The Birth of IMAGE

Fred White

Adager Corporation
Sun Valley, Idaho 83353-3000 • *USA*

www.adager.com

IMAGE was the first database management system implemented on a minicomputer. Subsequent to its release in November 1974, it rapidly achieved acceptance and is acknowledged as one of the foremost database management systems in the world.

There are many reasons for this leadership: reliability, capability, performance, ease-of-use, and ease-of-installation, to name a few.

The importance of IMAGE to the HP3000 is reflected by the number and variety of applications developed by users, OEMs and HP which are IMAGE-based.

It would be nice to think that IMAGE's conception was the result of brilliant foresight on the part of HP marketing or HP management. However, such was not the case. Nor was it brilliant foresight on the part of development personnel.

Rather, it would appear to be the by-product of chance circumstances involving a handful of people whose combined personalities and experiences provided the necessary environment for its conception to occur.

It is the purpose of this paper to discuss the people and circumstances which let to the "birth of IMAGE" and to recount some of the challenges which we overcame.

Prologue

It was the summer of 1971. HP was still and instrument company with few employees who knew what a collating sequence was.

The systems lab was furiously preparing for the November 1972 "happening" (i.e. the introduction of the HP3000 with its

Conception

60-plus programmers sharing three prototype 3000s and a few teletype terminals.

A special section, headed by Lee Johnson and independent of the systems lab, was working on other lower-priority, second-release projects: SIS (Student Information System) SAS (Student Accounting System), QUERY (a file inquiry and reporting system), ISAM (an indexed sequential file access method), the sort/merge facility and the editor.

Dick MacIntire and Lee Bollinger made up the QUERY team; Jon Bale and I made up the ISAM team. Both groups were busy drafting their project specifications.

Dick and Lee became interested in our ISAM project as they saw in it a possible solution to the perceived performance problems of their inquiry facility. Dick proposed that we combine both projects into a single QUERY management project and that Jon and I supply file access methods tailored to meet QUERY requirement s. This proposal was accepted in late September. In retrospect, this appears to have been the moment of IMAGE's conception.

Early pregnancy   It was the Fall of 1971. The project team held countless meetings and engaged in endless arguments over QUERY syntax and semantics, file organizations and data element definitions. All of this was a by-product of our efforts to draft project specifications.

The first draft of our specifications was reviewed in November.

We spent another month revising our specs to include additional features and to clean up our definitions. During this time it became increasingly apparent that much of our difficulties in drafting project specifications were due to the fact that Jon and I were "files" oriented while Dick and Lee were "inquiry and reporting" oriented.

Three milestones occurred in December of 1971: the second draft of our specifications was reviewed, the Project was split along the original lines (Dick and Lee with QUERY and Jon and I with IMAGE), and Bob Mayer joined the IMAGE team.

The IMAGE project was still responsible for providing an access method for QUERY. The major difference was that this access method would not be "buried" within QUERY but would be accessible to all HP3000 languages.

This approach allowed each team to develop its own set of project specifications as well as exercise better control of its project. Our meetings were also smaller and shorter.

It was at this time that it dawned on us that we were embarked on a database management project.

This period began in January of 1972. Dick MacIntire was the project manager of both projects.

Dick and Lee Bollinger jointly developed the QUERY project specifications which were completed in the Spring of 1972. At that time, however, a decision was made to develop IMAGE on the 2100 to "show capability." The suggestion to do this is attributed to Jim Treybig, now president of Tandem Corporation. Lee Bollinger was spun-off as project manager and was joined by Bob Brown in June of 1972. Carol Fuquay and Phil Taylor were added in September of 1972.

Jon, Bob and I continued developing IMAGE specifications which were reviewed twice more before achieving final approval in late March of 1972.

Our purpose was to provide a system which was, first and foremost, reliable and which was easy to use, had a critical mass of capabilities, had reasonable performance characteristics and an extendible access method usable from all HP3000 languages. We required that it honor all of the file management security provisions and that it extend them by providing privacy and security at the database, dataset and data item levels.

All of this had to be provided on a system with as little as 64K bytes of memory and with the sure knowledge that if we embellished it too much, it wouldn't fly.

Agreements from earlier reviews enabled Jon to begin the coding of DBSCHEMA in late January. I began coding DBOPEN in early March and Bob began coding DBUTIL in March or April.

This period, which extended into November of 1972, was characterized by an extreme shortage of computer and terminal resources. The systems lab had three prototype HP3000s, one of which was named PP2. The few available terminals were all hard-copy teletypewriters.

The competition for these resources was greatly relieved when Mike Green introduced an auction technique whereby each programmer could bid for machine time. Each Friday weekly "calendars" (1 per system), were posted. Each programmer was provided with 100 "virtual" chips each week to use in bidding for the machine hours of his choice. Time slots in consecutive 1/2-hour increments were bid for by filling them in on the appropriate "calendar." This bidding was frozen at 4:30 pm to cover all hours up to 4:30 pm of the next working day. If some other programmer made a higher bid (at least 10 chips)

for your time-slot (or the same bid for a proper subset of your time-slot), your earlier bid was cancelled. Any time-slots not auctioned off in this way could be reserved by anyone at no cost.

Since the personnel in our section were not working on first release projects, we were not allowed to participate in this auction. We were, however, allowed to join the 4:30 pm rush to reserve some of the unsigned-for time-slots. Even in this we were limited to time-slots on PP2, the system that was frequently cannibalized to keep the other systems running.

This meant that our use of machine-time was normally limited to nights, weekend s and holidays. It also meant that we didn't know what our machine-time schedule would be until late afternoon. This put quite a crimp in our family and social lives.

All of this, together with the fact that IMAGE was being developed in parallel with MPE, the file system, the SPL compiler and the editor, lead us to develop IMAGE on 80-column cards. In this way we could maximize our utilization of the few machine hours we had. During the day, we could key punch all new code and be all set to go, using the machine solely for compilation and testing.

$INCLUDE files were not supported by SPL in those days but were easily implemented with cards containing our global declarations.

Such was life in the "good old days."

Late pregnancy   This period started in November of 1972 (when it was realized that we were slipping schedule) and ended with product release in the fall of 1974.

Ed Estes and Mary Berner were added to the project team in December of 1972. Both of them were brought up to speed in January of 1973.

Mary coded DBUPDATE before being transferred to the QUERY/3000 project which had re-appeared on the scene.

Ed worked with Jon in developing the routines common to DBPUT and DBDELETE and went on to code DBDELETE.

Bob coded DBGET prior to leaving the project in the fall of 1973.

Jon completed DBPUT by mid-1973 and then took on the chore of performing product quality assurance. From IMAGE's nearly bug-free record, it is apparent that Jon was the perfect choice for this job.

Sometime in 1973, more terminal and computer resources became available and our working life became somewhat more structured.

By the fall of 1973, IMAGE was ready for extensive testing. Actually, it was already being tested by SIS and SAS which were both IMAGE-based projects. Our major holdup turned out to be the late assignment of a technical writer to the project. Sam Boot was assigned in October 1973.

Having authored the IMAGE/2100 manual (which was more like QUERY) Sam quite reasonably assumed that the IMAGE/3000 manual would be easy. After some start-up difficulties he realized that the HP3000 and IMAGE/3000 were much more complex. Before he could hope to understand IMAGE/3000 he had to at least understand how to log on and run HP3000 programs and acquire experience and familiarity with the HP3000 file system, the Account/Group/User structure and the file security mechanisms.

Happily, Sam did this and, with assistance from Jon and me, was able to complete the IMAGE/3000 manual by the fall of 1974.

In December of 1973 our section was disbanded. The IMAGE project was returned to the systems lab and was informed that all members except one would be re-assigned to other projects.

Lee Johnson was elected to move to Cupertino to tackle development of a divisional order processing system. Each of us was invited to choose between returning to the Systems Lab or joining Lee in this new project.

Having been in the Systems Lab before, I elected to remain with Lee. Jon Bale was the only team member who returned to the Lab. This was very fortunate for the IMAGE project, since, in addition to helping Sam in developing the manual, Jon made a major clean up and consolidation of all the IMAGE code. He improved the documentation and the performance and unearthed a few previously undetected bugs. He also developed many of the test programs still employed by quality assurance.

All members of the project team are proud of our achievements, a few of which are:

Achievements

All IMAGE/3000 intrinsics are callable from SPL, standard FORTRAN and standard COBOL. This was achieved by:

1. not using typed procedures
2. not using option variable

New standards in software interfaces

3. not using value parameters (using word-reference parameters only)

In contrast, less than 15 of the 119 MPE intrinsics have these characteristics and about 8 of those (such as DEBUG) owe this to the fact that they are parameterless.

All IMAGE/3000 intrinsics have extremely similar parameter sequences making them easier to learn and remember and less bug-prone.

All major IMAGE intrinsics perform extensive parameter consistency checking and provide ten words of information describing the result of the call. Supplementary intrinsics may be used to transform this information into human-readable form.

Contrast this with the MPE intrinsics, some of which don't even set the condition code and all of which return little, if any, information to the caller.

The principal IMAGE/3000 intrinsics all have a mode parameter which provide an open-endedness to the IMAGE sub-system. Subsequent enhancements are readily added by simply adding another valid mode value. This also made it possible to retain backward compatibility. With few exceptions, any program which ran on IMAGE/3000 in 1974 can run on IMAGE/3000 today.

IMAGE/3000 data structures (memory resident, disc resident and tape resident) include identifying information which is used by IMAGE for fault detection and is (or can be) used by other systems (such as OPT/3000) and users to assist them in their use of the HP300.

New standards in data structures and data structure handling

For instance, the first 6 characters of IMAGE extra data segments are of the form "IMAGEX" where X = 1, 2, 3 or 4 as determined by the type of segment. IMAGE uses these to ensure that it is dealing with the correct extra data segment. They also come in handy in identifying IMAGE data segments in any memory dump. DBUNLOAD, although developed in 1972 and 1973, has pr oven to be one of the two most reliable tape handling routines provided by the HP3000. It generates labeled tapes, handles multiple reels with the ability to restart at the beginning of a reel and provides end-of-dataset progress reports. DBUNLOAD also verifies all write recoveries.

DBLOAD, the other most reliable tape handling program, verifies all reported tape read recoveries. It also performs tape record sequence checking and extensive repositioning efforts in the face of sequencing or reading errors as well as progress reporting on a dataset-by-dataset basis.

Although far from perfect, the IMAGE/3000 manual had examples (which actually work) in COBOL, FORTRAN, and SPL.

The manual description of each IMAGE intrinsic includes a description of each exceptional (or error) condition.

Contrast this, for example, with the descriptions of the file management intrinsics where the user is referred to FCHECK and/or the Appendix for a table of error codes which are supposed to help you determine the error. The two tables don't coincide and, even more important, the user is left to guess which error codes might result from a particular intrinsic call, and the circumstances which might cause them.

These internal features didn't just happen. Each was thoroughly discussed, carefully planned for, and exhaustively tested by the members of the development team.

The presence of such internal features distinguishes the output of product developers from the output of programmers.

Product developers are thorough in their efforts to anticipate and cope with system, user and operational problems. Programmers are simply solving a programming problem, not unlike a graduate school project.

Product developers also know that "project completion" is analogous to giving birth. It is only the first step of a life-long process. Projects are "complete" when the product dies.

It is interesting to note that little, if any, effort has been made to include any of these features in any subsequent HP3000 software.

Santayana once said "He who forgets the mistakes of the past is doomed to repeat them."

To this could be added "He who doesn't recognize the mistakes of the present is doomed to propagate them."

One of the better HP/3000 manuals

Giving birth