

The Cray Y-MP EL: An overview

By Clive England



At the end of a virtual tour of The National Museum of Computing, I saw an old friend. Standing quietly by the very impressive Cray-1 from the Royal Aircraft Establishment was a grey and red cabinet about 5 feet tall with a big red button on top. This diminutive Cray supercomputer is the self-contained, air cooled, Y-MP EL. This “Entry Level” machine provided a hardware and software environment replicating its much larger and more expensive Cray cousins. During the 1980s the Cray-1 successor the Cray XMP, and later the Cray YMP, provided the backbone of scientific supercomputing around the world. As chip technology progressed in the late 1980s, along with an economic slump, and competition from the likes of Convex, the market opportunities for smaller more cost-effective scientific computing platforms became apparent. Starting with the Cray Y-MP EL (Entry Level) designation, later systems became the Cray EL98 with more

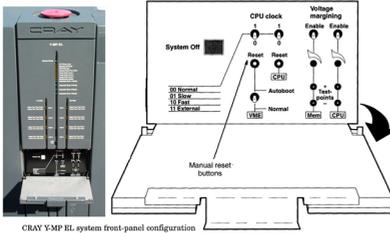
memory and processors. A smaller cabinet version was released as the Cray EL92. Key features across the EL range were vector-enabled processors, multiple sets of fast registers, fast and parallel memory interconnects but with no virtual memory or paging.

The famous designer Seymour Cray split from Cray Research to pursue the Cray-3 project in '89 while Cray Research developed the XMP then YMP and later C90 & T90 architectures throughout the '80s and '90s. Cray supercomputers were sold for their ability to crunch huge number sets in the shortest possible time. Delivering accurate weather forecasts, shaping a aircraft wing, or solving complex financial equations, all rely on high performance processors to deliver timely results. The EL described here is one of the smallest Cray machines but does have Seymour Cray-inspired architecture in its bones.

A technology start-up called Supertek created a CMOS implementation of the Cray XMP processor architecture with a view to complementing the Cray XMP at a much lower price-point. Launched in '89, Supertek managed to sell about ten “Supertek S-1” machines and then started development of a Cray YMP processor implementation. After Cray acquired Supertek in 1990, the S-1 was briefly released to market re-branded as the Cray XMS. In practice the processor emulation was well received but the input / output subsystem (IOS) was not robust enough for production supercomputing. Using a single RAID storage device, seek contention severely hampered system performance. A couple of years later in '91 the Cray Y-MP EL was released; an evolution of the XMS with a similar VME bus based input / output subsystem but emulating the later Cray YMP processor architecture. Offloading input / output operations to a specialist processor was a tactic deployed across the Cray system range, supported by direct memory transfers between IOS processors and main memory. This solution also facilitated two separate system development threads allowing the Cray range to be refreshed by alternating the release of systems, with new main processors or new ranges of storage and network devices.

Hardware tour

Weighing in at 635 kg with a 6K Watt power consumption, the Cray EL had a one square meter footprint and could be delivered through most doorways and lifts. One Eastern European university however did have to strip the skins, remove the disk and tape drives then have six large workmen manhandle the chassis up two flights of stairs.



CRAY YMP-EL system front-panel configuration

At the front of the chassis the most obvious feature is the big red emergency “power off” button on the top of the cabinet. As this button is just about at leaning height for taller folks some sites put a machined Perspex tube around the button, other sites used part of a two litre Coca-Cola bottle or even an upside-down painted Coconut shell to prevent unexpected shutdowns.

Just below the red button are the system fault and status lights and a concealed flap covering the reset, master clear button and a couple of diagnostic switches.

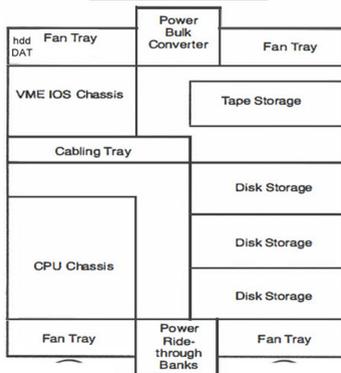
Looking at the specific hardware internals, both the left and right sides of the main cabinet are hinged to reveal, on the left the input / output subsystems (IOS) cage above the processor and memory boards and, on the right, the disk and tape storage devices.

The left-hand side of the cabinet holds the two types of processor found in the EL. The full width VME card cage, near the top, could be sub-divided into up to four independent input / output subsystems. Each IOS consisted of a Master I/O Processor, an I/O buffer board and customer selection of VME peripheral cards. The first IOS also connects to disk drive and system service QIC or DAT tape and provides the console connection point for the system. The IOS disk contained the VxWorks based IOS operating system, the main system configuration files and diagnostics. The main processor and memory cards in the lower cage ran the Unix based UNICOS operating system to perform the actual customer workload.



Left side view

CRAY
RESEARCH, INC.



Rear Xray view
Cray YMP-EL

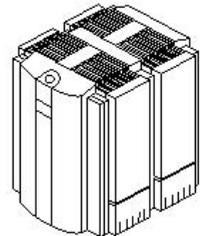


Right side view

The tall cards at the base of the left side of the system are the main processors and memory boards. The connection between the master IOS and first processor was via a 40 Mega bytes per second rainbow-coloured ribbon “Y1” cable. An Ethernet card was also included in the first IOS with fibre optic and Hyper channel cards also available. Early EL systems were available in one to four processor versions with memory sizes from 256 to 1024 Mega bytes with a memory bandwidth of 4.2 Giga bytes per second.

The main memory cards, available in four sizes, have multiple banks with error and port conflict resolution hardware. Each processor in the system has four parallel memory ports. Each port performs a specific function, allowing different types of memory transfers to occur simultaneously. To further enhance memory operations, the bidirectional memory mode allows block read and write operations to occur concurrently. Memory logic hardware included SECDED (singlebit error correct, doublebit error detect) error detection to avoid silent corruption errors. When a single bit error is detected, the correction logic resets that bit to the correct state and passes the corrected word on. When a double bit error is detected, to avoid unpredictable results, a memory error flag in the running program exchange package is set and referred to the operating system.

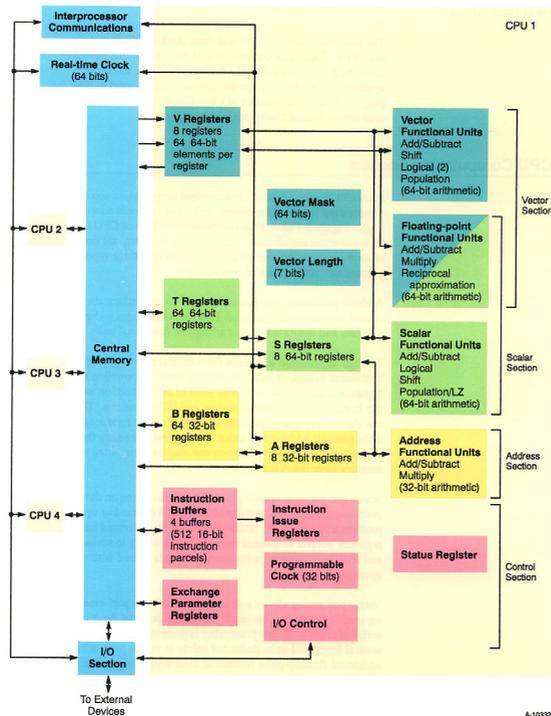
The right-hand side of the cabinet holds the main hard drives and the optional tape units. An autoloading low-profile reel tape drive or a cartridge 3480-type drive and an 8mm helical scan tape system were the tape options offered. Early EL systems had two types of main storage hard drives, IPI for high performance or ESDI for larger storage size. Only two IPI drives fitted in a drive bay but eight ESDI drives could be put in the same space. Later EL94 and EL98 systems had SCSI connected DD5-s 3.1Gb Seagate disk drives. Disk array sub-systems with hot spares and auto remapping of flaws were available. Customers could later purchase an SCSI card to attach external generic hard drives. Customers could order double or triple wide main cabinets to hold extra input / output processors and peripherals.



Cray processors

The Cray processor internal architecture was built to be fast, concurrent and feature rich. To the modern eye, the system looks very old-school with its different registers for different data, no virtual memory or fast context switching but the system was designed to match the larger Cray YMP processors whose development heritage traces directly back to the innovative Cray-1. The later Cray architectures built on the Cray-1 processor design by adding more parallel processors, more and longer vector registers, and more “other” registers. A flexible instruction set combined independently and concurrently, in the right conditions, scalar, vector and address data calculations.

Each of the processors included multiple registers to hold data and functional units to process data (see diagram). Vectors, being a list of values, play a critical part in achieving high performance calculations. The eight sets of V registers are each 64 words long and 64 bits wide. The processor is able to stream from memory directly to the V, T, A, or B register sets, between vector registers, via a functional unit and back to memory or to an output device. These streaming operations in combination with pipelined functional units allow up to 64 calculations to be performed in hardly more time than it takes to pull the values from memory.



CRAY Y-MP EL System CPU Block Diagram

The system memory is arranged into sections and banks to allow simultaneous and overlapping memory references. Simultaneous references are two or more references that begin at the same time; overlapping references are one or more references that start while another memory reference is in progress.

Special processor operations such as gather/scatter, bit population count and leading zero count supported specialised software used in signal processing at certain valued government customer sites. An Exchange package and Exchange registers combine to hold the state of a processor at the point that a processor switches to another task. Having many registers made context switching a slow operation, predisposing the machine to run best with large industrial strength workloads.

The interplay of registers, memory references and functional units resulted in a complex processor architecture. Constructed from CMOS Very Large Scale Integrated Circuits (VLSI) the processors ran with a 30 nano second clock speed delivering a peak performance of 133 Mega Flops (Million of floating point operations per second) enabling an EL98 system to achieve a prestigious one GigaFlop rating.

What it was like to service these rare and complex machines?

When I was a software support engineer for the UK Cray EL systems, a customer asked “Why did the process accounting fail to run last night?”. I started looking around the system using the remote support modem for access and the system was behaving normally. However the cron.log file showed many, time triggered, but rejected cron jobs. The immediate cause was the cron job queue reaching its maximum of 25 outstanding jobs. This was the reason that accounting process had not run but what was filling the cron queue? It was clear from the ps command output that olcrit, the on-line processor health check command, was the problem. Run hourly, from cron system diagnostics would run for a second or two to check the system health. Here were 25 sets of diagnostics queued waiting to run. The olcrit diagnostic is attached, to each processor in turn, using the ded(icate) command. Each olcrit program would start, wait until its nominated processor exchanged to the kernel then would run processor functional checks for a second and exit. Here was a queue full of diagnostics waiting to get into processor 2 that was stalled and not exchanging into kernel space. A site visit was arranged to check the contents of

the program counter in the stalled processor by running another diagnostic program on the master IOP. This showed that the processor was hung on a gather/scatter instruction, one of the two instructions (test/set being the other), that would cause early EL processors to stall. As a result of this incident and processor hangs at a couple of other sites, all the "PC-04" chips on every EL processor board were manually replaced at customer sites.

Software tour

The Cray EL has two operating systems; in the IOS a real time operating system based on VxWorks, and in the main processors UNICOS. VxWorks controls the I/O peripherals and most importantly, after cold startup checks are completed, started, stopped and diagnosed the main processor boards. During the start cycle the IOS would halt the main processors, push a boot loader into main memory and deadstart the main processors. The boot loader locates the root filesystem and the main operating system, UNICOS, would be started into single user mode. The final step was jumping to multiuser mode and the system was ready for service.

Cray EL systems ran UNICOS, a "full cream Unix", that was bit compatible with the UNICOS run on the full size multi-million dollar Cray systems. Designed to be multi-user from the outset UNICOS has a full suite of user resource controls able to track the memory usage, processor usage, process count, disk quota and even the quantity of printout generated. User controls are affected via a user database with fields and security procedures way ahead of the traditional Unix "/etc/passwd" password file. Full system process cost accounting was available for sites that needed to charge end users by resource usage. Process accounting was also a helpful diagnostic tool for retrospectively exploring system performance issues. Complaints of system slowness over the weekend or unexpected long program run times could be placed into context for system administrators and scheduling decisions adjusted. Extensive system logs and routine diagnostic runs would detect any emerging hardware errors before they became serious.

UNICOS came pre-installed on the system drives but could be reinstalled from scratch using the maintenance tape system. Early system updates included a tedious 10 hour UNICOS re-linking stage but this was mitigated by using "chroot" on an alternative root filesystem and later replaced with a binary distribution for major upgrade cycles.

Another UNICOS feature, inherited from bigger systems and optional for the EL was Multi-Level Security (MLS). MLS allowed Unicos to isolate groups of users into specific compartment(s) and level(s). Special commands and privileges were needed to move files between sub-sections. Restrictions on the output of commands such as ps and ls helped separate communities preventing back channel communications. The concept of an all-powerful root user was replaced by various system administration privileges. Popular with defence and government sites MLS UNICOS was not widely used by commercial customers.

Built into UNICOS was an early disk volume manager providing a couple of valuable performance features. Combining large numbers of disk drives into a single usable volume was facilitated by volume configuration options in UNICOS applied at boot time. Filesystems could be arranged so that file directory meta-data resided on faster

drives with the bulk storage on larger slower drives. The concept of using large chunks of memory as a solid-state drive, had been a part of Cray architecture since the early '80s. On the EL this feature appeared as the ability to add large chunks of memory as filesystem Idcache space improving the performance of high turnover system or work files.

Filesystem storage could be managed with individual user quotas and the Data Migration Facility. DMF pushed files on to a tape drive and released the disk space. Migrated files could be automatically recalled when later accessed. Used mostly on the larger sites that had access to an automatic tape library, DMF was a complex system that did cause some support headaches but could extend storage space from Gigabytes to Terabytes. Database controlled DMF, was no substitute for extra working storage but, given the expense of disk drives at the time, it was very cost effective when storing and using large datasets and for system backups. Backing up the residual file stubs and the DMF database, not using DMF, was considered good practice.

The real gem of Cray software was the Programming Environment. Inherited from the established Cray range, compilers were provided for C/C++, Fortran, Pascal and ADA backed by a host of development and code analysis tools. Getting the most out an expensive Cray machine involved carefully looking at how application programs performed. Some programs would port easily from workstations and immediately gain a moderate speed up but that alone would not justify the extra cost of a Cray. Techniques such as vectorisation, parallelism and use of optimised libraries could make dramatic program performance improvements. To ease the transition Cray compilers had features that would automatically recognise performance opportunities or the programmer could add directives that would invoke parallelism or direct the compiler to perform loop combining and loop unrolling. Tools such as :

- * ja - reported comprehensive statistics on time spent in program modes and memory usage.
- * flowtrace - reported which sub-routines are the best candidates for in-lining.
- * jumpview - reported which sub-routines use the most time.

With the correct compilation flags enabled, the compilers listed the performance areas identified in a program. This snippet shows the CFT77 compiler identifying the 'S' scalar outer loop and 'V' vectorizable inner loop that has been unrolled 'r'.

```

26      26.                               ** Fill Matrix 1
27      27.
28      28. S-----< .   do 40 i = 1, maxdim
29      29. S Vr-----<   do 50 j = 1, maxdim
30      30. S Vr                mat1(i,j) = cosvec(i) * sinvec(j)
31      31. S Vr----->50   continue
32      32. S----->40   continue

```

First seen in Cray language compilers, OpenMP exploits many parallel performance features. Today OpenMP is widely used on modern Intel-powered systems; by including a few well-chosen OpenMP directives, a program can get the most out of modern processors without having to rewrite the whole program as a multi-threaded application.

UNICOS provided a rich runtime work administration environment on all systems large and small. Centred around the Network Queuing Environment, an industry standard batch queuing service, systems could be loaded up with work from a diverse user base and run at full steam 24/7. Features such as checkpointing could allow long running jobs to be frozen and restored. System administrators trying to both fill the machine and provide the best turnaround for priority clients used "The Fair Share Scheduler". The workload balancing conundrum resulted in multiple work scheduling and queueing arrangements. One site reported "Fair Share is the standard scheduling algorithm used for political resource control on large, multi-user UNIX systems. Promising equality, Fair Share has instead delivered frustration to its users, who perceive misallocations of interactive response within a system of unreasonable complexity."

The Los Alamos site analysts went on to create an Opportunity Scheduler. Opportunity Scheduling differed from Fair Share scheduling in three key respects:

- * Past usage (individual or group) had no bearing on present service,
- * Resources were allocated to groups, not to individuals, and
- * Processor and memory scheduling are coordinated.

The site later reported "Opportunity Scheduling first made its appearance in October 1995 and has been running on all of our production systems since April 1996. The scheduler is running on two YMP-8s, a J-32, and a T-94. The modified kernel supports both Fair Share and Opportunity, Scheduling, switchable via a runtime flag."

Sites that used Cray EL systems

National Center for Atmospheric Research - A weather forecasting customer.

"At NCAR we have clustered four Cray EL series machines. The cluster consists of three EL98s, and a smaller EL92. The primary objectives for the cluster are to provide a comfortable development and production environment for work which does not require the computational power of a YMP. The development environment emphasizes quick response, fast turnaround and large memory for both interactive and batch work. The production environment provides similar goals, with less emphasis on turnaround. Fast turnaround is achieved by scheduling and load sharing."

Other weather sites with EL systems were the United Kingdom Meteorological Office, the European Medium Range Weather Forecasting (ECMWF) at Shinfield and Rosgidromet, the Russian Federal Service for Hydrometeorology and Environmental Monitoring in Moscow. All these weather sites had full size Cray systems alongside.

Fuji Capital Markets - A step into financial engineering.

"Fuji Capital Markets has ordered a CRAY EL94 mini-supercomputer system. The global derivatives firm will use the new system for derivatives simulation related to its trading operation, according to James Kennedy, Fuji Capital Markets Corporation's managing director. The CRAY EL94 system will be used to run Monte Carlo-based applications developed in-house by Fuji Capital Markets Corporation's technology group. The system will enable the firm to enhance pricing and hedging of its current and new derivatives products, as well as provide real-time turnaround on pricing requests from Fuji Capital Markets Corporation's customers. Kennedy said with this new Cray Research technology Fuji Capital Markets Corporation will be able to run hundreds of thousands of "yield

curve” scenarios for more accurate statistical analyses of Fuji Capital Markets Corporation derivatives products. The system will also be used for “stress testing,” a complex analysis used in defining Fuji Capital Markets Corporation’s portfolio risks. We were impressed with Cray and their level of understanding about our problems. In fact, we gave Cray’s expert staff a benchmark based on our in-house code and they quickly vectorized and optimized it for the CRAY EL94 system. The code runs much faster on the Cray Research system than it did on our workstations.... The new Cray Research system will be integrated with the 125 workstations on Fuji Capital Markets Corporation’s wide area network that links the firm’s London and Hong Kong offices.”

Other Financial sites with EL systems included Freddy Mac and Atlantic Portfolio Analytics.

Hawtal Whiting - Engineering heartland.

March 1, 1994 -- Cray Research, Inc. announced that “Hawtal Whiting Group, an engineering design consulting firm based in the United Kingdom, has installed a CRAY EL98 departmental supercomputer at its Basildon, Essex, headquarters, and has also ordered a next-generation Cray Research departmental system due out in 1995.

Hawtal Whiting will use the CRAY EL98 system primarily to run structural analysis applications. The CRAY EL98 system will enable the company to analyse substantially larger models in significantly less time. Colin Cox, senior principal engineer, said Hawtal Whiting selected the CRAY EL98 system after evaluating it against other supercomputers as well as networked workstation offerings. “The Cray Research system offered the best performance across the range of applications we typically run,” said Cox.”

The press release ended with this teaser ... “Hawtal Whiting purchased the CRAY EL98 system under a recently introduced Cray Research program that offers customers lower combined pricing on orders of a CRAY EL98 system and a next-generation departmental supercomputer expected to be available for delivery in the first quarter of 1995. These next generation systems are expected to be 6 to 12 times more powerful than the CRAY EL98 supercomputer at approximately the same price point.”

Other Engineering customers included: Raytheon Aircraft Company, Canadair, Lotus Engineering, Motorola’s Automotive and Industrial Electronics Group for material science and British Gas’ Gas Research Centre based in Loughborough, England.

The EL was popular with universities, academics and automotive sub-contractors being purchased on a departmental budget and installed without special cooling and power arrangements. EL systems landed at Cranfield University, Bedford, Queens University Belfast, University of Texas and many others. The Cray EL on display at TNMoC was originally at Swansea University where it was used to develop and run computational fluid dynamics codes.

An insider reports on features that annoyed Cray EL customers:

1. There was considerable resentment amongst some of the early EL customer base that an expensive Sun workstation was provided with the system exclusively for the use of the hardware engineers when servicing the machine. This relic of big system thinking eventually led to a change of policy whereby the maintenance workstation became

- optional and system maintenance activity is run directly from the master IOP.
2. The Cray system install and configuration tool, a menu-driven program, could not in its early releases be used to configure the disk layout or adjust certain critical system configuration files.
 3. The first few UNICOS upgrades took upwards of ten hours of system downtime and could leave a system dead in the water if it failed. In later EL releases the upgrade process was changed to proceed whilst the systems were in production using a spare filesystem.
 4. The VME bus IOS boards could be error prone on some systems and difficult to diagnose and repair. The IOS buffer board located between the main processor and master IOP was a particular soft spot on some systems.

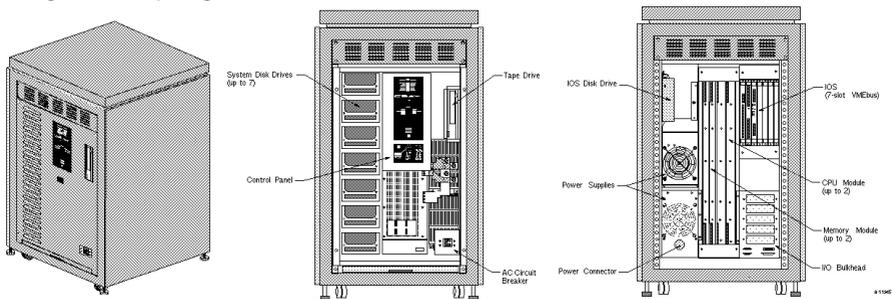
Evolution of the Cray EL into EL92, EL94 and EL98

Published in 1993, Cray Channels (a Cray marketing and technical magazine) announced upgrades for Cray EL customers. Modules could be replaced to upgrade systems into a Cray EL98 or new systems purchased with more processors and memory.

“Cray Research’s strong-selling Cray EL entry level supercomputing system, 130 orders booked in its first full year, got even better this spring. In March the company introduced the Cray EL98 system, with twice the peak performance of the Cray EL configurations at the same price and with four times the central memory capacity. Because the Cray EL98 system uses the same chassis as Cray Research’s previous entry level system, customers who already have a Cray EL system can easily add or exchange modules to upgrade their processing power. Cray Research currently has 11 customers upgrading to Cray EL98 systems; seven of these organizations are upgrading to fully configured eight processor systems. With up to eight processors and 4096 MegaBytes of central memory, the Cray EL98 supercomputer provides a peak performance of over 1 Giga Flops. To deliver the highest levels of sustained performance, the balanced architecture of the Cray EL98 system provides four Giga Bytes/s of total memory bandwidth in combination with one Giga Bytes/s of I/O bandwidth to storage” ... “Pricing for a two-processor Cray EL98 system begins at \$340,000 in the U.S”

The ex-Swansea University Cray EL on display at The National Museum of Computing was upgraded from an EL94 configuration to an EL98.

The Cray EL was repackaged into the least expensive Cray system ever. “With a U.S. starting price of \$125,000 and a footprint of about four square feet, the EL92, is limited to two processors and 512 Mega Bytes of memory.” The parts and arrangement looked similar to the EL but were not interchangeable. The EL92 later doubled up to the EL94. The power and cooling was suitable for an office environment but the system was just too big and noisy to go under a desk.



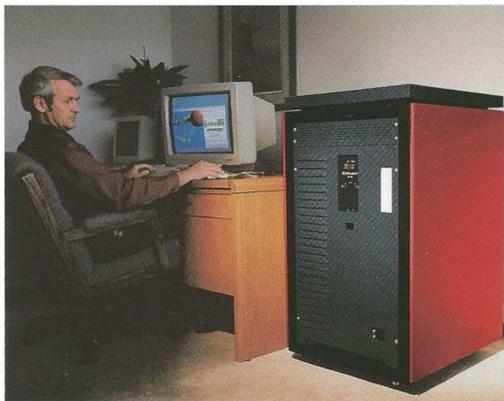
In conclusion

The Cray EL range provided an expanded customer base for Cray systems and complemented the larger Cray systems providing an identical software development and production environment. Adjusting the company from creating a handful of big complex systems each year to shipping two or three systems each week did cause big manufacturing changes in Chippewa Falls, Wisconsin. Never a mass market machine the Cray EL established a range of air-cooled departmental systems that brought supercomputing power to mainstream engineering, climate, academic and financial customers. With an eye to the future of bigger systems one 1993 press release noted “The EL92 systems can also be ordered with the CRAY T3D Emulator, a software tool that helps programmers using the CF77 programming environment to develop and test applications for the CRAY T3D massively parallel system.”

The Cray EL range was a sales success: “In 1992, the company’s first full year in the entry-level supercomputing market, Cray Research booked orders for 130 CRAY EL systems, exceeding its original target of 100 orders for the year, according to Bob Ewald, CEO of Cray Research. Cray Research’s entry-level systems have brought in more than 70 new-to-Cray Research customers in the aerospace, automotive, chemical, financial, construction, utilities and electronics industries, as well as universities and environmental and general research centres.”

Personally, in eight years of working at Cray Research, I was only ever a small chip in a big machine, but I always had a proud feeling like a car mechanic who works for Lamborghini. Dealing directly with high value customers allowed me to make a difference and taught me some valuable customer service skills. I was always learning, never knew enough and valued the really clever colleagues throughout the organisation that worked together to deliver exceptional computers into customer production environments.

By the end of the ‘90s the air cooled EL range evolved into the Cray J90 and, at the top of the computing power tree, the big vector T90 machines were giving way to the massively parallel Cray T3E systems but that’s another story. Search for “Cray FAQ Home page” for other Cray history notes and information. I would be very interested to hear your stories about the Cray machines you came across and what problems they solved for you.



Pictured left is a Cray EL92, the cheapest Cray Research system ever made.

Article by Clive England
Date: 31/01/2021
email: Clive@spikynorman.net

First Published in “Resurrection” Journal
of Computer Conservation Society Issue
93 - Spring 2021

CRAY
RESEARCH, INC.