HPC at NCAR: Past, Present and Future

Tom Engel
Computational and Information Systems Laboratory
National Center for Atmospheric Research

26 May 2010

ABSTRACT: *The history of high-performance computing at NCAR is reviewed from Control Data Corporation's 3600 through the current IBM p575 cluster and new Cray XT5m, but with special recognition of NCAR's relationship with Seymour Cray and Cray Research, Inc. The recent acquisition of a Cray XT5m is discussed, along with the rationale for that acquisition. NCAR's plans for the new NCAR-Wyoming Supercomputing Center in Cheyenne, Wyoming, and the current status of that construction project, are also described.*

KEYWORDS: Cray-1A, Cray-3, legacy systems, XT5m, NCAR

1. Introduction

NCAR's use of high-performance computing (HPC) resources predates the coining of that term, as well as many others in routine use in recent years, such as "supercomputer" and "Moore's Law." Indeed, NCAR has had a long relationship with the Cray name, whether it is the man, his machines, his companies, or those companies keeping his namesake and legacy alive. From the CDC 6600, the first system installed at the NCAR mesa laboratory, to the Cray-1A and ultimately the Cray-3, NCAR's Computational and Information Systems Laboratory (CISL), formerly known as the Scientific Computing Division (SCD), used Seymour Cray's systems to advance computational capability in the atmospheric sciences.

That relationship persisted through several generations of Cray systems, but the mid 1990s brought change. On March 23, 1995, SCD issued its Accelerated Computing Environment (ACE) Request for Proposal (RFP). A day later Cray Computer Corporation filed Chapter 11. Less than a year later, SGI acquired a struggling Cray Research, Inc. (CRI).

In early 1996, NCAR completed the evaluation phase of ACE, and began subcontract negotiations with Japanese supercomputer manufacturer NEC. However, the ACE procurement was suspended when SGI/Cray filed a dumping complaint with the Commerce Department. A lengthy legal battle ensued, with anti-dumping duties imposed upon Japanese supercomputers and subsequent appeals, which culminated in a U.S. Supreme Court decision in February 1999 that let the dumping decision stand. During the three-year hiatus, NCAR continued necessary incremental augmentation of its HPC capacity with various Cray, SGI and IBM systems.

In 2000, SCD issued the Advanced Research Computing System (ARCS) RFP, hoping to recoup the HPC capacity losses that Moore's Law had imposed during the dumping investigation. That procurement and a subsequent Integrated Computing Environment for Scientific Simulation (ICESS) procurement, both of which were strongly competed, resulted in IBM displacing Cray in NCAR's HPC environment. The last Cray system was decommissioned in September 2002. Despite Cray's absence from NCAR, their relationship was maintained via cooperative scientific and model development activities between NCAR and other HPC centers with a Cray presence, most notably with the Community Climate System Model (CCSM) and Weather Research Forecast (WRF) model.

NCAR has again, in April 2010 after an eight-year absence, joined the esteemed ranks of the Cray Users Group. CISL determined during the preceding winter that the acquisition of a Cray XT5m system was advantageous, not only to evaluate it and the Cray software stack, but to provide a local development platform for NCAR scientists collaborating with institutions having much larger Cray XT systems.

Looking to the future, NCAR's partnership with the state of Wyoming will provide a new computational facility, substantially expanding CISL's ability to house high-performance computing and data storage and archival resources. The design of the facility was completed in April 2010 and groundbreaking is expected in June. Installation of HPC resources is currently anticipated for the first quarter of calendar 2012.

1. Past

NCAR's Founding

Meteorology departments were established at the Massachusetts Institute of Technology (MIT), the University of Chicago and other U.S. universities during the 1930s to investigate the physical principles governing the behavior of the atmosphere. World War II, with its massive sea, land and air operations, emphasized the importance of weather forecasting to military operations. After the war, though, the field of atmospheric science lost ground in comparison to other sciences, with 90% of American meteorologists employed by the federal government, mainly in weather forecasting. Basic research, the number of PhDs, and new students entering college in the atmospheric sciences was extremely low. In 1956, the National Academy of Sciences brought together a group of distinguished scientists to evaluate the state of meteorology in the United States. That group determined that it was imperative to establish a national institute, later called a national center, for atmospheric research to be operated by a consortium of universities with the support of the National Science Foundation (NSF)[1].

In 1960, the University Corporation for Atmospheric Research (UCAR) was founded with fourteen member universities, to manage the National Center for Atmospheric Research. One of the first tasks that Walter Orr Roberts faced, after his appointment as the first Director of NCAR, was to construct a facility. At the time, NCAR was loosely housed in some University of Colorado (CU) campus research buildings. Roberts worked to secure over 560 acres of Table Mesa, overlooking the Boulder valley. He envisioned a building that would be commensurate with the dramatic Fountain Formation sandstone outcrops that define Boulder's southwest skyline. The Deans of the Schools of Architecture at UCAR member universities unanimously selected Ieoh Ming Pei to be NCAR's architect. Groundbreaking occurred on June 9, 1964, and scientists, as well as computing equipment, began moving into the building in late 1966. The NCAR Mesa Laboratory was dedicated on May 10, 1967.

Figure 1. The National Center for Atmospheric Research, under construction in 1965.

Early Computing at NCAR

Before the NCAR Mesa Laboratory was completed, NCAR scientists occupied various buildings around Boulder, including a CU research building, the Armory building on 17th Street, an former CU women's dormitory, and CU's cyclotron building. The first computer system purchased by NCAR was a Control Data Corporation (CDC) 3600, with a clock speed of 0.7 MHz and 190 kilobytes of memory. That system was installed in November 1963 and located in a CU research building named PSR-2. The system was shipped before CDC's operating system was ready. The computing division's staff wrote a rudimentary operating system to utilize the machine.

The 3600 was CDC's team-designed, instruction-compatible follow-on to the CDC 1604, which had been designed by Seymour Cray. Seymour was already working on the 6600, starting over with a clean slate and targeting a system fifty times faster than the 1604 – but the technical hurdles with early silicon transistors (rather than germanium transistors, as was used in the 1604), system packaging and cooling led to numerous delays. Adding to that was Seymour's decision in 1962 to move the 6600 design and fabrication from the Twin Cities to Chippewa Falls, WI, so he and his small team could work without the constant interruptions and corporate distractions. The efforts of that small team (characterized

[1] The rationale was presented in what has become known as NCAR's 'blue book', the recommendations of that group to the NSF. A fascinating read for history buffs. See http://www.ncar.ucar.edu/organization/about/assets/bluebook1959.pdf

by Thomas Watson, IBM President, as "thirty-four people, including the janitor!") paid off in August 1963 when the CDC 6600 debuted – capable of three million instructions per second, it outperformed the then-champ, the IBM Stretch by a factor of three. The CDC 6600 retained the distinction of being the fastest computer in the world until Seymour's CDC 7600 successor was introduced in 1968.

The first computer system installed at the NCAR Mesa Laboratory was the CDC 6600, serial number 7. NCAR took delivery of S/N 7 at the end of December 1965 and it was operated in a CU research building until being installed at the Mesa Laboratory in December 1966. The 6600 was a revolutionary machine, not only for its speed, but also for its Freon refrigerant cooling technology. Having only 64 instructions, one can argue that the CDC 6600 was the first reduced instruction set computer (RISC); but that was well before the term was invented. It has been said that RISC actually means Really Invented by Seymour Cray.

NCAR's 6600 had 655k bytes of memory (65k 60-bit words; CDC display code used 6-bit bytes) and a clock speed of 10 MHz. Ten peripheral processor computers for managing input and output supported the central scientific processor. Input was via punched cards and seven-track magnetic tape; output was via two line printers, a cardpunch, a photographic plotter and magnetic tape. The CDC 6600 had an interactive CRT display console that allowed users to view results as they were produced.

In 1970, NCAR benchmarked the IBM 360 Model 195 and the CDC 7600, Seymour's follow-on system to the 6600. Seymour's system beat its competition, and CDC delivered serial number 12 of its 7600 line to NCAR in May 1971. S/N 12 had the same memory footprint as its predecessor 6600, but with a clock speed of 36.4 MHz and a sustained computing power of five times that of the 6600. NCAR's 7600 ran the same NCAR-developed operating system as the 6600 and, with NCAR's FORTRAN 70 compiler, most NCAR users could run on either platform without changes.

The CDC 6600 and 7600 each served NCAR for nearly twelve years, overlapping residency by six years. The 7600 was found to be less stable than the 6600. This caused SCD and NCAR scientists to develop, over a period of several years, robust job and file recovery procedures in the operating system and the equivalent of checkpoint-restart in the models themselves.

In the early 1970s, Control Data Corporation dominated the scientific computing industry. CDC had diversified, was servicing the 6600s, manufacturing and selling the 7600s, had new computer design teams struggling on the 8600 (Seymour's four processor follow-on to the 7600) and the Star-100, and CDC was ramping up its services offerings. These, along with a lawsuit against IBM, were financially demanding. IBM had attempted to stymie the sales of the 6600 with promises of a new IBM 360 model purported to be as fast as the 6600. William Norris, CDC CEO, waged and ultimately won a lengthy legal battle against IBM, claiming they were marketing phantom machines and willing to sell systems at a loss in an attempt to drive CDC out of business. In addition, the 8600, still being assembled with discrete components, was proving to be difficult to build, and to cool. In early 1972, Seymour uncharacteristically wrote a memo to William Norris describing his aversion to the machinations of a large corporation, as CDC had become, the problems he could not solve in the 8600, and that, in effect, he needed to move on. Seymour Cray left CDC later that spring and, along with six other engineers, founded Cray Research. Borrowing ideas from the 8600's cooling technology and the Star-100's vector architecture, and deciding to use integrated circuits rather than discrete components, that small team set out to create a new supercomputer which would later be called the Cray-1.

The Cray-1A

By the mid 1970s, NCAR's CDC systems were saturated and demand for additional computational resources was increasing. G. Stuart Patterson, SCD Director, NCAR, UCAR and the NSF were working on the future; they were constructing an addition to the NCAR Mesa Laboratory's computer center, and were assessing the state of the scientific computing industry. The first Cray-1 was being completed and Lawrence Livermore and Los Alamos National Laboratories were locked in a battle over who would purchase it. Seymour decided to break the logjam by *giving* S/N 1 to Los Alamos for six months, saying that after their half-year evaluation period they could either buy it, lease it, or give it back. NCAR learned of the early results of LANL's testing and entered negotiations to purchase the next system built by Cray.

On Wednesday, May 12, 1976, UCAR President Francis Bretherton signed an $8.86 million contract with Seymour Cray for a "new 5^{th} generation computer," culminating over three years of planning by NCAR. That contract resulted in the delivery of the Cray-1A S/N 3[2] on July 11, 1977, and made Cray Research a revenue-generating company. S/N 3 had a clock speed of 80 MHz, a peak computational rate of 160 megaflops, one megaword (eight megabytes) of memory, and 4.8

[2] Cray-1A S/N 2 was never built. Its chassis was torn down and sliced up for "marketing trinkets" after Los Alamos scientists convinced Seymour to add SECDED to the Cray-1's memory.

gigabytes of "high-speed" DD-19 disk. The Cray-1A chassis itself consumed 110 kWatts (see Table 2); with the cooling equipment and disk, the total power consumption was double that of the Cray-1A chassis. That system arrived at NCAR without software, and Cray Research's software development organization spent the remainder of 1977 on-site at NCAR developing the Cray Operating System (COS) Version 1 and the Cray Fortran Translator (CFT) Version 1.

Figure 2. Seymour Cray looks on as Frances Bretherton, UCAR President, signs the contract for the Cray-1A, May 12, 1976.

SCD had decommissioned the CDC 6600 two months before S/N 3's arrival. Once NCAR scientists were able to begin running on S/N 3, they were amazed at its speed – five times that of the 7600; and it proved to be a much more reliable and stable system. As time went on, SCD had to encourage scientists to run on the 7600 while the Cray-1's workload became backlogged.

In early 1983, SCD decided to decommission the 7600 and replace it with a second Cray-1A. Leading up to that decision, NCAR and SCD had engaged in a lengthy debate over whether a fast scalar processor system would be a more suitable augmentation to S/N 3's vector capabilities, since data processing was a significant part of the NCAR workload. But the price-performance of the Cray-1A won out. In May 1983, Cray-1A S/N 14, a "previously owned" system, was delivered to NCAR, and its scientists now had two Crays.

Multitasking Comes to NCAR

Increased demand on NCAR computing in the 1980s came on many fronts: higher model resolution, the addition of more physical processes, handling of water phase changes in the models, and early concerns over global climate change (at that time, the "next ice age"). In October 1986, the Cray-1A S/N 14[3] was replaced by a Cray X-MP/48, with a 256-megaword SSD and DD-49 disks. The new system's price totalled just over $20 million. The X-MP ran COS and had a peak speed of just under 1 gigaflop per second and represented a three-fold expansion of the computational resources at NCAR. The X-MP also began a revolution in NCAR modelling: the use of more than one processor by a single application. The pioneer of multitasking at NCAR was the team of Bert Semtner and Robert Chervin. Their work on parallelizing a global ocean model helped shake out performance issues, race conditions, error- and end-conditions in Cray's new multitasking libraries. As a result of this work, in 1987 Robert Chervin received Honorable Mention at the first Gordon Bell Awards, for sustaining 450 megaflops (54% of peak) with a global ocean model on the Cray X-MP/48.

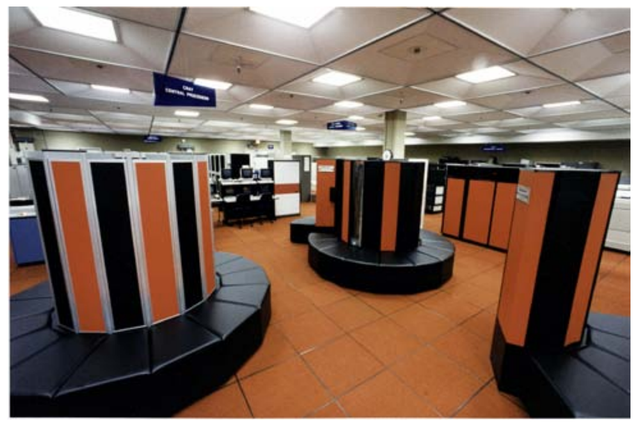

Figure 3. Wide-angle view of the Cray-1A S/N 3 (left) and the Cray X-MP/48 with SSD (right) at NCAR (circa 1987).

Subsequent to the acquisition of the X-MP/48, SCD used some research funding to acquire a Thinking Machines Corporation (TMC) CM-2 to explore emerging parallelization technologies and their applicability to atmospheric and oceanic codes and basic fluid dynamics algorithms.

One day in early 1987, the new SCD Director Bill Buzbee walked into the author's office (who, at the time, was a Cray site analyst at NCAR) and asked if there was any way the SSD could be used to alleviate system-level I/O bottlenecks, specifically, "How much effort would it take to modify COS to swap to SSD instead of disk?" He said that, since most of NCAR's models had restart capability built in, the risk of losing the workload due to a power failure might be worth the payoff. Ten days later, we booted a new, NCAR-modified version of COS that used an SSD partition for swap. Within days, NCAR was

[3] The Cray-1A S/N 14 can now be seen at the Smithsonian Air and Space Museum in Washington, D.C.

observing a nearly 30% improvement in system throughput. Buzbee characterized it as a "zero-cost augmentation of NCAR's computational capacity" – allowing NCAR's $20 million investment in the X-MP to be stretched further.

On a cold winter day, Friday, January 27, 1989, SCD laid an old friend to rest. After serving NCAR for nearly twelve years, the system that made Cray Research a moneymaking company and would have put NCAR at the top of the "top500" list, had it existed then, was decommissioned. On Wednesday of the following week, the Cray-1A S/N 3 was powered off.[4]

In June 1990, the X-MP/48, after serving NCAR for seven years, was replaced by a Cray Y-MP/864, named "shavano" for one of Colorado's "14ers". With eight six-nanosecond processors, sixty-four megawords of memory, and a 256 megaword SSD, the system nearly tripled the computational capacity and capability at NCAR, just as the X-MP had upon its arrival. The Y-MP also served as the system that brought Unix to NCAR's HPC environment. Cray Research had introduced UNICOS on the Cray-2 in March 1986, and subsequently had ported it to the MP line and declared that it was the operating system of the future. NCAR's Y-MP was first run with COS, and then UNICOS was introduced on a partition of the system, with the system eventually being converted wholly to UNICOS.

In June 1991, NCAR obtained a Y-MP/216, named "castle" (another Colorado "14er"), with special funding from the Model Evaluation Consortium for Climate Assessment (MECCA) – a predecessor to the Climate Simulation Laboratory. That system was dedicated to climate simulation, and ran UNICOS from its installation.

In the years following the acquisition of NCAR's two Y-MP systems, SCD again used research funding to acquire experimental parallel systems, including a TMC CM-5, and IBM RS/6000 cluster and an eight-processor IBM SP-1; but the user community's production computing demands continued to favor the "big iron".

In 1994, NCAR acquired a Cray Y-MP/8I (antero) and a T3D/64 which, after a couple month overlap, replaced the Y-MP/216 (castle), and were used exclusively by the Climate Simulation Laboratory (CSL). In October 1995, to augment its computing capacity, NCAR acquired two Cray J90/16 systems, one dedicated to the CSL (aztec) and one for the general NCAR community (paiute). Over the Thanksgiving holiday of 1995, the CSL's J90 system began crashing erratically while running the CSL workload. After a month-long investigation, it was determined that the instability was due to an inter-processor design issue, which Cray subsequently fixed. Cray made it up to NCAR by rebuilding the 'aztec' system with twenty processors.

Parting of the Crays

While NCAR was working on its X-MP and planning the acquisition of its Y-MPs, Cray Research was undergoing some radical changes. Just like CDC had, Cray Research had experienced nearly unimaginable success and growth, it was approaching 5,000 employees worldwide. Cray Research was funding the MP line and the development of the Cray-3, a massively parallel T3D, and a new Y-MP follow-on system called the C90. The company was being stretched, and financially stressed. CRI cut Steve Chen's design budget, and he left Cray in 1987 to found Supercomputer Systems Inc. While Steve had been the public face of the X-MP and Y-MP products, insiders realized that Les Davis, just as he had done with Seymour's earlier systems, was the man who had brought the machines to production. The effect of Chen's departure, and the October 1987 stock market crash, tarnished the previous public and investment community's view of Cray Research. Additionally, Seymour's Cray-3 work in Chippewa Falls, WI, had run into technical issues and, in August 1988, he decided to move the project to Colorado Springs, CO, purchasing a building and land from INMOS, a British government funded computer development project.

Then, in May 1989, John Rollwagen, CEO of Cray Research, after meeting with Seymour to discuss concerns over the financial drain of funding multiple design and development efforts, announced that Seymour and the Colorado Springs Cray-3 operation would be spun off from Cray Research as an independent company: Cray Computer Corporation. With $100 million in cash and $50 million in assets, Cray Computer Corporation (CCC) set out to complete the design of the Cray-3, using gallium-arsenide (GaAs) semiconductor material, three-dimensional modules, and cooled, as had been done with the Cray-2, in a chilled Fluorinert bath.

CCC bought integrated circuit fabrication assets from Gigabit Logic, and built its own gallium-arsenide chip fab, and by August 1990, it was producing its own GaAs die. Refinement of the fab process to improve yields, and working out the myriad issues with the assembly of the Cray-3 logic boards, slowed the assembly of a Cray-3 system. Eventually, Cray Computer assembled a four-processor system (S/N 5) and an eight-processor system (S/N 7).

Seymour realized that the Cray-3 may have missed its niche in the marketplace and began working on the Cray-

[4] The Cray-1A S/N 3 is still on display in the first basement of the NCAR Mesa Laboratory. Visitors still love to have their picture taken standing next to, and inside, the machine.

4, which would double the clock speed, but use the same semiconductor material and circuit board fabrication processes to put an entire CPU on a single module. In the interim, he felt it was necessary for potential customers to “kick the tires” on the Cray-3 – to prove that GaAs and the exotic 3-D module fabrication processes could hold up under a workload. Because of its proximity to Colorado Springs, Seymour approached NCAR with an offer.

The Cray-3

As he had done with Los Alamos and the Cray-1, Seymour offered to *give* a Cray-3 to NCAR, saying “If it will run at NCAR, it will run anywhere.” So, on May 24, 1993, CCC delivered the Cray-3 S/N 5 to NCAR, which SCD named “graywolf”. S/N 5 was a “two-octant tank” configured with two processors (instead of the four it could accommodate), 128 megawords (1 gigabyte) of memory and 20 gigabytes of disk; its processors ran with a two-nanosecond clock (500 MHz), and the system used 90 kWatts. The Cray-3 was the first system with a peak computational capability of 1 gigaflop per processor. The system ran Cray Computer’s CSOS (Colorado Springs OS) and had Fortran and C compilers.

Figure 4. Seymour Cray and the Cray-3 S/N 5. Photo (c) 1993 Stephen O. Gombosi, used with permission.

The early going for the Cray-3 at NCAR was rough. After some additional refinement in the manufacturing processes to improve the system’s reliability, and the repair of a floating square root problem (a design error in hardware round-off discovered by NCAR’s climate model), the system proved to be a stable workhorse. NCAR scientists, primarily from the Climate and Global Dynamics division, used S/N 5 for additional climate modelling and weather model development for over a year.

Unfortunately, Cray Computer Corporation exhausted its, and a considerable sum of Seymour’s, money and was unable to obtain additional funding from investors or prospective customers. On Friday, March 24, 1995, just a week after demonstrating the floating-point section of a Cray-4 processor, clocked at 980 picoseconds (1.02 GHz), to visitors from the Geophysical Fluid Dynamics Laboratory, Cray Computer Corporation filed Chapter 11.

ACE in the Hole, Tragedy, and Industry Turmoil

In 1994, the National Science Foundation and NCAR agreed to create an accelerated computing environment for use by the newly created Climate Simulation Laboratory (an outgrowth of MECCA), to be funded by the U.S. Global Change Research Program. In the months leading up to March 1995, NCAR had been busy formulating the requirements for its Accelerated Computing Environment (ACE) procurement. In many ways, NCAR anticipated that Cray Computer Corporation’s Cray-4 could be a strong competitor for ACE. NCAR released the ACE RFP the day before CCC filed Chapter 11. Subsequently, NCAR looked to fifteen other U.S and foreign supercomputer suppliers for ACE proposals.

Despite its shedding of the Cray-3 design effort in 1989, Cray Research continued to struggle. Its T90 follow-on to the C90, and last “big iron” vector system, was pricey, and not particularly reliable. Cray’s financial bleeding continued, losing $226 million in 1995, and though they submitted an ACE proposal, Cray Research sought “the refuge of a large corporate parent with deep pockets.” CRI was acquired by Silicon Graphics, Inc., in February 1996 for $740 million.

After Cray Computer’s financially-induced failure, Seymour had the opportunity to start over, and in late summer of 1996 he founded SRC Computers, Inc., with a vision of creating supercomputer with a ‘large, flat, high-performance memory subsystem’ with commodity microprocessors: a Cray with “Intel Inside.” Just a month later, on Sunday, September 22, 1996, Seymour’s Grand Cherokee was struck by a car attempting to pass him on the entrance ramp to I-25 from North Academy Boulevard in Colorado Springs; it rolled three times. Emergency crews spent an hour and a half extracting Seymour from the wreckage. Seymour’s neck was broken, and he had sustained massive head injuries. Seymour Cray died on Saturday, October 5, 1996, at age 71.

The ACE RFP was released to fourteen U.S. and two foreign supercomputer vendors. ACE defined three- and five-year scenarios. The three-year scenario provided $13.25 million for a sustained performance of five gigaflops by mid 1996. The five-year scenario provided $35.25 million for systems to meet the sustained performance goals of five gigaflops by mid 1996, twenty-five gigaflops by October 1997 and fifty gigaflops by October 1999. Four vendors responded and three of those proposals were within in the competitive range. NCAR issued guidelines for Best and Final Offers (BAFO) in October 1995, requesting vendors focus on a slightly amended five-year scenario and that they perform final Live Test Demonstrations in February 1996.

Subsequent to the final LTDs and BAFOs, NCAR entered subcontract negotiations with Federal Computer Corporation (FCC), a domestic 'reseller' of the NEC SX-4. Later that year (July 29, 1996) SGI/Cray filed a dumping complaint with the Commerce Department against NEC. That complaint suspended the ACE procurement process, and disallowed NCAR from presenting benchmark results at the SC96 conference.

In his testimony to the International Trade Commission in August 1997, Dr. Bill Buzbee stated that NCAR "concluded that FCC offered and demonstrated overwhelmingly superior technical performance and low risk relative to the Cray Research five-year offer." NEC was able to demonstrate 13 gigaflops on a prototype SX-4 on the ACE community climate model benchmark and had proposed a phased installation of four 32-processor SX-4 systems over three years. Cray had proposed an ensemble of vector and scalar systems, only one of which could be benchmarked, with a sustained performance well short of the SX-4's.

On September 26, 1997, the ITC voted to impose anti-dumping duties on imported Japanese supercomputers. Commerce then set anti-dumping duties of 454% against NEC, for "pricing its supercomputers at 80% below cost," of 173% percent against Fujitsu and of 314% percent on all other Japanese firms. NEC and its HNSX U.S. subsidiary appealed the decision to the federal Court of International Trade, claiming that the ITC had excluded IBM, SGI and other vendors from their analysis. That court rejected the lawsuit. In November 1998, NEC petitioned the U. S. Supreme Court to review the federal court's decision. On December 22, 1998, the federal Court of International Trade overturned the ITC ruling, stating the ITC failed to show a causal relationship between Cray's deteriorating performance and the small number of Japanese supercomputers sold in the U.S and stated that the ITC must reconsider the dumping penalties. Ultimately, on February 22, 1999, the U.S. Supreme Court rejected NEC's appeal and let the ITC decision stand.

NCAR still needed to augment its computing capacity and capability. The hopes it had hung on the ACE procurement were being slowly strangled in a legal/political noose while the dumping battle raged. SCD decided to acquire, via sole-source procurement processes based primarily on maintaining compatibility with its current computational technology, a series of systems that would keep NCAR's scientific activities on track. The first acquisition was a Cray C90/16 to replace the CSL's Y-MP/8I; SCD kept the 'antero' name for the new system, despite the four-fold increase in computational capacity. The C90 was placed into production in January 1997, along with a new Cray J90se/24 (ouray). Two months later, NCAR's Cray T3D was upgraded from 64 to 128 processors. Finally, in March 1998, SCD acquired a fourth J90 system, the J90se/24 named chipeta. As it would turn out, chipeta would be the last system NCAR would acquire with the Cray nameplate, though two SGI 2000 systems would be purchased in the subsequent months.

Transition

In 1998, Dr. Bill Buzbee retired from SCD and NCAR selected Al Kellie, former director of the Canadian Meteorological Centre, to lead SCD. At the time, NCAR's computational future was still hobbled by the last battles of the dumping battle, and it was apparent that NCAR would be in practice, if not in fact, limited to acquiring future supercomputer systems only from domestic suppliers. NCAR needed additional computational capacity and the state of the U.S. supercomputer industry was distressing. Though SGI had flexed its financial muscles in 1996 by acquiring Cray Research, it had failed in the following years to define itself, or to capitalize upon the technologies it had acquired. By the time the dumping battle was concluded in the U.S. Supreme Court, it was not clear to SCD and NCAR whether the United States was going to stay in the supercomputing business. Conversely, it was clear that Japan's Earth Simulator, which began a two and one-half year construction phase in October 1999, would displace all other systems (in fact, after it was commissioned in March 2002, it would retain the #1 position on the Top500 list for an unprecedented two years).

NCAR's need for additional computing power, its "buy American" restrictions, and the indications that the domestic supercomputer market, and its future, would be dominated by distributed-memory cluster systems, led SCD to a two-fold decision. That was to first begin the development of the next competitive RFP for HPC resources, and second, as an interim augmentation of NCAR's computational capacity, acquire a POWER3-based IBM SP RS/6000 cluster (the only safe and viable alternative at the time). The IBM SP, blackforest, was delivered to NCAR on August 11, 1999, two months after a small test system, babyblue, had been installed to help SCD and NCAR become familiar with its operational, administrative, and user-environment peculiarities.

After the primary NCAR user community applications had made the transition to the IBM SP, SCD decommissioned the Cray C90 in November 1999. The Cray T3D had been decommissioned in June. This left only the Cray J90 systems, two of which would be powered off the following June, and the last one of which would be decommissioned in September 2002.

21st Century

As SCD was developing the technical requirements for its Advanced Research Computing System (ARCS) RFP, the domestic supercomputer industry continued its turmoil. SGI was self-destructing, and Tera Computer Company bought the Cray brand and product line from SGI in March 2000 for $15 million in cash, a $36.3 million promissory note, and a million shares of stock. Tera then renamed itself Cray, Inc.

SCD released its ARCS RFP on November 1, 2000. Only three vendors submitted ARCS proposals: IBM, Compaq and SGI. Subsequent BAFOs gave IBM the edge over the proposed Compaq solution. The ARCS agreement with IBM, valued at $24 million, resulted in four phases of equipment. The first phase was an upgrade to and expansion of the POWER3 systems acquired in 1999. ARCS Phase 2 added a large POWER4-based IBM p690 cluster, named bluesky, which put NCAR in the #10 position on the November 2002 Top500 list. A subsequent Phase 3 expansion of that system to fifty frames in October 2003 allowed climate scientists to run the Intergovernmental Panel on Climate Change (IPCC) simulations for the 4th Assessment Report that was published 2007. Several dozen NCAR staff members who contributed to that work, and the scientific findings of the simulations which were in part conducted on bluefire, shared in the Nobel Committee's recognition of the IPCC and former Vice President Al Gore with the 2007 Nobel Peace Prize. The fourth and last phase of the ARCS agreement augmented bluesky's computational capacity with a POWER5-based p575 cluster, named bluevista, in August 2005.

A few months before bluevista was installed at NCAR, SCD began planning for its next RFP. In December 2005, SCD released its Integrated Computing Environment for Scientific Simulation (ICESS) RFP with award valued at $15 million. Again, only three vendors submitted proposals: Cray, IBM and SGI. Evaluation of the best value embodied in the ICESS proposals swayed between Cray and IBM, with the BAFOs giving only a slight edge to IBM. Contract negotiations led SCD and UCAR's Contracts office to again teeter to consider Cray's solution, but ultimately the negotiating teams reached an agreement that would be acceptable to all parties, including UCAR's NSF sponsor.

The ICESS subcontract resulted in a three and a half year agreement with equipment delivery taking place in two phases. Phase 1, delivered in October 2006, was a POWER5+ p575 system, named blueice, to carry NCAR forward until IBM's POWER6-based HPC cluster could be deployed.

The same month that the first ICESS equipment came to NCAR, the NCAR Director's office completed the restructuring of NCAR. Formerly, the organization had been comprised of a dozen divisions and institutes; the restructuring consolidated those divisions into a smaller number of laboratories. The Institute for Mathematics Applied to Geosciences (IMAGe) and SCD were merged into the Computational and Information Systems Laboratory (CISL).

2. Present

Currently, CISL's HPC resources are comprised of the seventy-seven teraflop IBM Power 575, POWER6-based, cluster (bluefire) and a twenty-three teraflop IBM BlueGene/L system (frost); the latter of which serves as an NSF TeraGrid resource. CISL is deploying a Globally Addressable Data Environment (GLADE) that will, by summer 2010, provide nearly two petabytes of common data storage for all NCAR systems. In addition, the data held in NCAR's Mass Storage System (MSS) is approaching ten petabytes, with a growth rate of approximately two petabytes per year. The NCAR-developed MSS software is being replaced by HPSS (the High Performance Storage System), and CISL plans to move all data in the NCAR MSS into HPSS by January 2011.

During 2009, CISL determined that, in preparation for the NCAR-Wyoming Supercomputing Center (NWSC), it was of strategic importance to evaluate a number of new technologies, including alternate batch and resource management subsystems, filesystems, as well as explore wide-area network (WAN) filesystems. Aided by unsolicited proposals received in the fall from both IBM and Cray, but not wanting to deplete resources targeted for the NWSC, CISL was able to pool enough money and support to acquire a Cray XT5m as that test platform. NCAR scientists with resource allocations at Department of Energy (DOE) laboratories supported the Cray acquisition as well, opining that a local development platform would be a valuable asset for their collaborative DOE work.

On Monday, April 26, 2010, after nearly eight years absence, NCAR had an old friend return in the form of a single cabinet Cray XT5m system, named lynx. The name was chosen after a CISL staff member suggested “A lynx is just a small jaguar.” The system installation went exceptionally smoothly, and the two-day acceptance test was completed the following Thursday. CISL is now familiarizing itself with the XT5’s hardware and the Cray software stack. CISL plans to provide first “friendly user” access to the system on June 1, 2010.

Figure 5. The new Cray XT5m, lynx, at NCAR (no, it’s not Larry from Frontier Airlines).

3. Future

SCD had expanded its computing facility in 1976 to accommodate the Cray-1A. Shortly after the first ARCS equipment drop in 2000, SCD realized that it would ultimately outgrow its present Mesa Laboratory computing facility, but a more pressing issue was the electrical power available to the facility’s equipment, which was, and still is, capped at 1.2 megawatts. That cap forced the installation of the IBM POWER6 system, bluefire, to be performed in two stages; the first third of the equipment was installed and run through acceptance testing, then bluefire’s predecessor POWER5+ system was powered off in order to power up the remainder for acceptance testing.

In 2003, SCD began the process of evaluating its future facility options. SCD explored the expansion of the Mesa Lab facility and co-location facilities in the Boulder/Denver area. NCAR considered partnerships for a new facility with a number of organizations along the Front Range, giving CU-Boulder and the University of Wyoming particularly close scrutiny. Ultimately, with support from NSF and the UCAR Board of Trustees, NCAR chose in January 2007 to construct a new NCAR-Wyoming Supercomputing Center (NWSC) in a partnership with a consortium consisting of the University of Wyoming, the State of Wyoming, the Cheyenne–Laramie County Corporation for Economic Development, the Wyoming Business Council, and Cheyenne Light, Fuel and Power Company. The Wyoming partnership is providing $20 million for facility construction and $1 million annually for facility operations. The National Science Foundation will supply the balance of the expected $76 million total cost of the NWSC.

To be located on a 24-acre site in the North Range Business Park west of downtown Cheyenne, Wyoming, the NWSC is intended to be a world-class center for high performance scientific computing in the atmospheric and related geosciences. Consistent with its mission, and NCAR’s work on global climate assessment and stewardship, the facility will be a leader in energy efficiency, incorporating the newest and most efficient designs and technologies available.

The NWSC facility is modular in design, approximately 150,000 sq.ft. in size, and will ultimately accommodate up to four raised floor data center modules. Upon initial build-out, anticipated in late 2011, the structure will accommodate two raised floor areas, each of approximately 12,000 sq.ft., along with associated mechanical and electrical support spaces, a network operations center, disk storage space, and general office and public display areas. Commensurate with the modular design and construction plans, only 12,000 sq.ft. of raised floor will be installed during the initial facility construction phase, and the facility will be provisioned with electrical and mechanical equipment sufficient to support only this first raised floor area. An additional environmentally controlled area of 6,000 sq.ft. will be available for location of data archival equipment. See Table 3 for additional structural details.

The NWSC incorporates a mechanical system that utilizes 100% recirculation of air coupled with cooling towers to meet computing equipment cooling needs. The facility will support both air and liquid (chilled water) cooled HPC technology. This design solution, in combination with the installation of a high-efficiency cooling plant and fan systems, will enable the facility to achieve compressor-free cooling for the majority (96%) of the year and to realize a PUE[5] value of less than 1.1. At present, LEED Gold certification of the facility is planned, and NCAR and the NWSC project design team are exploring options to obtain LEED Platinum (the highest level) certification.

CISL is currently working on the requirements for the computational and storage equipment to be placed into the NWSC in 2012 and anticipates the release of RFPs for the acquisition of that equipment in December 2010. Resources will include high-performance computing, data analysis and visualization systems, large scale, high-performance filesystems and disk storage, and a large HPSS-based tape archive. A budget of approximately $36 million is currently targeted for those resources. Based upon the science requirements CISL has gathered, general expectations can be loosely characterized as:

HPC:	1 to 1.5 petaflops peak
Disk Storage:	12-15 petabytes
Archive:	50 petabyte capacity initially, ~25 petabytes/year growth rate

Conclusion

NCAR is glad to be again associated with the distinguished colleagues of CUG, and CISL looks forward to Cray's response to its NWSC-1 procurement.

Acknowledgments

The author would like to thank colleagues, users and vendor staff for a wealth of information from which, along with the author's and their first-hand experience, this paper is derived. Some supplementary information was taken from Charles J. Murray's book, "The Supermen: The Story of Seymour Cray and the Technical Wizards behind the Supercomputer", 1997, John Wiley & Sons, Inc. publisher.

About the Author

Tom Engel is a High Performance Computing Specialist with the Operations and Services Division of the Computational and Information Systems Laboratory of the National Center for Atmospheric Research. He is a former employee of Cray Computer Corporation and Cray Research, Inc., with degrees in the atmospheric and physical sciences. Tom can be reached at the National Center for Atmospheric Research, P. O. Box 3000, Boulder, CO 80307-3000 USA, e-mail: engel@ucar.edu.

[5] Power usage effectiveness (PUE) is the ratio of the total amount of power provided to a data center divided by the power load delivered to the HPC, disk and related IT equipment. This quantity is a key measure of the overall energy efficiency of computing center operations.

Delivered	Retired	System	Clock (MHz)	Memory (GB)	Peak GFLOPs
1963/11/01	1966/06/01	CDC 3600	0.7	0.000032	0.001333
1965/12/30	1977/05/01	**CDC 6600 S/N 7**	10	0.000655	0.010000
1971/05/01	1983/04/01	**CDC 7600 S/N 12**	36.4	0.000655	0.036400
1977/07/11	1989/02/01	**CRI Cray 1A S/N 3**	80	0.008	0.16
1983/05/01	1986/05/01	**CRI Cray 1A S/N 14**	80	0.008	0.16
1986/10/01	1990/05/01	**CRI Cray X-MP/4**	117.6	0.064	0.9408
1988/09/01	1993/04/01	TMC CM-2/8192 (capitol)	7	1	7.168
1990/06/01	1997/06/30	**CRI Cray Y-MP/8 (shavano)**	166.67	0.512	2.667
1991/07/16	1994/09/01	**CRI Cray Y-MP/2 (castle)**	166.67	0.512	0.667
1992/02/01	1997/09/01	IBM RS/6000 cluster	66	0.512	0.528
1993/03/01	1996/11/01	TMC CM-5/32 (littlebear)	32	1	4.096
1993/08/01	1995/09/01	IBM SP1/8 (wildhorse)	62.1	0.064	1.99
1993/10/01	1995/03/25	**CCC Cray-3/4 (graywolf)**	500	1	4
1994/06/01	1996/12/01	**CRI Cray Y-MP/8I (antero)**	166.7	0.512	2.667
1994/07/01	1997/03/01	**CRI Cray T3D/64**	150	4	9.6
1995/10/01	2000/06/30	**CRI Cray J90/20 (aztec)**	100	1	4
1995/10/01	2000/06/30	**CRI Cray J90/16 (paiute)**	100	0.5	3.2
1997/01/01	1999/11/30	**CRI Cray C90/16 (antero)**	243.9	1	15.61
1997/03/01	1999/06/01	**CRI Cray T3D/128**	150	8	19.2
1997/01/01	2001/11/30	**CRI Cray J90se/24 (ouray)**	100	1	4.8
1997/05/01	1999/05/14	HP SPP-2000/64 (sioux)	180	8	46.08
1998/03/01	2002/09/03	**CRI Cray J90se/24 (chipeta)**	100	1	4.8
1998/06/01	2002/07/15	SGI Origin2000/128 (ute)	250	16	64
1999/04/14	2004/09/30	SGI Origin2000/16 (dataproc)	250	16	8
1999/06/18	2000/04/13	IBM SP/32 (16) WH1 (babyblue)	200	32	25.6
1999/08/11	2000/05/13	IBM SP/296 (148) WH1 (blackforest)	200	296	236.8
1999/11/01	2002/02/19	Compaq ES40/32 (prospect)	500	8	32
2000/04/13	2005/01/10	IBM SP/64 (16) WH2 (babyblue)	375	64	96
2000/05/13	2001/10/25	IBM SP/604 (151) WH2 (blackforest)	375	302	906
2001/10/26	2005/01/10	IBM SP/1308 (318) WH2/NH2 (blackforest)	375	702	1962
2002/01/09	2007/01/17	IBM p690/16 (1) Regatta-H (bluedawn)	1100	32	70.4
2002/07/15	2008/04/23	SGI Origin3800/128 (chinook/tempest)	500	256	128
2002/10/31	2003/10/17	IBM p690/32 (38) Regatta-H/Colony (bluesky)	1300	2560	6323.2
2003/10/02	2007/11/30	IBM p690/32 (2) Regatta-H/Federation (thunder)	1300	128	332.8
2003/10/18	2007/03/05	IBM p690/32 (50) Regatta-H/Colony (bluesky)	1300	3328	8320
2004/07/12	2009/12/04	IBM e325/2 (128) Opteron Linux cluster (lightning)	2200	528	1144
2005/01/31	2008/11/30	IBM e326/2 (64) Opteron Linux cluster (pegasus)	2200	272	580.8
2005/03/15	2009/09/05	IBM BlueGene/L (1024/2) (frost)	700	1048.6	5734.4
2005/08/16	2010/09/22	Aspen Nocona/InfiniBand Cluster (coral)	3200	96	563.2
2005/08/27	2008/09/29	IBM p575/8 (2) POWER5/HPS (otis)	1900	32	121.6
2005/08/27	2008/09/29	IBM p575/8 (78) POWER5/HPS (bluevista)	1900	1248	4742.4
2006/10/18	2008/06/17	IBM p575/16 (2) POWER5+/HPS (icecube)	1900	64	243.2
2006/10/18	2008/06/16	IBM p575/16 (112) POWER5+/HPS (blueice)	1900	4832	13312
2008/08/13	2008/11/12	IBM Power 575/32 (4) POWER6/IB (firefly)	4700	640	2406.4
2008/11/12	2012/06/06	IBM Power 575/32 (6) POWER6/IB (firefly)	4700	896	3609.6
2008/04/24	2008/06/16	IBM Power 575/32 (43) POWER6/IB (bluefire_1)	4700	4608	25868.8
2008/06/17	2009/02/24	IBM Power 575/32 (127) POWER6/IB (bluefire)	4700	12096	76403.2
2009/02/25	2012/06/06	IBM Power 575/32 (128) POWER6/IB (bluefire)	4700	12224	77004.8
2009/09/06	2012/03/31	IBM BlueGene/L (4096/2) (frost)	700	4194.3	22937.6
2010/04/26	2013/04/29	**Cray XT5m (912/76) (lynx)**	2200	1296	8129.6

Table 1. A tabular history of high-performance computing systems at NCAR.

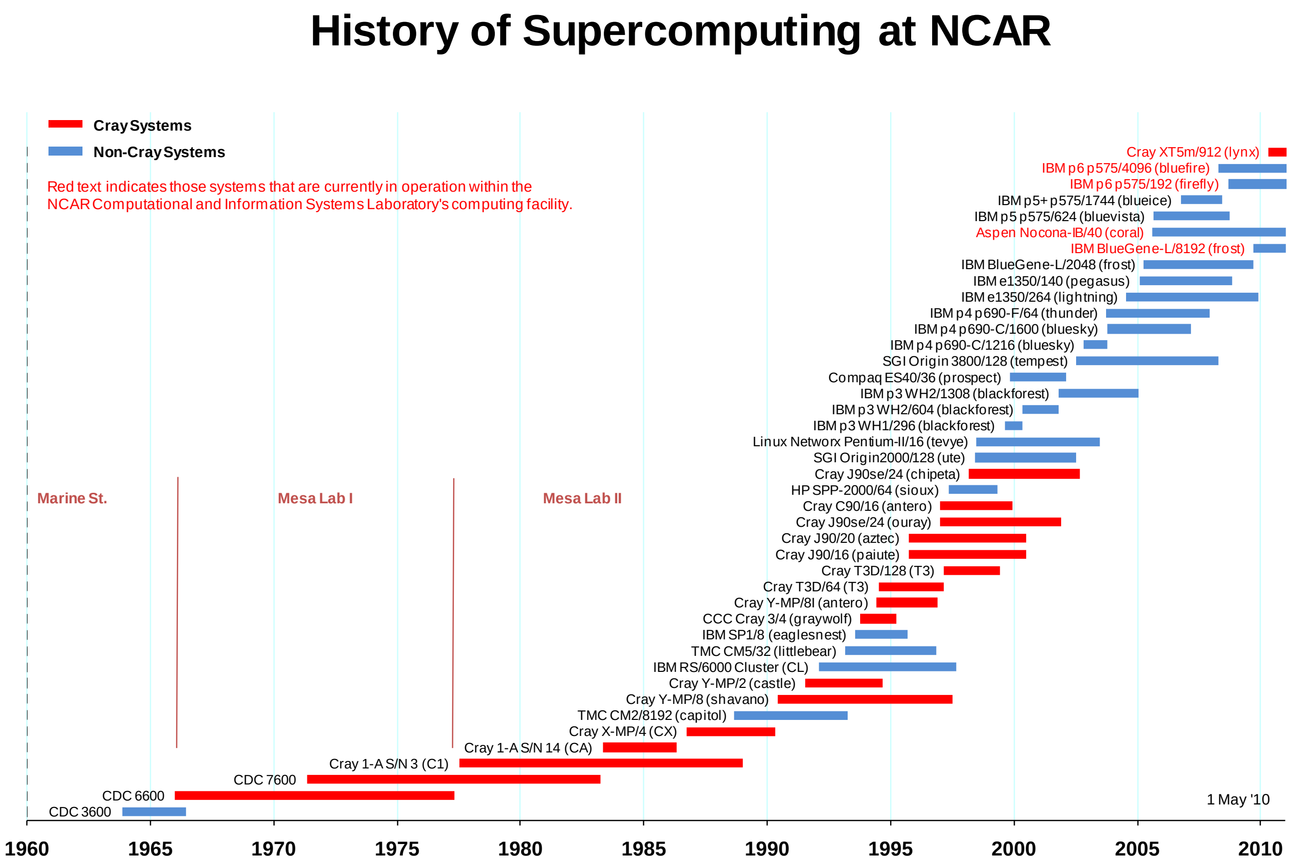

Figure 6. A graphical history of high-performance computing systems at NCAR.

Technology	Year	
Cray-1A S/N 3	**1977**	
Processor Speed	80 MHz	
Random Access Memory	8 MBytes	
Disk Storage	4.8 GBytes	
Power	110 kWatt	(~2x for cooling and disk)
iPhone 3GS	**2009**	
Processor Speed	600 MHz	7.5 x Cray-1A
Random Access Memory	256 Mbytes	32 x Cray-1A
Flash Storage	16 GBytes	3.33 x Cray-1A
Power	3 Watt (max)	0.000027 x Cray-1A

Table 2. Thirty years of technological development

NCAR-Wyoming Supercomputing Center Structural Data	
Size of structure	153,302 square feet
	17,680 penthouse
	170,982 square feet
Height of structure	62 feet to top of penthouse
Square footage of raised floor	2 raised floor modules, 12,000 square feet each
Depth of raised floor	10 feet
Height from the raised floor surface to the raised floor area ceiling	12 feet
Height of interstitial space above raised floor ceiling space to hard roof	Average 9 feet
Size of Administration / Visitor Center space	9,550 first floor
	6,445 second floor
	15,995 net square feet
Size of the Network Operations Center (NOC)	3,560 net square feet
Size of the HPSS tape archival (Mass Store) area	2,380 net square feet
Size of the HPSS expansion area	4,026 net square feet
Size of the Mechanical / Electrical Utility Plant (CUP) space (Total CUP numbers do not include data room service floor or HPSS service floor)	39,846 - Chiller floor
	11,394 - Main floor
	31,142 - Electrical floor
	17,654 - Penthouse
	100,036 - Total CUP
No elevation changes (ramps) from the loading dock to the equipment raised floor space	0

Table 3. Structural information for the NWSC.