To be presented at the International
Conference on "Very Large Databases",
Framingham, MA, September 22 - 24, 1975

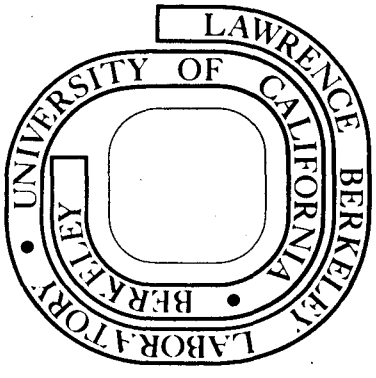
LBL-4233
c.1

HARDWARE AND SYSTEM ARCHITECTURE FOR A VERY LARGE DATABASE

Robert Healey and Bradford Heckman

August 1975

Prepared for the U. S. Energy Research and
Development Administration under Contract W-7405-ENG-48



LBL-4233
c.1

DISCLAIMER

HARDWARE AND SYSTEM ARCHITECTURE FOR A VERY LARGE DATABASE

BY

ROBERT HEALEY AND BRADFORD HECKMAN

LAWRENCE BERKELEY LABORATORY (LBL)

UNIVERSITY OF CALIFORNIA

BERKELEY, CALIFORNIA 94720

This paper describes practical experience in structuring a large amount of data (about 1000 reels of 7-track 556 binary code decimal magnetic tapes of 1970 census data). The hardware and software systems were given. The goal was to structure the database for online devices and to make the retrieval as efficient as possible. This paper explains only the database architecture finally chosen.

The original data records were altered via some aggregation and elimination, plus the addition of region names to identifier records. The latter was accomplished by utilizing the master enumeration district list (medlist) tape provided by the Bureau of the Census. It contains a hierarchical code list which relates each state, county and other area name to codes in the original identifier records. As this alteration to the data record structure was made, the output was written to 7-track or 9-track binary tapes, a form more readily usable on the CDC-6600 and CDC-7600 computer systems.

Two online mass storage systems were used. The first is called the program storage system (PSS) at LBL and is a collection of computer programs which allows data storage to and retrieval from IBM 2321 data cell random rewritable access devices. They use the data medium of a strip of magnetic tape roughly 1 foot long and 2 inches wide. Each data cell contains 2000 such strips, stored in a circular array of 10 cells which contain 200 strips each. Strips are accessed by first rotating the array to the proper position, then mechanically picking up the selected strip and wrapping it aroung a rotating drum. The strip is then read or written as if it were a small drum. The system also uses a dedicated CDC 854 disk pack for the data cell index access. A data cell's storage capacity is over 3 billion bits or about 40 of the above mentioned magnetic tapes. Access time is about 1/2 second and the transfer rate is over 400 kilobits/second.

The second online mass storage system used is cleverly called the mass storage system (MSS) at LBL and is a collection of computer programs which allows data storage to and retrieval from an IBM 1360 photodigital system random access device. It uses the data medium of 35 by 70 millimeter film "chips" (thus the local term for the device is the "chipstore"). About 5 million bits are recorded on a chip in 18 seconds by writing on it with a cathode ray beam. The encoding

algorithm is a visual analog of the phase-encoded recording technique used for magnetic tape. The data is saved on a disk while up to 8 chips are simultaneously developed online in 2.5 minutes. (The limitation of 8 chips is set because the recording is permanent.) After the chips are developed they are read and compared with the saved data. If the comparison fails, the chip is rejected and a new one written. The chipstore can hold 2250 boxes of 32 chips each or over 1/3 of a trillion bits (about 4000 of the above mentioned magnetic tapes). Reading the chipstore is much faster than writing. Access time is 3 seconds and the transfer rate is 2 megabits/second. Like PSS, MSS uses a CDC 854 disk pack for the chipstore index access. Unlike PSS, MSS allows data (in the form of boxes) to be taken offline and placed back online later.

As the converted census tapes were copied to the chipstore, the addresses of the records were collected by state and census "file" to form an index with a list of key codes. (For example, if a record is from the "4th count" census file for San Mateo county in California, the key codes would be 46 for a county designation and 81 for San Mateo county itself.) Because of its relatively large size, this index of records by state and census file was itself stored by MSS. Next, the chipstore addresses of the indices were themselves collected to form a state/file index, and this index, because of its relatively small size, was stored by PSS.

The multilevel indexing scheme described above allows random retrieval of any census record stored on the chipstore. For example, if the 4th count data record for San Mateo county in California is requested, the address of the state index for the correct census file is first retrieved by PSS. Next, the chipstore address of the actual datarecord is retrieved from the chipstore by specification of the appropriate keys necessary for San Mateo county. Then the actual record is randomly retrieved form the chipstore. This procedure requires a small fraction of the time needed to load and sequentially scan a magnetic tape for a wanted record. But the full power of the random retrieval capability becomes apparent when aggregation is done of records originally from 2 or more different magnetic tapes. This is because the physical location of data on the chipstore has negligible effect on retrieval time.

General purpose programs have been developed to randomly retrieve these data via input directives. Among other options these programs allow record aggregation and filtering by population size.

One of the uses of this system is national planning by the U. S. Department of Labor. A text summary and/or a subset of possible fixed tables are generated from retrieved data records and/or aggregation(s) of data records. In 1973 more than 50 million print lines were produced in less than 3 months real time. High quality maps demonstrating census characteristics are also produced. To date maps have been completed up to the federal region level, and a national production of SMSA atlases is underway.

The computer cost for generating the text summary and all possible fixed tables (about 150 pages of computer output) is typically about $10.00 for 1 record. Less than $1.00 of this cost is for the actual data retrieval, the rest being for the report generation from that retrieved data. For example, if 100 records are aggregated, the cost

of actual data retrieval would still be less than $5.00.

Although the initial cost of processing the original magnetic tapes and developing the random indexing system was high, it has more than paid for itself through increased retrieval efficiency. Not only does the architecture described herein allow random data retrieval from online devices, but experience at our installation has demonstrated that retrieval is may times more reliable than with systems utilizing magnetic tape.

REFERENCES

1. Austin, Donald M., Kranz, Suzanne G., and Quong, Carl. An Overview of the LBL Socio-Economic-Environmental-Demographic Information System, report #3699, Lawrence Berkeley Laboratory, March, 1975.
2. Healey, Robert. 1970 Census Data. Lawrence Berkeley Laboratory online documentation, Berkeley, California, August, 1974.
3. Lawrence Berkeley Laboratory conputer center consultants. Lawrence Berkeley Laboratory computer center online documentation, Lawrence Berkeley Laboratory, Berkeley, California, updated frequently.
4. U. S. Bureau of the Census. 1970 Census Users' Guide. U. S. Government Printing Office, Washington, D. C., 1970.

TECHNICAL INFORMATION DIVISION
LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720