The Programmer as Navigator

Charles W.Bachman

Summarised by Nikhil Reddy. P, CS16B018

January 18, 2018

1 Introduction

A new basis for understanding is available in the area of information systems. It is achieved by a shift from a computer centered to the database centered point of view. This new understanding will lead to new solutions to our database problems and speed our conquest of the n-dimensional data structures which best model the complexities of the real world.

2 Into the Database

- The earliest databases, initially implemented on punched cards with sequential file technology.

- Sequential file technology,Start with the value of the primary data key, of the record of interest, and pass each record in the file through core memory until the desired record, or one with a higher key, is found.

- The revolution in thinking is changing the programmer from a stationary viewer of objects passing before him in core into a mobile navigator who is able to probe and traverse a database at will.

- Direct access storage devices also opened up new ways of record retrieval by primary data key. The first was called randomizing, calculated addressing, or hashing.

- The programmer who has advanced from sequential file processing to either index sequential or randomized access processing has greatly reduced his access time.

- The programmer's training should be a full-fledged navigator in an n-dimensional dataspace.

3 Database Management systems

Database Management involves all aspects of storing, retrieving, modifying, and deleting data in the files on personnel and production, airline reservations,-data which is used repeatedly and updated as new information becomes available. These files are mapped through some storage structure onto magnetic tapes or disk packs and the drives that support them.

- Database management has two main functions.

1. First is the inquiry or retrieval activity that reaccesses previously stored data in order to determine the recorded status of some real world entity or relationship. This data has previously been stored by some other job, seconds, minutes, hours, or even days earlier, and has been held in trust by the database management system.It has a continuing responsibility to maintain data between the time when it was stored and the time it is subsequently required for retrieval.

2. The second activity is to update, which includes the original storage of data, its repeated modification as things change, and ultimately,its deletion from the system when the data is no longer needed.

- In addition to a record's primary key, it is frequently desirable to be able to retrieve records on the basis of the value of some other fields. For example, it may be desirable, in planning ten-year awards, to select all the employee records with the "year-of-hire" field value equal to 1964. Such access is retrieval by secondary data key.

- This equality of primary and secondary data key fields reflects real world relationships and provides a way to reestablish these relationships for computer processing purposes. The use of the same data value as a primary key for one record and as a secondary key for a set of records is the basic concept upon which data structure sets are declared and maintained.

- The Integrated Data Store (I-D-S) systems and all other systems based on its concepts consider their basic con- tribution to the programmer to be the capability to associate records into data structure sets and the cap- ability to use these sets as retrieval paths.

- Performance is enhanced by the so-called "clustering" ability of databases where the owner and some or most of the members records of a set are physically stored and accessed together on the same block or page. These systems have been running in virtual memory since 1962.

- The significant functional and performance advantage is to be able to specify the order of retrieval of the records within a set based upon a declared sort field or the time of insertion.

- As navigator he must brave dimly percieved shoals and reefs in his sea,which are created because he has to navigate in a shared database environment.

4   Shared Access

- Shared access is a new and complex variation of multiprogramming or time sharing, which were invented to permit shared, but independent, use of then computer resources.

- Shared access is a specialized version of multiprogram- ming where the critical, shared resources are the records of the database

- The pressures to use shared access are tremendous.Two problems would maintain the pressure for successful shared access. The first is the trend toward the integration of many single purpose files into a few integrated databases; the second is the trend toward interactive processing where the processor can only advance a job as fast as the manually created input messages allow.

- Of the two main functions of database management, inquiry and update, only update creates a potential problem in shared access.Once a single job begins to up- date the database, a potential for trouble exists.

- The two basic causes of trouble in shared access are interference and contamination.

  1. Interference is defined as the negative effect of the updating activity of one job upon the results of another.One job running an accounting trial balance while another was posting transactions illustrates the interference priblem

  2. Contamination is defined as the negative effect upon a job which results from a combination of two events: when another job has aborted and when its output has already been read by the first job.

- A critical question in designing solutions to the shared access problem is the extent of visibility that the application programmer should have.

5   Interesting Projects

1. Z511 Project

- Description

  The project involves designing a complete database management system to address a practical database need and implementing a relational database based on that design. Your database system should be designed to perform general information management tasks such as systematic collection, update, and retrieval of information for a small organization (e.g. music/video/book store). Students should work in a group of 3 or less.

2. Khipu Database Project

- Description

  The Khipu Database Project began in the fall of 2002, with the goal of collecting all known information about khipu into one centralized repository. Having the data in digital form allows researchers to ask questions about khipu which up until now would have been very difficult, if not impossible, to answer.

6 Future Scope

1. A Vision for a Converged Database

- An ideal database architecture would support multiple data models, languages, processing paradigms and storage formats within the one system. Application requirements that dictate a specific database feature should be resolved as configuration options or pluggable features within a single database management system, not as choices between disparate database architectures.

2. Disruptive Database Technologies

- Extrapolating existing technologies is a useful pastime, and is often the only predictive technique available. However, history teaches us that technologies don't always continue upon an existing trajectory. Disruptive technologies emerge which create discontinuities that cannot be extrapolated and cannot always be fully anticipated.

- There are a few computing technology trends which extend beyond database architecture and which may impinge heavily on the databases of the future.They are :

  (a) Universal Memory : A technology should arise that simultaneously provides acceptable economics for mass storage and latency then we might see an almost immediate shift in database architectures. Such a universal memory would provide access speeds equivalent to RAM together with the durability, persistence and storage economics of disk.

  (b) Quantum Computing : Quantum computers could break existing private/public key encryption schemes seems increasingly likely, while quantum key transmission already provides a tamper-proof mechanism for transmitting certificates over distances within a few hundreds of kilometers.

7 Reference

Page acm turing award to Charles W.Bachman,1973