *data definition language* (DDL) and a *data manipulation language* (DML). Many products other than DB2, such as IBM's Application System (AS), support interfaces to SQL and thus to DB2. SQL statements can be used to define, retrieve, and update data in DB2 databases. In addition to the basic data definition and data manipulation statements, there are also SQL statements that provide for processing related to security, integrity, and recovery and for the physical administration of DB2 databases. SQL statements can be executed interactively from a terminal, or they can be embedded in an application program.

Having a single data language that encompasses all functions needed to both support and use databases helps make DB2 easy to use while still providing a powerful set of facilities. All SQL statements use the same syntax. This allows a person to perform a full range of functions, from simple retrieval of data at a terminal through application system development to complete administration of a DB2 system, without having to learn several languages.

## THE DB2 CATALOG

DB2 maintains information about the data that it manages in a set of tables known as the *DB2 catalog*. Since the DB2 catalog consists of DB2 databases, authorized users can access and manipulate data in the catalog using SQL statements in the same manner as they manipulate data in application databases. Data in the catalog can be used to determine what data items exist in the DB2 data-