

Cray supercomputer FAQ

From: newsposter@SpikyNorman.net (Gannett)
Newsgroups: [comp.unix.cray](#), [comp.answers](#), [news.answers](#), [comp.sys.super](#)
Subject: Cray supercomputer FAQ
Followup-To: poster
Date: Sun, 4 Mar 2001 12:02:31 +0000
Organization: UUNET WorldCom server (post doesn't reflect views of UUNET WorldCom)
Message-ID: <leppi26.aph46w6v0st0N@useraq60.uk.uudial.com>
X-Trace: lure.pipex.net 983708317 29996 62.188.136.120 (4 Mar 2001 12:18:37 GMT)
X-Complaints-To: abuse@uk.uu.net
NNTP-Posting-Date: 4 Mar 2001 12:18:37 GMT
X-Newsreader: MacSOUP 2.3

Archive-name: computer/system/cray/faq
Posting-Frequency: monthly
Last-modified: March 2001
Version: 1.06
URL: <http://www.SpikyNorman.net/>
Copyright: (c) 1999 "Fred Gannett"
Maintainer: Fred Gannett <CrayFaq@SpikyNorman.net>

Cray Research and Cray computers FAQ Part 3

Cray Research and Cray computers FAQ Part 3

- * What's in a name ?
- * What's in a number ?
- * Where did the first ones go ?
- * Who has/had the most Cray systems ?
- * What is a T94/SSS ?
- * What went overseas ?
- * Keeping it cool
- * What's a Mega word ?
- * What is an SSD ?
- * What is SEC-DED ?
- * Binary compatibility
- * How did users shape the design of Cray machines ?
- * Why was it hard to program Cray machines ?
- * What is boundary scan ?
- * How is a T3d different from a Beowoulf/NOW/PC cluster ?
- * What is dumping ?
- * How do you start a Cray system ?
- * Instructions for starting a Cray EL system
- * Chips off the same block
- * Could you choose the colour of your Cray machine ?
- * What was Ducky Day ?
- * What were the code names rain, gust, drizzle, cyclone etc ?
- * What was the physically smallest Cray machine ?
- * What was the physically largest Cray machine?
- * What was the computationally largest machine that Cray made ?
- * What was the Cray connection with Apple ?
- * How many people worked for Cray Research ?
- * Who ran Cray Research
- * Was a Cray supercomputer value for money ?
- * Did Cray fail? It was bought out by SGI in 1996
- * Trademark Disclaimer and copyright notice

This Cray supercomputer Faq is split into sections, Part 1 describes

Cray supercomputer families, Part 2 is titled "Tales from the crypto and other bar stories", part 3 is "FAQ kind of items", part 4 is titled "Buying a previously owned machine" and part 5 is "Cray machine specifications". Corrections, contributions and replacement paragraphs to CrayFaq@SpikyNorman.net Please see copyright and other notes at the end of each document. Note: Part 3 is the only part posted to newsgroups, the latest version of this and the rest of the documents can be located in http format just down from:

<http://www.SpikyNorman.net/>

What's in a name ?

Cray Research (CRI), the original founded Seymour Cray Jr. in 1972 bought by SGI in 1996. Now considered a separate business unit within SGI from 1999. Cray Computer Corporation, spun off from CRI to Colorado Springs. 1989 .. 1995 Cray Labs research arm of Cri in Boulder in the early 1980s. Cray Communications, Not related, changed name to Anite corp. Cray Electronics ditto. CraySoft, a venture to exploit the compiler technology and expertise of the application software development group. This lead to the interesting fact the each time a program was compiled on a T3e the compile script checked to see if the machines was running Solaris! SuperTek, original developer of the XMS, was bought in 1989 by CRI Floating point systems, original developer of the CS6400, folded in 198?, assets bought by CRI. Was sometimes known as Cray Research Superservers (CRS) SSI, Supercomputer systems Inc, founded by Steve Chen chief architect of the YMP

What's in a number ?

All CPU chassis, SSD chassis and in the larger machines IOS chassis had serial numbers but the CPU chassis number was taken as the main system designation. From the serial number it is possible to get a general indication of the model type as different ranges of machines were grouped into ranges of numbers. The structure of the model designation varied but often indicated the max. number of CPUs that could be installed, the chassis size, and sometimes the memory size.

Cray Research System Serial numbers

| | Machine type |
|------|-------------------------------|
| <100 | Cray 1 |
| 1nn | XMP 2 CPUs |
| 2nn | XMP 4 CPUs |
| 3nn | XMP 1 CPU |
| 4nn | XMP 2 CPUs |
| 5nn | XMP 1 CPU (SE) |
| 6nn | Cray/ELS XMS 1 CPU |
| 10nn | YMP 8 CPUs, Model D IOS |
| 11nn | XMP/EA 4 CPUs |
| 12nn | XMP/EA 2 CPUs |
| 13nn | XMP/EA 1 CPU (SE) |
| 14nn | YMP 2 CPUs, Model D IOS |
| 15nn | YMP 4 CPUs, Model D IOS |
| 16nn | YMP 2 CPUs, Model E IOS |
| 17nn | YMP 8 CPUs, Model E IOS YMP8I |
| 18nn | YMP 8 CPUs, Model E IOS YMP8E |
| 19nn | YMP 4 CPUs, Model E IOS |
| 20nn | Cray 2, 2/4 CPUs |
| 2101 | Cray 2, 8 CPUs |
| 24nn | Cray YMP/M94 4 CPU |
| 26nn | Cray YMP/M92 2 CPU |
| 28nn | Cray YMP/M98 8 CPU |
| 3nnn | SV-1 |
| 40nn | C90 16 CPUs |
| 42nn | C92A 2 CPU |
| 43nn | D92A 2 CPU |
| 44nn | C94A 4 CPU |

| | | |
|------|------|---------------------------|
| 46nn | C94 | 4 CPU |
| 47nn | D94 | 4 CPU |
| 48nn | C98 | 8 CPU |
| 49nn | D98 | 8 CPU |
| 5nnn | ELs | |
| 60nn | T3d | |
| 61nn | T3d | Included in a YMP2E + IOS |
| 62nn | T3d | |
| 63nn | T3E | 300 - liquid cooled |
| 65nn | T3E | 600 - air cooled |
| 66nn | T3E | 600 - air cooled |
| 67nn | T3E | 900 - liquid cooled |
| 68nn | T3E | 900 - liquid cooled |
| 69nn | T3E | 1200 - liquid cooled |
| 70nn | T94 | 4 CPU |
| 71nn | T916 | 16 CPU |
| 72nn | T932 | 32 CPU |
| 9nnn | J90 | |
| 95nn | J90 | 32 CPUs |

Notes:

The < 100 series can be further subdivided, among the Cray-1, Cray-1A, Cray-1S, and Cray-1M lines. SN101 was the first XMP with 2 cpus

Eugene R. Somdahl posted in c.u.c The 1Ms were 1Ss with MOS memory. I believe only six were built. Plus one that wound up as "scrap" running CTSS for several years in a basement in Chippewa Falls. This was the machine on which much of the Y-MP (and later Chen's MP project) design work was done.

XMP/EA: systems had 500 series serial numbers. Production ran from sn501 (19 Jul 87) through SN515 (15 Jan 89).

Cray-2: Later, a small number of variants occurred: 2 and 8 processors, 128 MW and 512 MW, Dynamic RAM (DRAM) and Static RAM (SRAM) don't know what SNs Only Serial numbers Q1,Q2 and 2001..2029 built.

SN1001 Had slower clock speed than later YMP 8s.

SN1040: Originally a Model D IO system it was converted to Model E IOS for use as a T3d host then later went to Moscow.

Cray-2 SN2000 (only a single CPU) and SN2101 (only 8 CPU) now reside in the computer museum at Moffett field CA. There were 4 single cpu Cray2 systems built, Q1 -> Q4. Only one was shipped to a customer, one was in Eagan and 2 were used as test systems in Chippewa.

The 1700 series was known as a Y-MP8I because it could hold 8 CPU's with 1 SSD Section and 4 IOSE clusters in a single integrated chassis. One 1700 chassis was rewired to hold three 1600 type machines. That wiremat wasn't pretty. Served in CCN as software development machines ICE, FROST and SUBZERO.

The 1800 series known as the Y-MP8E with similar module layout to a 1000 series but could hook up to a Model E IOS. It would typically hook up to a 700 series Model E IOS which had a chassis very similar to an 1800 that holds 4 SSD sections and 8 IOSE clusters.

IOSE/SSDE boxes usually came in three serial number ranges 7xx, 8xx and 9xx. The 8xx is a short chassis like the 1600 that held 1 SSDE section and 2 or 3 I/O clusters. The 900 chassis looked like a 700 chassis but was only wired for SSDE.

A couple of these machine types (1600, 1900 and 6100) could support a JSSD (SSDE32i/SSDE128i). That was a 32 or 128 Mword SSD that fit in a single chassis slot.

The T3D 6100 series integrated 128 EV4's 2 Y-MP CPU's and 4 IOSE clusters into 1 chassis, and a 6200 which I believe is like a 6100 without the Y-MP/IOSE integration.

SV-1 may comprise of a cluster of systems resulting in the use of more than one serial numbers.

SV-1 SN3001 to SN3007 and SN3213 to 3220 ... note that SN3217-3220 comprise SN3501 (super-cluster) at NIH.

In 1978 5 Cray-1 systems were installed. In 1996 350 Cray J90 systems where shipped the large part of the total of 415 J90 systems. Some J90 systems are being converted to SV1 chasis just to keep the records complicated.

If you know sub ranges of serial numbers or special variants from the above pattern please email the document author. Details of C3, C4 and CRS6400 numbers would also be appreciated.

The small serial number ranges indicate that the phrase "Hand built in Chippewa Falls" was never far from the truth. With SGI winding down development of the big Cray machines will we ever see Cray SN10,000?

Where did the first ones go ?

The early adopters of a technology are the crucial customers that can make or break a developing technology company so it is interesting to look where the first few Cray-1 machines were installed. In 1978 5 Cray-1 systems were installed.

The first Cray one SN1 went to LASL in 04/76 for a six month evaluation after which a short-term contract was signed. In 09/77 SN1 was replaced with SN4, a 1000k word machine with SECDED memory protection hardware. SN1 Ran an OS called DEMOS written in MODEL.

SN2 was gutted and never shipped.

Looking at the Cray 1 installed base in the 1976 to 1979 time frame.

Date, Place, application

04/76, LASL Los Alamos NM, Nuclear research later replaced by SN3

07/77, NCAR CO, Atmospheric research

10/77, ECMWF UK, Weather forecasting initially SN1 later replaced in 10/78

01/78, US DOD, Defence research

07/78, US DOD, Defence research

04/78, NNFEC National magnetic fusion energy centre Livermore CA, magnetic fusion research

09/78, United computing systems Kansas City MO, Commercial computing services

10/78, MOD UK, Defence research Used SN1 replaced 04/79

01/79, Lawrence Livermore Lab Ca, Nuclear research

04/79, Cray Research Inc., Development and bench marking

The machines above were 500k or 1000k Words. These machines were so far ahead of their time that many of them worked at a number of sites, sometimes being leased for short periods until the customer-ordered machine was ready.

Most non-government Cray-1 systems were front-ended by CDC CYBER 170 systems, IBM MVS or VM machines or DEC VAX systems.

Apart from the first Cray-1, which travelled the world, the first machine of most product types typically served in the Cray computer centre (CCN) first at Mendota Heights and later at 655 Lone Oak Road, Eagan.

The first (prototype) Y-MP/8D, sn1001, served in the Cray computer centre as the compute platform "Mist" (later changed name to "Gust" as "Mist" is a rude word in German).

The first (prototype) X-MP, sn101, remained on the floor of the 1440

building until 1989 (17 Aug 82 - 25 May 89). ... It was subsequently scrapped.

The first Y-MP2E, sn1601, travelled to the UK and served in the Cray UK data centre as a software development machine.

Pittsburg super computer centre PSC had the first customer T3D and T3E.

Cray 2 customers included NMFECC, NASA Ames, University of Minnesota, US DOD, Harwell Laboratory UK, Aramco, University of Stuttgart in Germany and Ecole Polytechnique (Paris France).

Quoted from From HPC wire Subject: 9037 CASH-STARVED CRAY COMPUTER CLOSES, SEEKS CHAPTER 11 Mar. 27 CCC did make one tentative sale, to Lawrence Livermore National Laboratory (LLNL) for support of the Department of Energy's nationwide research program. In December 1991 when CCC was unable to meet delivery/performance goals, the order was cancelled.

A small (4-CPU) CRAY-3 was later placed at the National Center for Atmospheric Research (NCAR) where, after some time and effort, it became ready for production work. This was not followed, however, by an actual sale.

Who has/had the most Cray systems?

According to the SuperSites list in Aug 1999 The NSA runs a lot of Cray machines.

| | | |
|---|------------------------|--------|
| National Security Agency, Fort Meade, Maryland,US | | |
| 1) | Cray T3E-1200 LC1084 | 1300.8 |
| 2) | Cray T3E-900 LC1328 | 1195.2 |
| 3) | Cray SV1-18/576 | 576 |
| 4) | SGI Origin2000/250-864 | 432 |
| 5) | Cray T3E-1200 LC284 | 340.8 |
| 6) | 3 * Cray T932/321024 | 174 |
| 7) | Cray T3E-750 LC220 | 165 |
| 8) | Cray T94/SSS-256K | 100 |
| ... snip | | |
| 20) | 4 * Cray C916/161024 | 64 |

I can't possibly comment but the only other place where so many Cray machines have been co-located is in the 655-D machine room of CCN located in the Cray Eagan building (now Wham!net). The population of that centre changed over time but will have had just about every sort of Cray Research machine at one time or another. As the CCN data centre was used to house and manage customer machines there was often more than one of each current production system.

What is a T94/SSS ?

The SSS designation on a T90 indicates a special configuration that was developed for particular customer. Its actual "speciality" may be seen in a section of an ARPA document on the net. The document mentions a processor in memory development for the Cray-3 for use in " This machine would be suited for parallel applications that require a small amount of memory per process. Examples of such applications are laminar flow over a wing, weather modeling, image processing, and other large scale simulations or bit-processing applications." It is thought (by me anyway) that this development was updated into T90 technology.

What went overseas ?

To the eastern block.

- * An EL went to Prague.
- * 2 ELs and later a YMP-4E and two J90s went to Poland.
- * An EL and later a YMP-8E went to Moscow.
- * A C90 went to China.

* An XMP went to India for weather forecasting.

All with the agreement of the export control authorities.

The first mainframe class Cray machine was installed in Moscow the same week that the Moscow branch of "Planet Hollywood" opened.

International subsidiaries as of December 31, 1995. It could be reasonably assumed that the offices were for the supply and support of Cray Research computers in the host country. Not in the list is the UAE and the large Aramco site that had a Cray-2 followed by a C98-DRAM machine.

International offices

| | |
|---------------------------------------|-----------------|
| Cray Research A.B. | Sweden |
| Cray Research Scandinavia A/S | Norway |
| Cray Research (Australia) Pty. Ltd. | Australia |
| Cray Research B.V. | The Netherlands |
| Cray Research (Canada) Inc. | Canada |
| Cray Research Europe Ltd. | United Kingdom |
| Cray Research France S.A. | France |
| Cray Research GmbH | Germany |
| Cray Research Japan, Ltd. | Japan |
| Cray Research (Korea) Ltd. | Korea |
| Cray Research (Malaysia) Sdn. Bhd. | Malaysia |
| Cray Research de Mexico, S.A. de C.V. | Mexico |
| Cray Research OY | Finland |
| Cray Research, S.A.E. | Spain |
| Cray Research S.R.L. | Italy |
| Cray Research (Suisse) S.A. | Switzerland |
| Cray Research (UK) Ltd. | United Kingdom |

Keeping it cool

The development of Cray cooling technology allowed each technology generation to increase the circuit board density. "Someone (perhaps Gary Smaby? I truly don't remember) once said that Cray Research was primarily a refrigerator company."

Cray-1: Single sided boards clamped to copper plates placed in aluminium racks that had cooling fluid in tubes.

XMP: Double side sandwich boards clamped to twin copper plates placed in aluminium racks which had cooling fluid in tubes.

Cray-2,3,4: Immersion cooling. The CPU and memory boards sat in a bath of electrically inert cooling fluid.

YMP, C90, T3d LC, T3e MC: Double-sided circuit boards clamped to hollow aluminium boards in which the cooling fluid circulated.

E1,J90,T3eAC,SV-1: Blown air cooling.

T90: Immersion cooling. The CPU and memory boards sat in a bath of electrically inert cooling fluid.

So the main forms of cooling were conduction to external cooling, conduction to internal coolant, blown air cooling and total immersion cooling. There was some research done into sprayed coolant methods for J90 modules, reported in an internal technical symposium paper, but this did not make it into a publicly available product.

What's a Mega word ?

All Cray machines (except the CS6400 range) had 64 bit words. Each word could hold 1 integer, 1 floating point numeric value or 8 characters. All arithmetic and logical operations were done in 64 bit maths. Memory sizes were generally designated in Mega Words. 32 MWd can be thought of as 256 Megabytes. However really the memory was 72 or 80 bits wide at the hardware level See SECDED.

What is an SSD

Bigs gobs of memory that attached via very high speed channels to vector CPUs. These extra memories, often 4 to 16 times the size of central memory could be utilised as disk IO cache, swap space, extra memory segments for programs or even RAM based file systems. These memories were accessed on a YMP at up to 1000 Mbytes/sec and are often used to transparently hide the IO required for an out-of-memory solution. Sites that used SSD as root/usr file system disk cache often saw little of no physical disk activity even when the system was stressed. A trickle sync mechanism was employed to prevent these vast disk caches from becoming stale.

Some later SSDs were built from J90 technology memory boards and some from T3E boards.

What is SECCED ?

Single bit Error Correction, Double bit Error Detection. This was a hardware scheme used in all Cray machine (except the very first C1 and CRS systems) to allow single bit memory errors to be effectively ignored and for detecting other memory errors to prevent memory data corruption. Each 64 bit word had 8 other bits that were used in this advanced parity checking/correction method. If a double bit error was detected a syndrome byte indicated where the error had occurred. SECCED on the C90 has 64+16 bits and allowed correction at the 4-bit level.

Binary compatibility

Binary compatibility, the ability to run programs compiled on one machine on another, extended back one generation only. You could run XMP binaries on a YMP but not on a C90. The J90 counted like a YMP.

There were some cross compilation libraries available so that you could make binaries for different architectures i.e. it was possible to build YMP native codes on an XMP. Cray always recommended recompilation on the newer system in the case of an upgrade in order to take advantage of the new hardware features.

Execution compatibility across the range looks like :

```

C1 --> XMP --> XMPema --> YMP --> C90 --> T90
      |           |           ^
      V           V           |
      XMS --> ELs --> J90 --> J90se --> SV1
                *
                *
                T3d ==> T3e

APP --> CS64 --> CS64000
C2 .   C3 .   C4

```

For binary compatibility you can go across one arrow or up/down one arrow only.

Note ==> the T3e used the same programming models as the T3d and so was source code compatible with only minor exceptions.

Note ** An MPP emulator was available for ELs to develop MPP codes while waiting for the T3 product line to arrive. The emulator managed about 4 MPP nodes but used the native CRI arithmetic.

Note The C90 could run J90 binaries if compiled without the scalar cache optimisation. The T90 comes with either Cray Floating Point CPUs or IEEE CPUs (or both). T90 IEEE arithmetic CPUs will not support C90 native codes.

The C4 had IEEE FP and went back to lots of registers instead of local memory. There was no binary compatibility between the C2, C3 and C4.

How did users shape the design of Cray machines ?

System owners asked for and got machine instructions to do "population counts", leading zero and later a bit matrix multiply functional unit. BMM was provided as an option on C90 and standard on all T90 CPUs.

Hardware based gather/scatter instructions were included in the EL and later systems. This instruction allowed the compilers to manage array indexed data more efficiently.

See also What is a T94/SSS ?

Why was it hard to program Cray machines ?

There has always been lots of Unix src code sloshing round on the net since well before Linux and the admirable open src code movement was invented. However there were a couple of things about Cray machines that made porting codes to Cray machines tricky. We won't get into what you had to do to your algorithm to get the best out of a Cray machine but just examine a few things that made the conversion of codes to Cray machines interesting.

Firstly there was the word size, one rather large size fitted all, integers and floats were represented in 64 bits, whereas most other machines used 16 or 32 bit for ints and floats. This would not cause a problem in well written codes unless assumption were made about the range of an integer or the relative sizes of integers to character data types.

Cray machines are word addressable, not byte addressable. Although the compilers used transparent word division to mimic byte addressable constructs, advanced character handling was never a natural process on a Cray machine.

When the original arithmetic units were designed for Cray PVP systems, the IEEE floating point standard for arithmetic overflow and underflow did not exist. The IEEE standard compromises performance in favour of ease of use and so may not have made it in anyway. The Cray floating point format provided a different range of numbers and level of accuracy that tripped up some programs.

There are also other subtle considerations that have caused headaches for many programmers porting codes to Cray machines. Known collectively as the "Cray effect", they are the combination of algorithm scaling problems, cyclic accumulation of errors and parallelism interdependencies that seem to show up most times you take an apparently well behaved small program and run it longer harder and further than possible on a conventional system.

A source adds :

Possibly the most subtle problem in porting programs to UNICOS was that some very "bright" person put a #define for "WORD" in a header file required by the kernel. The first thing to do in porting a program to UNICOS was to find if it used WORD as a #define or as a typedef and work around that.

Cray machines could not support "alloca", so minor magic had to be applied to programs using "alloca."

In the very early days, many C programs suffered from the "nUxi" problem, but that was hardly unique to Cray machines.

At various times, the following languages were available on Cray machines, FORTRAN 66,77,90,HPF, Cray C, Standard C, C++, Pascal, Ada, Perl, but never GNU C or Basic. There was an implementation of APL for the CRAY-1, and an implementation of SNOBOL 4. Neither was ever "officially supported." LANL is reoprted to have ported Common LISP.

What is boundary scan ?

This is a low level logic feature that is used to initialise and test the logic state of a board in a machine. Originally introduced on the EL it was later adopted and used on the J90 and T90 machines.

Scan files are unique to each CPU/memory board type. On the early EL systems boundary scan tapes with pre-built scan initialisation files for various part numbered CPUs were available but had to be updated whenever a later revision of board was used.

As the number of CPU module types mushroomed, Cray engineering realised the service problems associated with scan tapes. Releasing the scb command and the associated ASIC data files structures allowed service personnel to rebuild scan files of a machine in the field.

If you know that your hardware is sound you can use scb command to regenerate boundary scan files, as long as you have the /install/asic files and scb command.

The bscan command is used to read and write the scan state off/onto a board to progress hardware problems.

How is a T3d different from a Beowulf/NOW/PC cluster ?

This section could also have the title "The return of the Killer (network of) micros." The battles between the proponents of the "network of workstations" and the "its only a supercomputer if it costs more than 1M\$ per CPU" crowd rattle endlessly round the halls of [comp.sys.super](#) but we are not going to discuss the merits or otherwise of MPPs V. Vector supers. This section is purely to describe some of the differences between modern implementations of MPP technology. The term cluster can be considered to refer to any collection of generic parts pushed together to make a compute engine. e.g. Beowulf, NOW, SP2. An understanding of the applications that you will be using in an MPP system is vital when deciding what performance attributes are important to you.

There are four main areas of differences between a cluster and a T3e, interconnect bandwidth (and low latency), single system image, programming models, and IO ability. Some of these differences may erode, as cluster technology develops, but for now these differences stand to justify the extra 0s on the price.

Interconnect bandwidth and low latency: To make effective use of an MPP computer or cluster the problem/program has to split the work between multiple CPUs. Inevitably the whole task will require some communication between the nodes working on the different parts of the answer. Sometimes this co-ordination is only a very small part of the process, with just a distribution of sections of the problem, at the start, and the collection of the answers at the end, but with other types of problem close synchronisation and large amounts of inter processor data passing is required.

This is where understanding the speed and latency between the processing nodes becomes important. Inter processor synchronisation speed depends on network latency. Inter processor data passing speed depends on network bandwidth. Total communication bandwidth depends on the interconnection topology and sophistication of the routing algorithm.

The amount of communication required by a cartoon film render farm, for instance, would be at the low end of the communication requirements scale whereas a large finite element or fluid dynamic problem would be at the top end requiring low latency interconnects and high bandwidth. Combining these factors is the ratio of communication time to processing time. Delays in synchronisation and bandwidth become less significant if the calculation time between communication is long.

See The Gannett's law document for a description of a possible method of assessing problems and cluster computer solutions.

Single system image: The T3e has a single system image, this is not to be confused by the fact that the operating system is distributed across an number of CPUS. One of the main benefits is that there is just one OS to upgrade, monitor, boot and tune. Another even more important benefits is the availability of global memory data segments, closer synchronisation between the processing elements and shared filesystem access. The downside of a single system image is a high degree of inter node dependence that can make hardware problems tricky to find and fix. The close PE co-ordination makes command and task load balancing fast and efficient.

Programming models: The hardware support of a barrier tree mechanism allows fast multi-node synchronisation. The eureka hardware support allows problems with wide solution branching to be co-ordinated on "first to find the results" basis. Hardware support for atomic swap (and block memory move) allows for fast and accurate semaphore operation between shared memory locations. Clusters support the message passing programming models but also seen in the T3e are the shared memory programming models popular from symmetric multiprocessor systems.

Processor interconnect: The high level of processor and network integration allows the T3e to perform multi-node process moves, checkpointing and program debugging. Problems that require 1000s of process synchronisation per short problem step can only be realistically considered on the T3e.

IO ability: Clusters can work well in the cases where the IO is limited or centralised but many super computer applications are defined by large amounts of computation and huge amounts of data. The T3es provision of large shared file systems that makes sharing of huge datasets between computational processors easy. Support for huge raid disk arrays equally available across all CPUs in the system make shared file access routine. NFS cross mounting between 100s of processing nodes just doesn't provide the strength for many big problems. The emergence of multinode access filesystems on the SP2 have served to make IO a closer race.

So to sum up a T3e is best suited to problems whos parts require very high levels of synchronisation, co-ordination and communication. If the requirement is for chunks of computer power loosely coordinated over long time steps and budget is a big consideration generic cluster technology should be considered.

What is dumping ?

Copying memory contents to disk, copying disk contents to tape or selling computers at less than cost to stuff the opposition. CRI filed a governemnt investigated dumping complaint against NEC over the supply of system to NCAR.

How do you start a large Cray system ?

This is the short version, the full version fills a very long chapter and a whole day on the "Unicos kernel internals" course. The exact details varied a bit from generation to generation but the principals remained the same. Starting from a cold machine, the first thing to do would be to call the Cray engineer to start the power conditioning MG set. Next the check the cooling circuits and WACS panel on the side of the main cabinet, to see that the power supply rails and temperature levels are normal. Working from the MWS the engineer would then master clear the machine to initialise the processor and IO channel logic. After this the first code would be copied main memory and the DEAD START command issued. The first code would usually be diagnostics to set and check the internal logic of the CPUs and memory. The diagnostic stage was largely replaced by the boundary scan logic in later systems.

Finally the system would be handed over to the operators for the booting of

Unicos. The Unicos boot process went like this. Firstly all the CPUs are halted and, starting from word 0, the kernel is copied into main memory from the OWS or support disk, along with a parameter file describing the hardware environment, kernel parameters, disk and IOS layout. At this point CPU 0 would deadstart and run while the rest of the CPUs idled. The kernel parsed the parameter file and used it to locate the IOSs and root partition, both of which it would need for single user mode. Once in single user mode the rest of boot sequence (triggered by init 2) was much like most Unixs, spawn the service demons, initialise the networks, check and mount the rest of the file systems, spawn the console gettys.

There were many an interesting reasons why systems would not boot but most of them came down to hardware problems, incorrect parameter files and Unicos bugs on new hardware. Finding what was causing a system boot failure was of course a source of endless amusement.

Instructions for starting a Cray EL system

Check the "big red button" is up. Switch system power on at the power breaker on rear machine wait until system ready indicator lights up on front panel of the machine. Press the system reset buttons under the flap on the front of the machine. Let the console messages and diagnostics scroll up.

It should come to the BOOT> prompt. Type "load" this will scan the hardware looking for disks and controllers according to the config.sys file on the IOS disk. It will also run diagnostics if this is a cold start.

At this point you should have a IOS> prompt. You can now cat /bin/boot Look for the line that has something like
lu /sys/sys.ymp /sys/config.uni

/sys/sys.ymp is the kernel the other is the text parameter file which describes the disks/filesystems that Unicos will use. Then type "boot" to start Unicos into single user mode. The system should now arrive at the # Unicos prompt. Check your file systems with /etc/gencat. When this is complete go to multi user mode with init 2

Typing ^a on the console switched you between talking to the IOS and talking to the Unicos console unless either the scroll lock (right next to the delete key) had been pressed or the console line had been switched into remote support mode in which case the machine just looks hung.

A computer based learning training course for system operators was constructed that simulated the boot sequence and other basic EL maintenance procedures. See the annotated boot sequence in a near by document.

Chips off the same block

What was the difference between the Alpha chips used in the T3[de] and other Alpha chips ? Basically not much except in the liquid cooled T3[de] versions where the chip was packaged upside down. Normally the heat from the Alpha chip is dissipated through an air heat sink on the top of the chip but with the T3e providing cooling from inside the board the packaging was flipped to have the heat go down instead of up. DEC also provided a pin to allow CRI to run data fetches big-endian rather than little-endian.

```
Heat| ++++++ chips
      V ===== PCB board
      ----- Cooling fluid in hollow copper module
      ^ ===== PCB board
Heat| ++++++ Chips
```

There was some consternation in 1997 when Intel bought the Alpha chip fabrication plant, the prospect of putting "Intel Inside" on the case of the Cray T3e filled purists with horror.

What was Ducky day ?

This was an employee fun day that happened once a year at the Eagan buildings. The day was a holiday for all staff with organised sport and social events. The Tee shirt design of the day was decided by a cartoon competition. The origins "Ducky day" name came from a prank; a plastic duck was found in the rather grand water pond with sculpture outside the 1440 building. The facilities manager issued a rather stern warning so the next day the pond was filled to the brim with yellow plastic ducks.

What were the code names rain, gust, drizzle, cyclone etc ?

Up to about 1994 the computers in the Cray Research, Eagan and Chippewa Falls, machines rooms in were known by their serial numbers. For manufacturing this was fine but for the Eagan centre it was a pain as computers were replaced and upgraded on a reasonably regular basis. CCN in Eagan was more interested in running compute environments so it was decided that each type compute platform benchmarking, filestore, batch engine etc. would have an environment name instead. This meant that "Rain" as a compute platform started out as a YMP8e but later transparently changed into a C90 with less disruption to the end users. The exception to this was USS the Data Migration file server that was always called USS. The some of the biggest machines in Eagan were interconnected by very high speed networking, HYPER channel in the early days, and later (@1994) 200MB/s HIPPI connections.

Could you choose the colour of your Cray machine ?

In the early days of the company yes, there was even rumour of a cow hide covered XMP delivered to a Houston oil company. As time went on this was dropped as a customer option. Well almost - when there is that much money changing hands, if enough fuss is made, the exterior panels would revisit the paint shop. This did not apply to ELs which were all black and red. Well almost - one customer which had just upgraded a pair of YMPs (one green, one blue) for a C90 and an EL did manage to get the EL painted a rather fetching sky blue colour.

As for XMP/EA, sn501 and internal contact reports "We lobbied to gave it done up in denim (like the denim Jeans) & have a little red Levis tag attached. ... Management was not amused & it never happened."

One second user customer did have a bit of a surprise when their second user C90 arrived in a lurid deep rose/pink colour. The top of the C90, being a convex shape happened to sit just a couple of feet under a set of strip lights and resulted in lovely pink glow over a whole section of the machine room.

The Bell Labs Cray (XMP) was a wallpaper'ed IC design.

What was the physically smallest Cray machine ?

The smallest, commercially available machine, was the EL92 a repackaged version of the air cooled EL range which measured approx. 1.2m high by 0.6m wide by 0.6m deep and could run from a normal power outlet. Whilst not a big commercial success as it was a bit late to market, it was widely used for trade shows, software development and as loan equipment. Available in 2 and 4 CPU versions with 512Mb memory it was truly a deskside Cray. Good write up and pictorial in "Advanced systems" magazine July 1994.

What was the physically largest Cray machine?

The 16 CPU version of the C90 was a truly big machine standing at 2.5m tall and CPU cab just fitting in a 4m diameter circle. Together with the power and cooling equipment the whole system weighed in at approx. 12 (?) tonnes. There was one (or more?) systems delivered that consisted of four interconnected 16 CPU C90s.

What was the computationally largest machine that Cray made ?

In December 1998 the biggest machine was a 1048 CPU T3e assembled by Cray and released to a small community of scientists in 1998. This machine was subsequently delivered to an undisclosed customer.

The "Guinness World records 2000" has the Cray C90/16 as the fastest general purpose vector-parallel computer but we all know this should be updated to the T90/32.

What was the Cray connection with Apple ?

Cray and Apple, seemly at opposite ends of the computer spectrum, do have some subtle links. It was known that Seymour Cray used an Apple desktop some of the time when designing the Cray-2. It is also known that Apple had a sequence of Cray machines starting in March 1986 with an XMP/28 followed by another XMP in Feb 1991. A YMP-2E arrived later in 1991 and finally an EL from Dec 1993 to Jun 98. It is said that Apples first XMP was bought by Steve Jobs after he just walked into the Cray facility in Mendota Hights.

Originally purchased to help out on a computer on a chip project, the machines eventually earned their keep running MOLDFLOW an injection plastic modelling program (producing some results in the form of Quicktime movies) and later as a file server. Other applications were CFD codes for disk drive design improvement and one source reports "... they sometimes ran the first XMP as a single user MacOS emulator ... They had a frame buffer and a mouse hooked up to the IOP."

What is less known however is that the small active display panel on the T3d was an Apple powerbook. The powerbook ran a Macromedia presentation showing the T3e cube of cubes logo with an orbiting growing/shrinking sphere. The display at one site was changed to alternate with a presentation plaque display. It was rumoured that one site engineer ordered a collection of spare bits that, over time, comprised a complete new powerbook.

The recent (Sept 1999) launch on the www.Apple.com web site of the G4 Macintosh computers displayed a YMP-8D computer on the processor page. Whilst there was no direct reference to that particular machine the was a requote of the Seymour quote about "using an Apple to simulate the Cray-3" in a sidebar. (prob this should be Cray-2 ED). The G4 is being touted as a "Gigaflop (1 CPU) it is certainly up to at least 1992 supercomputer cpu speed. The YMP pictured on the site would have had 0.333 Gflops per cpu but was sold as sustaining 1 Gflop, for the whole machine, on real life applications. It remains to be seen if the G4 can match the memory size, memory bandwidth and IO capacity of this 8 year old Cray. Supercomputers these days do a Teraflop. There is however no doubt that it will be cheaper to buy.

The popular Macintosh telnet program developed by has an icon which is an XMP surrounded by a network with Macs. NCSA had a Cray accessed by Macs and thus needed to develop such a program.

How many people worked for Cray Research

This compiled from various SEC 10-K documents.

As of December 31, 1995, the Company had 4,225 full-time employees; 83 in development and engineering, 1,399 in manufacturing, 763 in marketing and sales, 915 in field service and 165 in general management and administrative positions.

As of December 31, 1994, the Company had 4,840 full-time employees; 1,172 in development and engineering, 1,531 in manufacturing, 849 in marketing and sales, 1,075 in field service and 213 in general management and administrative positions.

As of December 31, 1993, the Company had 4,960 full-time employees: 1082 in

development and engineering, 1798 in manufacturing, 728 in marketing and sales, 1157 in field service and 195 in general management and administrative positions.

Cray Research Growth Table

| Year | Employees | Locations | USA/International | New Systems | Installed systems |
|------|-----------|----------------------------|-------------------|-------------|-------------------|
| 1972 | 12 | | 1 | 0 | 0 |
| 1973 | 24 | | 2 | 0 | 0 |
| 1974 | 29 | | 2 | 0 | 0 |
| 1975 | 42 | | 2 | 0 | 0 |
| 1976 | 24 | | 4 | 1 | 1 |
| 1977 | 199 | | 6 / 1 | 3 | 4 |
| 1978 | 321 | | 8 / 1 | 5 | 8 |
| 1979 | 524 | | 10 / 3 | 6 | 13 |
| 1980 | 761 | | 13 / 4 | 9 | 22 |
| 1981 | 1079 | | 16 / 4 | 13 | 35 |
| 1982 | 1352 | | 17 / 4 | 15 | 50 |
| 1983 | 1551 | | 18 / 6 | 16 | 65 |
| 1984 | 2203 | | 20 / 8 | 23 | 84 |
| 1985 | 3180 | | 22 / 10 | 28 | 108 |
| 1986 | 3999 | | 28 / 12 | 35 | 138 |
| 1987 | 4308 | | 34 / 13 | 43 | 113 |
| 1988 | 5237 | | 35 / 18 | 52 | 220 |
| 1989 | 4708 | | 32 / 18 | 57 | 240 |
| 1990 | 4857 | | 34 / 18 | 52 | 263 |
| 1991 | 5395 | | 35 / 21 | 71 | 309 |
| 1992 | 4895 | | | | 446 |
| 1993 | 4960 | 6 + 153600 ft ² | Leased | | 505 |
| 1994 | 4840 | 6 + 190600 ft ² | Leased | | 638 |
| 1995 | 4225 | 6 + 211100 ft ² | Leased | | |

Data from Cray Research via www.cbi.umn.edu.

CCC had 350 workers at the time of its closure in 1995.

Who ran Cray Research ?

CEOs of Cray Research

| Who | Dates | Where are they now ? |
|-------------------|----------------|---|
| John Rollwagen | 1976..1992 | Now is a "Venture Partner" at Paul Venture Capital in St. Paul, Minnesota |
| Marcello Gummucio | 1992..1994 | Now at BeOS ? |
| John F. Carlson | Jan 1993..1994 | |
| Robert Ewald | Dec 1994..1996 | Now at Estamp an internet venture selling snailmail postage over the web. |
| J. Phillip Samper | May 1995.. ? | ? |
| Irene Qualters | ? | |

Mr. Ewald ... joined the Company (SGI) in 1996 as President of Cray Research and a Senior Vice President of the Company. Mr. Ewald was the President and Chief Operating Officer of Cray Research, Inc. from 1994 until 1996, when Cray was acquired by the Company. Prior to that, he was Chief Operating Officer of Cray's Supercomputer Operations and in 1993 served as Executive Vice President and General Manager, Supercomputer Operations. From 1991 through 1993, Mr. Ewald was Cray's Executive Vice President, Development. His daughter worked for part of the Craysoft organisation.

Les Davis can certainly be considered one of the spritual leaders of the traditional Crayons.

Was a Cray supercomputer value for money ?

The utilisation of most of the 16 cpu C90 systems I ever saw approached 98% at most sites, month after month, year after year so despite the high, typically USD 30,000,000 price customers got value for money.

Warning salesmen maths ahead, full of apples=oranges assumptions but go with me for while...

Lets work it out over 6 years the typical life of a C90. USD 30,000,000 purchase price + 3,000,000 * 6 Per year service cost = USD 54,000,000 total. That would be 9,000,000 per year over 365 days = USD 24,657 per day for the system. Per CPU that would be approx. USD 1,541 per day. In that day on that one CPU you could do 1 Gigaflop/s * 24 * 3600 * 98 % = 84,672,000,000,000 Flops in the day.

C90 Cost per 10^6 flops = 154,100 / 84,672,000 = 0.0018 cents per MFlops

Think that a workstation in 1994 would cost you approx. USD 8,000 + 800 PA and would be obsolete in 3 years. So that is USD 10,400 total or 4366.6 PA. Per day that's USD 12. Typically workstations are person driven, and people work 9 to 5 - 1 hours in a day, 5 days out of 7 that's a 40/168 = 23 % utilisation. Workstations in 1994 could manage about 5 Mflops/s on real codes. So the workload works out at 5,000,000 * 24 * 3600 * 23% = 99,360,000,000 Flops in a day.

PC Cost per 10^6 flops = 1200 / 99,360 = 0.012 cents per MFlop.

So it is cheaper to work things out on a C90 than a workstation just as long as you have enough work to keep it busy. We won't go into the fact that you can run much bigger problems on the C90 or that you can get the same work done faster as we seem to have won this point already.

Did Cray fail? It was bought out by SGI in 1996

No. Cray dominated, after inventing the super computing marketplace, for 20 years. Cray was *the* name in high performance scientific and engineering computing. Cray did seem to lose the technology focus in the later part of the 90s, the T90 was too expensive to build, the T3e cost too much to develop and the DEC deal scuppered the chances of world domination in departmental compute servers but there is no doubt that the sheer quantity of engineering and science completed on Cray hardware changed the face of the 20th century.

So did Cray Research fail? No. Whilst relatively few machines were ever made, even of the more popular models, they provided an unrivalled platform for tackling the toughest computational problems, which is of course, the mission of Cray Research.

Late breaking news As of March 2000 the Cray Research name and business has been sold to Tera, the inovative supercomputer maker from Seattle. We await further developments.

Trademark Disclaimer and copyright notice

Thank you for taking time to read these important notes.

Cray, Cray Research, CRI, XMP, YMP, C90, T90, J90, T3d, T3e, Unicos, plus other words and logos used in this document are trademarks which belong to Silicon Graphics Inc. and others. There is nothing generic about a Cray supercomputer.

Some of the ideas described in this document are subject to patent law or are registered inventions. Description here does not place these ideas and techniques in the public domain.

I wrote this document with input from a number of sources, but I get both the credit and the blame for its content. I am happy to read your polite correction notes and may even integrate them with the text so please indicate if you require an acknowledgement.

Personal use of this document is free but if you wish to redistribute this

document, in whole or in part, for commercial gain, you must obtain permission from and acknowledge the author.

Things to do: Detailed machine type specs, URLs and links to resources, War stories - Time warp, Fixed on site, T3e OS internals, T3e logical to physical, T3e go faster bits express message queues, Multi CPU Process relocation, How to program, vectorisation, macro tasking, micro tasking, auto tasking, Large memory, expand ccc, expand Sparc superserver,

March 2000 V1.0.5

Copyright (c) 1999 by "Fred Gannett", all rights reserved.

This FAQ may be posted to any appropriate USENET newsgroup, on-line service, web site, or BBS as long as it is posted in its entirety and includes this copyright statement.

This FAQ may be distributed as class material on diskette or CD-ROM as long as there is no charge (except to cover materials).

This FAQ may not be distributed for financial gain except to the author.

This FAQ may not be included in commercial collections or compilations without express permission from the author.

Downloadable from Gannett's home page

[[By Archive-name](#) | [By Author](#) | [By Category](#) | [By Newsgroup](#)]
[[Home](#) | [Latest Updates](#) | [Archive Stats](#) | [Search](#) | [Usenet References](#) | [Help](#)]

*Send corrections/additions to the FAQ Maintainer:
newsposter@SpikyNorman.net (Gannett)*

Last Update May 15 2001 @ 00:45 AM