Computer technology: From punch cards to clustered supercomputers

LAWRENCE M. GOCHIOCO, GX Technology, Houston, Texas, U.S.

In general, geophysicists are very analytical, methodical, and quick to adopt and implement new ideas and concepts. Indeed, the smartest geophysicists of the past generation were often so far ahead of the curve that they had to wait patiently for computer technology to catch up so they could test/implement innovative algorithms to solve complex 3D problems. Thus, there is a symbiotic relationship between the geophysicists and computer technology. In order to appreciate the current computational prowess that the energy industry employs, it is necessary to reflect on the past 20 years of progressive development. In particular, how the seeds were planted that eventually germinated into this explosive supercomputing status.

When I started my geophysical career 20 years ago, my first employer was a small data processing company that had limited financial resources. As a result, most of the computer equipment used to process seismic data was refurbished old hardware from the late 1970s. Data were saved on 12-inch reel tapes and were uploaded via tape drives at 1600 BPI. The computer was programmed via punch cards—a rather cumbersome ordeal because extensive processes required hundreds of them. Extreme care in handling was mandatory because it took only one out-of-sequence card to cause an entire processing job to crash or bomb. It was common then to have the punch cards safely stored in boxes when submitting the jobs.

At about this time, larger geophysical service companies and major oil companies relied heavily on mainframe computers (like the IBM-360) to conduct in-house data processing. Service companies were also fierce competitors in marketing their stand-alone seismic data processing systems, based on the 32-bit floating point array processor. The array processors expanded the computational aspect of some algorithms (e.g., *f-k* filtering and migration) through Fast Fourier transforms. Some of the hardware systems sold at that time were the TIMAP (GSI), LRS-3200 (Litton Industries), Phoenix (SSC), Tempus 32 (Geosource – Perkin Elmer), and Megaseis (Seiscom Delta). If you recognize the names and used some of these systems, you are a genuine old-timer.

By the end of 1983, some systems were capable of processing seismic data at computational speeds of 30 Mflops, I/O bus rates of 20 Mbytes/s, and calculating 1024-point complex FFT in 1.7 ms with 64 Mbytes of data memory. Memory size was still a major problem. Thus, it was common for these systems to have three tape drives. The first two drives were used as I/O with the third as a backup.

The geophysical industry was still reliant on the 12-inch reel (3600 ft) magnetic tapes as the media. Even though the read-write density increased by a factor of 4, from 1600 to 6250 BPI, it was common for major service and oil companies to maintain tape libraries of warehouse dimensions. Some company tape libraries contained tens of thousands of reels.

The geophysical service industry was able to capitalize on the incremental technology improvements in the fields of data acquisition and processing and larger 2D and 3D surveys began to be conducted. Continued software and hardware development eventually led to the deployment of tools that would enhance the interpretation process. By the mid-1980s, several service companies began marketing their first-generation seismic interactive interpretation workstations (IIW). The IIW allowed the integration of various geophysical and geologic data sets for improved analyses, which resulted in shorter interpretation cycle times. This was a major breakthrough in bridging the historic communication barrier between the geologists and geophysicists and soon resulted in a revolutionary concept called "multidisciplinary teams." However, it would take another few years before the full potential of the IIW would be realized and widely adopted by the petroleum industry.

By the mid-1980s, as more 3D seismic data were collected and 2D prestack time and/or depth migrations were gaining greater acceptance, pressure built to shorten processing turnaround times (which based on the biggest projects could last over a year). As a result, some companies acquired the fastest processor of that time—the CRAY YMP supercomputer. It became common for service companies to advertise their access to new CRAY supercomputer (hopefully to win more contracts), and some oil companies were proud to have one of their own (even designing buildings specifically to accommodate them). However, while the CRAY machine received all the media attention, an unforeseen revolution was brewing behind the curtain and the computer industry—the geophysical industry's introduction to the lowly personal computer (PC) circa 1986. There were many skeptics who felt the PC was just a simple tool (even a toy) and could never accomplish any but the most basic applications. However, a small group of visionaries saw the upside potential of PCs because of its tremendous cost advantage relative to the standard mainframe computers. Due to its limited processing speed and memory capacity, it was initially relegated to performing simple processing tasks in the field (like MicroMAX) and basic processing and modeling functions in the office.

Toward the end of the decade, the geophysical industry concentrated its resources in improving the multitasking capabilities (with enhanced graphics and memory) of the interactive interpretation workstations because multidisciplinary teams proved to be successful in addressing large 3D seismic data volumes that needed to be interpreted quickly. Concomitant with the events noted previously, the PC also experienced greater enhancements. Due to the PC's tremendous cost advantage and simplicity, geophysicists who grew tired of big oil company business cycles started their own companies. Microcomputer hardware and software improvements spawned many innovations such as the menu-driven user interface, multitasking, enhanced graphics, and window-based interactive processing techniques. These developments also facilitated the rapid growth in the fields of environmental, engineering, and mining geophysics. During this period, SEG membership in the petroleum industry declined because of poor business conditions, but membership related to near-surface geophysics grew.

As expected, CRAY supercomputers that we brought online in the mid-1980s were soon "pegged to the max" a few years later, as geophysicists implemented other geophysical technologies such as 3D reservoir characterization and other simulations, as well as implementation of the 3D prestack time migration and 2D prestack depth migration. From the late-1980s to early 1990s, geophysicists expanded their knowledge and experience from the traditional (acquisition, processing, and interpretation) base to include programming, system designs, etc. that brought geophysical expertise to other disciplines.

Because the petroleum industry was a major market for computer hardware, geophysicists leveraged this position and formed SEG committees to study and recommend to computer manufacturers what their needs were and were likely to be. An amazing thing happened—*they listened*. Geophysicists wanted more data storage and memory, standardization, and networking capabilities. They got it—and in pretty short order. Standardization was the key in keeping costs down because proprietary operating systems tended to lock the buyers into that vendor, limiting their choices. With standardization, the end users were able to buy the best performing machines at the right price from any vendor. For example, it became obvious to one SEG committee that the UNIX operating system would dominate the geophysical community and not the OS/2 system. Thus, this group supported the former operating system. The ultimate end result was that, finally, the manufacturers were building machines that geophysicists wanted and needed.

Simultaneously, the success of the IIW led to design enhancements and the introduction of RISC (reduced instruction set computer) workstations. The 1990s brought about stiff competition in the desktop (minicomputer) workstation market, especially among the UNIX-based SUN Sparc, DEC, and SGI machines. Later in the decade, PCs with the Linux operating system began to compete with the established minicomputer workstations. As usual, how the system was configured and networked to lower operating costs dictated which architecture would prosper.

By the early 1990s, the CRAY YMP supercomputers were showing their age, and were on their way to being replaced with MPP (massive parallel processors) systems. Traditional computers used a single CPU to solve computational problems. With MPP, many interconnected CPUs could be used to simultaneously process fragmented or partitioned data, resulting in faster job completions. The RISC workstations and MPP systems eventually became the computing workhorse in the early- to mid-1990s. Some key performance parameters of an MPP system, based on a 128-node Intel i860 supercomputer in 1992 were

- Peak computational speed = 10 Gflops (billion) floating point operations/s)
- Distributed memory = 2 Gbytes
- Disk storage = 96 Gbytes
- Online tape storage = 1.5 Tbytes

The shared-memory computer initially inhibited the performance of MPP systems because the increasing number of CPUs gaining access to the main memory became a bottleneck and performance was impaired. The problem was corrected by distributing memory between the processors and interconnecting them via a communications network to form a distributed memory bank. Combining the technologies of distributed memory and adding more and faster CPU chips increased computational speeds to 100+ Gflop. Unfortunately, even though MPP systems were computationally faster than their predecessors and unit processing costs ($/million seismic traces) dropped significantly, they were still expensive machines to acquire because all those parallel CPUs were housed inside one massive computer. Thus, MPP systems were only within the financial parameters of large corporations and government research institutions.

As usual, university professors, scientists, engineers, and students (probably including geophysicists) who were left out of the expensive MPP world searched for low-cost alternatives (circa mid-1990s). Using the same concept of parallel computing, they discovered that networking off-the-shelf (surplus?) PCs could achieve performances nearly as fast as MPP supercomputers at much lower costs. The networked or "clustered" PCs were easy to maintain and were scalable depending on the required computing jobs. In addition, the operating system (Linux) originally developed for this hardware design was *free*, and floated freely within the scientific community with continued enhancements. By the late 1990s, the geophysical industry began to adopt the PC-based "clustered supercomputer" architecture and operating system. Currently, they are now the supercomputer of choice. Typical data processing centers that employ the clustered supercomputer design can have more than 1000 installed CPUs and can attain peak performance of more than 1 Tflops (trillion floating point operations/s).

By coupling the clustered supercomputers and enhanced desktop workstations, geophysicists are now able to process and interpret very large-scale 3D seismic data volumes quickly and with confidence, using the latest 3D wave equation prestack depth-migration methods, 3D visualization techniques, more accurate 4D time-lapse and simulation studies, multicomponent processing, etc. In only about 12 years after its introduction to the geophysical industry, the PC became the dominant computing platform—imagine that. However, visionaries of the mid-1980s never doubted its upside potential because the *cost factor* has always been the driving force in the energy business and it favored the PC.

What will the next 20 years be like? I could only cautiously speculate, but I know it will exceed my current imagination. By then, I would be 66 years old (retired?). However, I know that I have selected a noble profession where I can see the positive influences brought about by the works of geophysicists. Not only do they help secure the global energy supply, but also make solid scientific contributions to society. Our innovations to signal processing theories, sophisticated 3D imaging algorithms, supercomputers, and 3D visualization techniques have advanced other scientific disciplines like medicine, meteorology, life sciences, etc. As new 21st century exploration challenges take us to other frontiers such as in greater oceanic depths and deep space/planetary exploits, I know for a fact that geophysicists will be involved tinkering to build a better mouse trap. TLE

Editor's note: Rather than go into a lot of detailed hardware specifications and performance, which is available in other publications, the author wrote this article to demonstrate the dynamic relationship between the geophysicists and computers from the point of view of an end user. Readers desiring more detailed information are referred to the special sections on computers previously published in TLE *(September 1986, April 1988, March 1990, January 1991, and December 1996).*

Corresponding author: lgochioco@gxt.com