

## ACE: A Distributed System To Manage Large Data Archives

Dr. Mike I. Daily

Mobil Oil Corporation, P. O. Box 650232, Dallas, TX 75265-0232  
 Phone: 214-951-2651 Fax: 214-951-3923

and

Dr. Frank W. Allen

E-Systems, Inc., P. O. Box 660023, Dallas, TX 75266-0023  
 Phone: 214-205-7510 Fax: 214-205-7200

### Introduction

Competitive pressures in the oil and gas industry are requiring a much tighter integration of technical data into E&P business processes. The development of new systems to accommodate this business need must comprehend the significant numbers of large, complex data objects which the industry generates. The life cycle of the data objects is a four phase progression from data acquisition, to data processing, through data interpretation ending finally with data archival (Figure 1.) In order to implement a cost effective system which provides an efficient conversion from data to information and allows effective use of this information, an organization must consider the technical data management requirements in all four phases. A set of technical issues which may differ in each phase must be addressed to insure an overall successful development strategy.

The technical issues include standardized data formats and media for data acquisition, data management during processing, plus networks, applications software and GUI's for interpretation of the processed data. Mass storage hardware and software is required to provide cost effective storage and retrieval during the latter three stages as well as long term archival.

Mobil Oil Corporation's Exploration and Producing Technical Center (MEPTEC) has addressed the technical and cost issues of designing, building and implementing an Advanced Computing Environment (ACE) to support the petroleum E&P function, which is critical to the corporation's continued success. Mobil views ACE as a cost effective solution which can give Mobil a competitive edge as well as a viable technical solution.

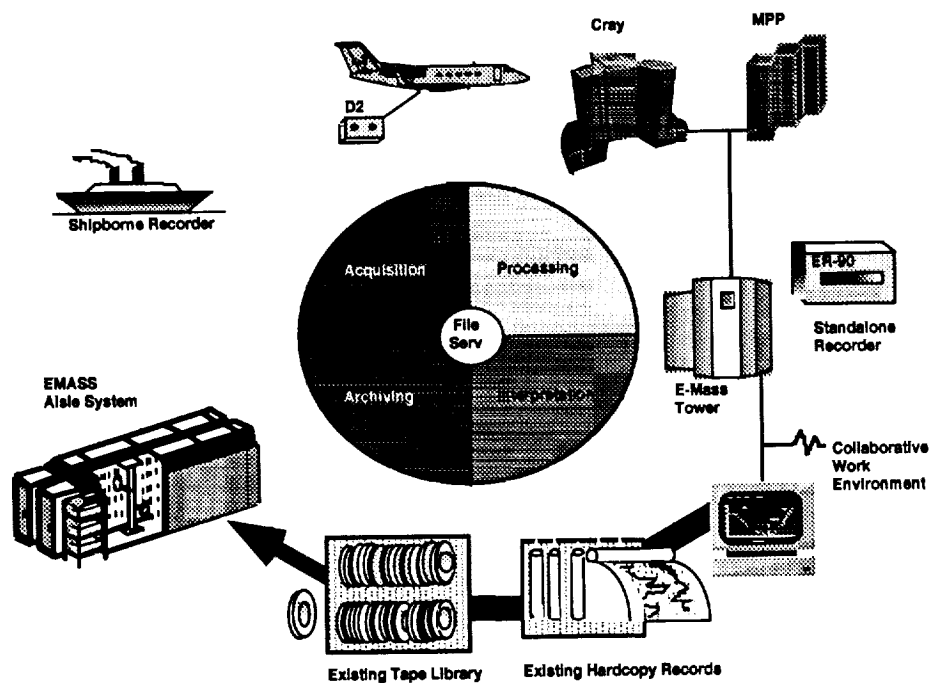
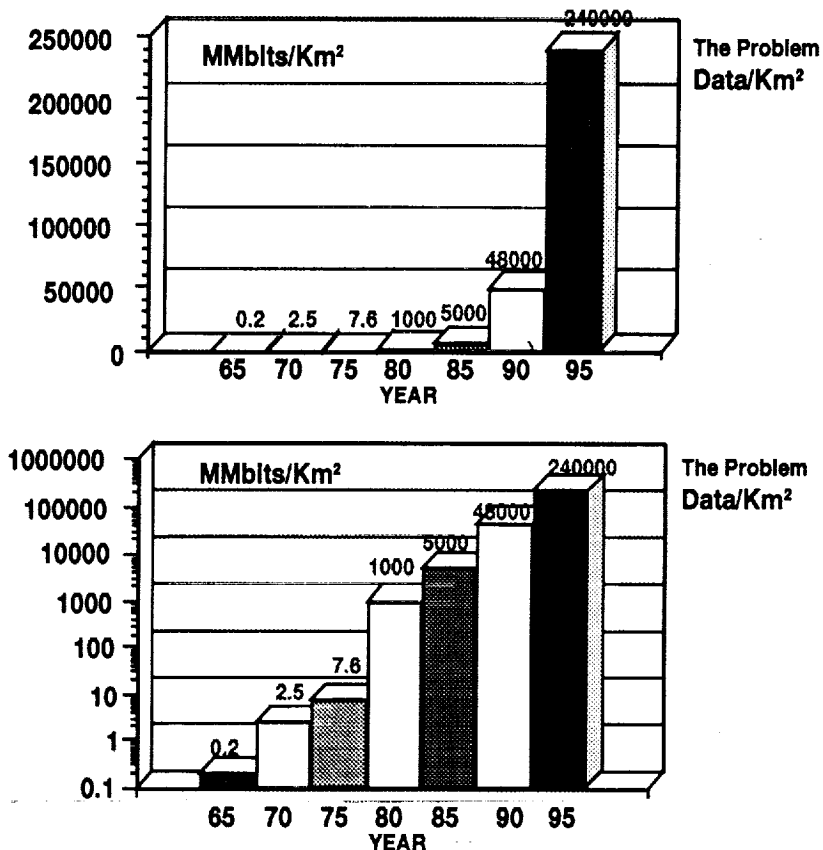


Figure 1. Data Life Cycle

**Acquisition**

The search for hydrocarbon accumulations requires an analysis of the earth's subsurface using the seismic reflection technique. Seismic data sets are acquired by land and marine crews over areas of interest and organized into surveys which are then transformed to 2-D or 3-D images of the subsurface. The increasing use of 3-D surveys in field exploitation has reduced the percentage of dry holes drilled from approximately 70% to 80% in the 1970's to 20% to 30% in the 1990's by providing more accurate and comprehensive geologic information. This reduction is significant when the cost of drilling a well in deep water exceeds \$100 million. But the trend to 3-D, and denser spatial data sampling has resulted in survey data sets which are terabytes in size. A single seismic acquisition vessel (there are approximately 90 in operation today) may collect 240 channels of seismic data every 12.5 meters using a 2 millisecond sampling rate. This amounts to 4 GBytes of data collected each hour or a terabyte (TB) every 10 days. As Figure 2 indicates, the trend since 1965 is a 5-10 fold increase each year in the amount of seismic data collected per square kilometer surveyed.



**Figure 2. Seismic Data Volumes from Martin Thompson & Ian Jack, Seismic 92**

The challenge in the acquisition phase is how to contain the increasing cost of seismic data collection and storage using a standard mass storage technology which is generic to the follow-on processing, interpretation and archiving phases. The storage media technologies in use today are 9-track tapes and 3480 cartridges. A 10 TB survey requires approximately 66,000 9-track tapes or 50,000 3480 cartridges with a total media cost of \$500K to \$1000K. To transport the survey to be processed, as well as to replenish the supply of media, requires a costly port call by the seismic vessel. Then the land/air transportation to the seismic processing center may well exceed \$100K. The bottom line to the acquisition contractor and ultimately the end user is a 3-D survey which may approach \$100 million in acquisition costs!

The acquisition contractors are always evaluating the latest storage media technologies for adaptation to seismic vessels. Of particular interest today are the helical scan technologies because of the higher media densities, faster transfer rates and increased reliability. Mobil is cooperating with acquisition contractors to evaluate the 19mm D-2 media technology. The D-2 technology enables the storage of a 10 TB survey on 400 small (25 GB) cassettes which cost \$17K, 140 medium (75 GB) cassettes which cost \$13K or 70 larger (165 GB) cassettes costing \$15K. Port calls will not be required to off-load the D-2 cassettes or to replenish the D-2 supply nearly as often and the projected transport costs to the processing centers will be two orders of magnitude lower than current costs. The ability to make a backup copy of the survey field data, something not done today, prior to transporting to the processing center is a key advantage to insuring the security of the survey against catastrophic loss. In the past, entire surveys have been lost in transit and had to be reshot. The backup of the survey will result in lower insurance premiums.

The data transfer rate and reliability of the D-2 technology is also important to the data acquisition process. Faster sampling rates and increased number of input channels in the future translate to higher bandwidth requirements. The D-2 recorder is capable of sustained transfer rates of 15 MB/sec and the reliability of the D-2 recorder has been measured at one permanent write error per TB with a 99% confidence. The features of the D-2 technology have led major oil producers such as Mobil, Shell and British Petroleum to request the development of a standard D-2 tape exchange format for seismic data by the industry standardization bodies including the Society of Exploration Geophysicists (SEG), International Association of Geophysical Contractors (IAGC) and Petrotechnical Open Software Corporation (POSC).

### Processing

The processing of seismic field surveys to develop 3-D images of the subsurface consists of several computation steps. But before the computations begin, the field data media must be manually mounted and the data transferred into the computational engine. This step can take months in the case of a 10 TB survey stored on 9-track or 3480 media due to the thousands of manually intensive tasks required and the relatively slow transfer speeds from the 9-track and 3480 recorders to the compute engine. Estimates of the cost of this step range from \$1 to \$2 per media when the manual handling, data administration and storage of active data are taken into account. Mobil has minimized these costs through the use of D-2 media and the EMASS® DataTower™ from E-Systems. The DataTower™ is a robotically controlled mass storage device about the size of a soft drink vending machine with a capacity of 5.7 TB of data stored on 226 small D-2 cassettes. The D-2 cassettes are accessed within 30 sec and loaded into one of four ER90™ D-2 recorders contained in the tower, each of which can transfer data at up to 15 MB/sec to Mobil's Convex C3220 file serving computer. In the future, Mobil plans to migrate all active and archived data to an EMASS® DataLibrary™ which is scalable to a 10,000 TB data capacity and bandwidth capacity which matches any commercially available supercomputer or MPP.

Each computational step to convert field surveys to image data requires careful analysis with intermediate data sets and partial test data sets created by different algorithms with multiple analytical parameters tuned for differing geophysical subsurface properties. A large 3-D survey can take months to process on the largest vector supercomputers. Mobil is reducing the time required for each of the computational steps by using a CM-5 massively parallel processor (MPP) from Thinking Machines. The EMASS file-serving Convex platform is connected to the CM-5 by an Ultra Network Technologies HiPPI channel which sustains a bandwidth of about 10 MB/sec.

It is desirable to store the interim results of the computational steps because the process is recursive, plus the results of step  $n+3$  may indicate that a return to step  $n$  is necessary because geophysical parameters used on step  $n+1$  were not optimal. Today the output of the current processing step is transferred to 3480 tape media because the amount of disk required to store these results is cost prohibitive and the earlier, interim steps are therefore deleted. The output of the current step normally requires many 3480 or 9-track media and results in additional manual intervention. Another I/O bottleneck occurs when the output data set on 3480 or 9-track media is used as the input stream for the next processing step. The supercomputer incurs an I/O wait while the slower I/O device transfers the data to the CPU and this I/O wait can amount to a yearly cost amounting to hundreds of thousands of dollars. This I/O wait is reduced significantly by using the ER90™ recorders as a virtual disk, storing the output stream and then transferring the data as an input channel to the next compute process.

The long, compute intensive processing steps are susceptible to errors inherent in the data storage media. Permanent write errors in the input data stream to the seismic processing procedures can cause abnormal termination of the processing and require restarts and/or reprocessing. The improved reliability of the ER90™ recorders and D-2 media reduces the risk of these occurrences and is a major reason why other major producers such as Aramco and Exxon are seriously considering the D-2 technology.

### **Interpretation**

Elements of the modern interpretation environment include high-performance X-based desktop displays, fast networks, tools for collaboration between remote sites, and seamless access to data. An advanced prototype environment, named MobilView, has been constructed to demonstrate key aspects of this environment.

Subject areas covered in the interpretation environment include:

- Relational (drilling, geoscience, and engineering)*
- Vector (downhole sensors, hydrography, political boundary)*
- Array (raw seismic, processed seismic, scanned photos and microfiche, scanned paper documents)*
- Other (grid/CAD files, multimedia, compound documents)*

The desktop user interface is geographical in nature, in line with users' mental models of the world. The user interface conforms to draft versions of an extension to the Motif style guide, which has been developed by an oil industry consortium known as the Petrotechnical Open Software Corporation (POSC). At the physical level, the underlying cartographic database has been organized using tree-structured tiling methods to ensure rapid data access over a wide dynamic range of scales.

The primary objective of MobilView is rapid viewing of large complex data objects, spanning a variety of formats. Secondary objectives include low-volume data ingest, file routing, and project archiving/recovery. Little emphasis is placed on actual computational processing, 3-D visualization, or hardcopy output. The viewing environment consists of a collection of Motif display programs ranging from purchased oil industry-specific tools to publicly-available image viewers. These are all integrated under a common shell and launcher environment that is fed by disk and cassette-based components of the storage hierarchy.

Image scanning and ingest in low volume are supported by a software environment that enables the user to pick an object (e.g., a seismic line) from the map or from a list, then scan in one or more hardcopy documents or images using a deskside scanner. The primary key association is made transparently and the user may then key in ancillary information about the scanned hardcopy. A browse-and-route function allows the user to browse through thousands of images and other large data types, and then route a file to destinations including the user's local workstation, a high-end processor (such as an MPP), or to a plotter.

Archiving functions have been developed to capture the results of long-running multidisciplinary studies. A named archive can be created and associated with a site, and files entered. Bulk data in any of the supported formats may be written to D-2 cassette and their associated metadata updated. Mandatory metadata includes an archive's geographic boundaries, thus enabling placement on the electronic map and use for later browsing. Upon selecting the archive, its contents are displayed for detailed browsing, display, or file routing over wide-area networks.

## Archiving

By emerging standards for archive size, the needs of a large oil company represent a medium-size (one Petabyte) problem. Data ingest consists of a mix of low-rate scanning input, medium-rate transcription from low-density tape, and direct insertion of D-2 tapes from offsite acquisition and processing activities. The upper limit of a petabyte is projected to consist of:

- hardcopy scanning = 400 TB total
- existing tape library = 200 TB total
- future (15 year) inflow = 400 TB total
  - interpretation results @10 TB/yr
  - acquisition/processing @15 TB/yr

The requirements for an archiving function include long effective media life, scalability, and reliability. Typical data has value for 15-20 years, comparable to the nominal lifetime of most magnetic tape media. Given the large number of files per tape on D-2, the likely failure mode becomes one of mechanical wear and tear. This occurs at approximately 1000 mount/dismount cycles, estimated at 2-3 years. The EMASS FileServ software enables automated transcription to be invoked after a specified number of cycles or at a given error rate threshold.

Scalability is important for supporting physically remote offices having relatively poor data communications service. Current plans are to configure non-robotic servers consisting of a pair of ER-90™ recorders managed by a RISC processor. With the ability of the recorders to use the large (165 GB) cassette, this gives a respectable 300 GB 'slow disk' facility. Bulk data transfer from the central archive could then be done on off-hours. The usual issues of synchronization, federation, etc. found in distributed database environments exist here as well,

At the high end, the archive must be designed to scale up from the present 10 TB systems to 1000 TB library systems. D2 Cassette replication will be needed to ensure backup and disaster recovery. Long-range technology planning for future media (optical tape, holographic) is simplified in a robotically-accessed environment having computer managed metadata. With increased data density from 3-D seismic data acquisition and the growth of full-motion video, the nominal one petabyte case may be overtaken by events later in the 90's.

## Conclusion

The oil and gas industry is currently one of the largest application areas for high-density mass storage technology. Current immaturity of the technology and standards forces the use of rather custom systems; by late in the decade, however, off-the-shelf one petabyte systems should be readily available. At the high end, Grand Challenge problems will spur the development of large integrated systems, while sub-petabyte systems will be commodity items in use by thousands of organizations. The seismic problem is a challenging one. The global competitiveness of the U.S. oil industry depends on solving this problem.

