From Kilobytes to Petabytes in 50 Years

"The day when the scientist, no matter how devoted, can make significant progress alone and without material help is past."
—E. O. Lawrence, founder of Lawrence Livermore National Laboratory, on accepting the 1939 Nobel Prize for Physics

THE history of Lawrence Livermore National Laboratory is inexorably tied to the evolution of supercomputers—the largest, fastest, most powerful computers in the world. Even before the Laboratory's gates opened for the first time in September 1952, founders E. O. Lawrence and Edward Teller recognized that computers were needed to better calculate the thermonuclear explosions for the nuclear weapons the "Rad Lab" in Livermore was destined to design.

Designing nuclear weapons and predicting their behavior has always been a difficult technical and scientific challenge. In a thermonuclear explosion, matter is accelerated to millions of kilometers per hour while experiencing densities and temperatures found only in stars. In addition, weapon designers needed to identify and understand the important physical properties of matter under these exotic conditions. With little experimental data available, Livermore's designers turned to computers to simulate and visualize the processes and the physics of nuclear weapons.

To fulfill its critical national defense mission, the Laboratory constantly sought out the most advanced computers with the most capability. In the 1990s, with the cessation of underground nuclear testing, advanced supercomputers figured prominently in plans for stockpile stewardship, helping scientists predict the behavior of the aging nuclear stockpile to better assess its safety, reliability, and security.

Mag Tape and Punch Cards

Livermore's first supercomputer, the Remington-Rand Univac-1, had 5,600 vacuum tubes and was over 2 meters wide and 4 meters long. Between April 1953 and February 1957, the Univac executed as many calculations as 440 human "calculators" could perform in 100 years if they worked 40 hours a week, 52 weeks a year, and made no mistakes. Memory, however, was an issue.

The Univac's memory consisted of mercury tanks that could store 9 kilobytes of data—a tiny fraction of what today's pocket-sized handhelds can hold. The code that performed all its operations was stored on magnetic tapes that had to be loaded into the machine in parts. Calculations could involve as many as nine tapes, and the nine reel mechanisms were troublesome, accounting for much of the machine's 25 percent downtime. Clearly, machines with more memory were needed.

With the arrival of the IBM 701 in 1954, scientists expected that nuclear explosives computations would run much faster. The IBM, which was the first fully electronic computer, was 12 times faster than the Univac, had twice the memory, and primarily used punch cards for input and output. Scientists took advantage of the improved capabilities to increase resolution and add more detailed physics, so the computational runs continued to average 100 hours.

A series of IBM machines followed the 701. The IBM 704—twice as fast as the 701—even played a part in the early

Weapons

Computations

Engineering

Physics

Chemistry & Materials Science

space race between the U.S. and the Soviet Union. Soon after the launch of the Soviet Sputnik I satellite in October 1957, the Laboratory received an urgent request to help predict when the satellite would come back to Earth. Livermore's IBM 704s were the only computers in the U.S. able to perform the calculations. Joe Brady, a now-retired Laboratory scientist, recalls, "We used two 704s for 70 hours straight, only stopping to rush outside to see the satellite orbiting overhead." Laboratory computation workers accurately calculated the satellite's plunge into the atmosphere in early December, an extrapolation of 58 days from launch. The 704s eventually gave way to IBM 709s, which were faster still, thanks to special-purpose input/output channels to speed up processing, and batch processing—a new technique that permitted many individual tasks to be processed without a human operator's assistance.

In the late 1950s, Edward Teller proposed that the Laboratory commission a computer from commercial suppliers. In May 1960, Remington-Rand delivered the Livermore Advanced Research Computer (LARC) built to Livermore's specifications. At that time, there was an international moratorium on nuclear testing, and upgraded computing capabilities were urgently needed by weapon designers. With a high-speed magnetic core memory for storing about 240 kilobytes and 12 auxiliary memory drums for storing about 24 megabytes more, the LARC had such dense wiring that technicians had to use special tools similar to surgical instruments to probe its insides. Next came the "Stretch," an IBM machine with about 780 kilobytes of memory that could perform 100 billion calculations in a day.

As the 1960s progressed, the computer market changed. Most manufacturers abandoned the highly specialized large-computer market of the national laboratories to concentrate on the computer needs of the rapidly growing business and financial markets. In 1963, the Laboratory turned to Control Data Company (CDC), which furnished all of Livermore's supercomputers for the next 15 years, including the CDC 6600 in 1964 and the CDC 7600—10,000 times faster than the original Univac-1—in 1969. The Laboratory received serial number 1 of each of the machines and, by using them, helped CDC ready their computers for the wider commercial market.

The Univac was the first computer to store information on magnetic tape. Running a program was a hands-on operation, with a physicist or programmer toggling console switches to execute the problem. Although highly accurate, the Univac was cantankerous, breaking down two or three times a day. Early workers regarded it as an "oversized toaster."

Entering a Parallel Universe

About this time, computers began exploiting computational parallelism. The CDC STAR-100s in 1976, followed by the Cray 1s, introduced vector architectures. Cray came out with the first closely coupled processor systems with its two-processor Cray X-MPs. The final Cray machine, installed at the National Energy Research Scientific Computing Center (now located at Lawrence Berkeley National Laboratory), had 16 central processing units (CPUs) and about 2 megabytes of memory.

In the early 1990s, massively parallel machines—that is, employing scalar architectures—such as the Meiko and the BBN (by Bolt, Beranek, and Newman) began to arrive at the Laboratory. As Mike McCoy, a deputy associate director for Livermore's Computation Directorate, explains, "About this time, we began looking at not just sheer capability, which has been the motivator at the Lab since day one, but price performance as well. Up to and including the Crays, we would depend on a single vendor to supply the capability we needed.

Nonproliferation

Energy & Environment

Lasers

Biotechnology

Stockpile Stewardship

Part of getting the price performance we needed involved moving away from specialized processors for parallel machines to commodity processor systems." The Meiko and the BBN were the first supercomputers of this type. Instead of using a few, enormous, one-of-a-kind processors, the Meiko and the BBN used many mid-sized workstation processors (the BBN, for instance, had 128 such processors). "We learned how to build software for parallel systems on these computers," notes McCoy. "These systems were what made us able to transition to the massively parallel ASCI [Advanced Simulation and Computing program, formerly called Accelerated Strategic Computing Initiative] systems."

In 1995, the Department of Energy and its defense laboratories—Livermore, Los Alamos, and Sandia—were directed to undertake the activities necessary to ensure continued stockpile performance in the absence of underground nuclear testing. DOE's ASCI program is a key component to meeting this challenge. The ASCI program is developing a series of ever more powerful, massively parallel supercomputers that employ thousands of processors working in unison to simulate the performance of weapons in an aging nuclear stockpile. The second ASCI supercomputer—the Blue Pacific, built by IBM—was received at Livermore in September 1996. It was installed, powered up, and running calculations within two weeks. IBM's ASCI White, which was delivered to the Laboratory in three stages during the summer of 2000, is currently the world's most powerful computer. Performing 12 trillion operations per second (teraops), it is 30 billion times faster than the Laboratory's very first computer, the Univac-1.

In late 1999, Livermore researchers achieved a major milestone with the first-ever three-dimensional simulation of a nuclear weapon's primary (the first stage of a hydrogen bomb) using the ASCI Blue Pacific. The simulation ran a total of 492 hours on 1,000 processors and used 640,000 megabytes of memory in producing 6 million megabytes of data contained in 50,000 graphics files. A second major milestone, a three-dimensional simulation of a nuclear weapon secondary, was completed on ASCI White in spring of 2001. Late in 2001, Livermore and Los Alamos met a third milestone on this system, coupling the primary and secondary.

Forward to the Future

With all that has occurred in the last 50 years, it's nearly impossible to predict what the far future will hold. "To meet ASCI's requirements, more powerful processors with more memory are needed to create a proxy of the world around us, from the microscale to the macroscale," says Dona Crawford, associate director of Computation. "At the same time, we are creating terabytes—soon to be petabytes—of data." Two trends, Crawford notes, need to continue into the near future. First, the Laboratory must acquire faster processors with more memory for simulation and modeling. Second, new ways must be created for storing, finding, visualizing, and extracting the data. "We need to merge high-end computing and high-end information technology," she concludes. "Scientific data management, in particular, is becoming more of an issue." (See the box on software development, .)

Within three years, the ASCI community plans to locate a 60-teraops machine with approximately 20,000 processors— the Purple machine—at Livermore in the soon-to-be-built Terascale Simulation Facility. Groundbreaking for this facility will occur in spring of this year. Beyond Purple lies a world of tantalizing prospects, including BlueGene/L (L stands for

The ASCI White, with power to perform 12 trillion operations per second, was delivered to the Laboratory during the summer of 2000.

Software Development

The supercomputers Livermore acquired were often the first of their kind—sometimes even prototypes of the final version—and had little support software. As a result, Livermore's scientists took the lead in developing software for operating the system (such as assemblers, loaders, and input/output routines) as well as for simulating and modeling physical phenomena. Because Laboratory users pushed the machines to their limits, Livermore's programmers had to find—or often invent—the most efficient programming and computing techniques. For instance, when certain aspects of the FORTRAN computer language turned out to be awkward or limiting for scientific applications, software developers created an enhanced version called LRLTRAN (Lawrence Radiation Laboratory FORTRAN). It took nearly two decades for many of the advanced features in LRLTRAN to be incorporated into standard FORTRAN. In addition, Livermore developed the time-sharing concept—in which a central processing unit (CPU) alternates between working on several jobs at once rather than one at a time—into its first practical use for supercomputers. The Laboratory also led the way in computational physics (the numerical simulation of physical phenomena) on supercomputers. Computer codes often hundreds of thousands of lines long are used to model complex processes that are too difficult or impossible to calculate exactly.

This expertise in codes continues today, with computer scientists writing or adapting codes for large parallel machines such as the Advanced Simulation and Computing (ASCI, for its former name, Accelerated Strategic Computing Initiative) systems. The sophisticated codes now under development promise a level of physical and numerical accuracy more like that of a scientific experiment than a traditional numerical simulation. In materials modeling, for instance, ASCI White will track 10 billion atoms simultaneously, beginning to predict what scientists will see when imaging materials through electron microscopes.

Interpreting, visualizing, and accessing the data are themselves challenges. From the early days of simple x–y plots to today's complex three-dimensional images, Livermore computer scientists have developed programs to help researchers access massive quantities of data in visual formats. This capability is particularly important for the future, given that ASCI-level supercomputers generate terabytes—soon to be petabytes—of raw data. As computers grow in speed, number-crunching capability, and memory, scientific researchers edge into data overload as they try to find meaningful ways to interpret data sets holding more information than the U.S. Library of Congress. Livermore's computer scientists are exploring techniques such as metadata, data-mining, and visualization to deal with the massive amounts of data.

(a) Results from Univac computations were spewed out as reams of numbers by a Remington-Rand typewriter modified to serve as an on-line printer. (b) Results from today's complex simulations are converted by powerful visualization software into three-dimensional detailed views, such as this one shown on the Livermore-developed PowerWall.

(a)

(b)

light), a machine 15 times faster than today's fastest supercomputers. "BlueGene/L would be a radical departure from previous machines," notes Mark Seager, program manager for ASCI Terascale Systems. BlueGene/L would use IBM's "system on a chip" based on commercial embedded-processor technology. Seager explains, "Embedded processors


A rendering of the Terascale Simulation Facility, which will house ASCI Purple, a machine capable of performing 60 trillion operations per second.

are optimized for low cost and low power and for usability in many configurations." McCoy notes that systems like BlueGene/L are the next big step in getting more performance at a lower price. "From ASCI Red to Purple, the systems use workstation processors targeted at the high-performance computing market. With BlueGene/L, we'd move from that curve to one using commodity PC processors. At the same time, we'd also move from using proprietary vendor software to open-source software such as the Linux operating system. These moves would result in considerably lower costs for the power we'd get—about $0.1 million per teraops for BlueGene/L, compare with White's $9 million per teraops or Purple's $3 million per teraops."

BlueGene/L would have 65,000 nodes or cells, 360 teraops—larger than the total computing power of the top 500 supercomputers in the world today—and between 16 and 32 terabytes of memory. "The questions facing us for BlueGene/L are: Can we build it? Can we write software for it? Can we write scientific simulations for it? We believe the answers are 'yes' to all," says Seager. Six times more powerful than ASCI Purple, BlueGene/L would open new
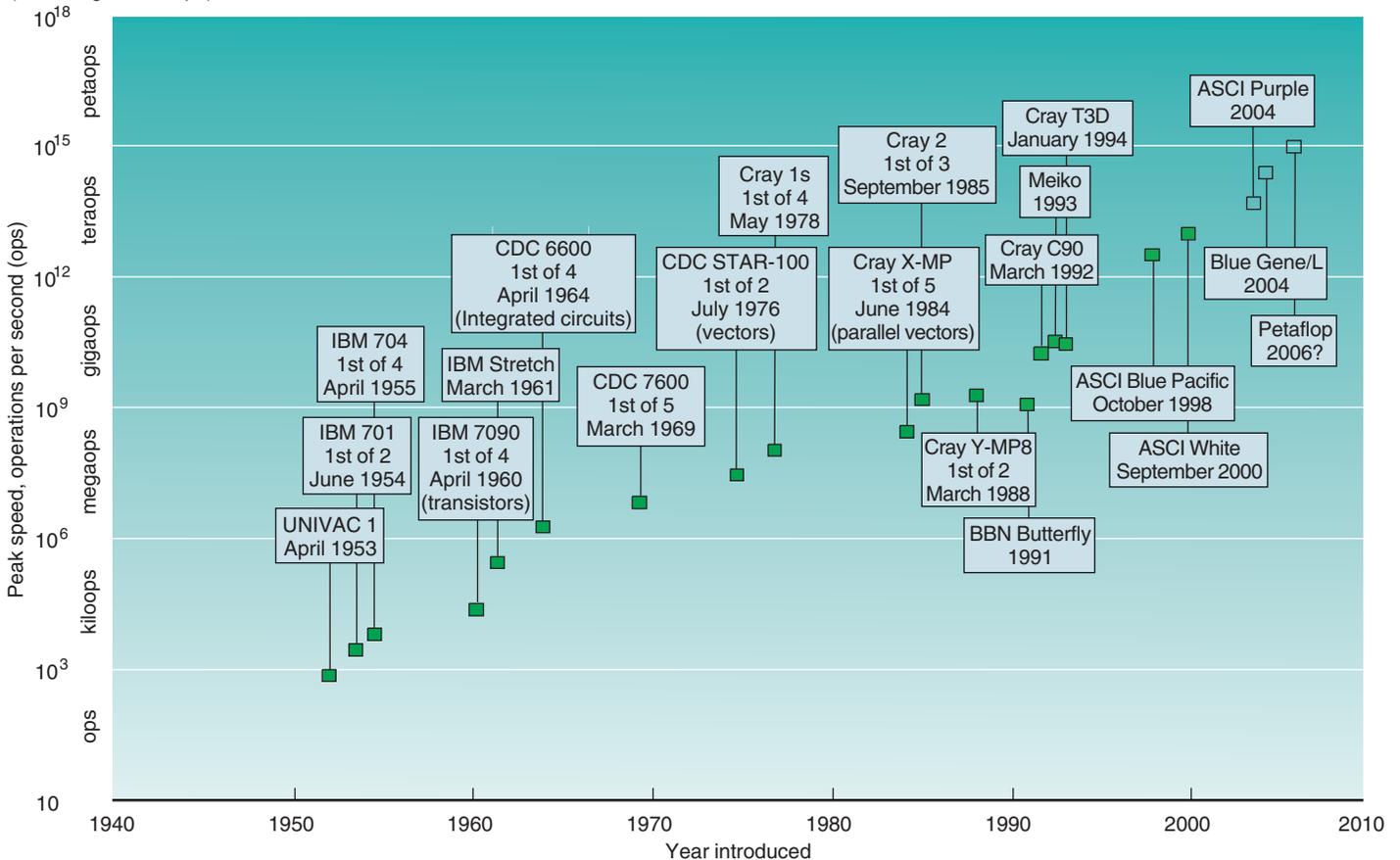
Timeline of Livermore's key supercomputers and their peak computing power.

From Personal Computers to Clusters

While supercomputers were always an integral part of Livermore's nuclear weapons design and stockpile stewardship efforts, other areas of the Laboratory also benefited from the computer revolution, particularly as computer systems became smaller, more powerful, and less expensive. In the 1970s, small microprocessor systems such as the PDP-11 began to be used in research tasks—digitizing oscilloscope traces, for example, and controlling experiments in chemistry labs. Then the personal computer, or PC, arrived, followed by more powerful microcomputers and workstations.

By the mid-1990s, many researchers in nonweapons areas were taking advantage of the relatively inexpensive and powerful desktop computers in their offices, or they used terminals tied to scientific workstations. Although having many advantages, these machines did not always have the necessary computational power, particularly for running three-dimensional simulations, which require the enormous computational horsepower of the latest generation of supercomputers.

Finally, in 1996, Livermore programs and researchers outside the stockpile stewardship effort gained access to unclassified Accelerated Strategic Computing Initiative–level terascale supercomputers through the Multiprogrammatic and Institutional Computing Initiative (M&IC). (See *S&TR*, October 2001, pp. 4–12.)

The M&IC acquired increasingly more powerful clusters, or groups, of computers such as the Compaq TeraCluster2000. As the Laboratory begins to celebrate its 50th year, Livermore researchers are at the forefront of simulating a wide range of physical phenomena in the unclassified arena, including the fundamental properties of materials, complex environmental processes, biological systems, and the evolution of stars and galaxies. Mike McCoy, deputy associate director for Integrated Computing and Communications, says, "Livermore Computing has become an institutional resource much like the library, a place where researchers from any program can expect resources to support their research."
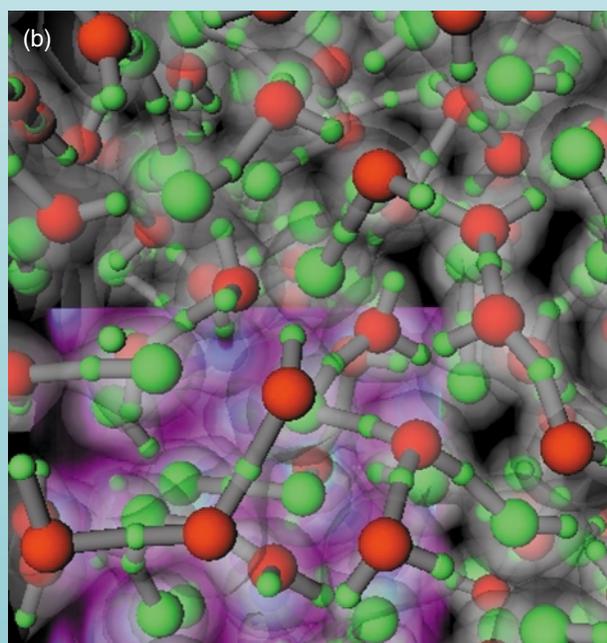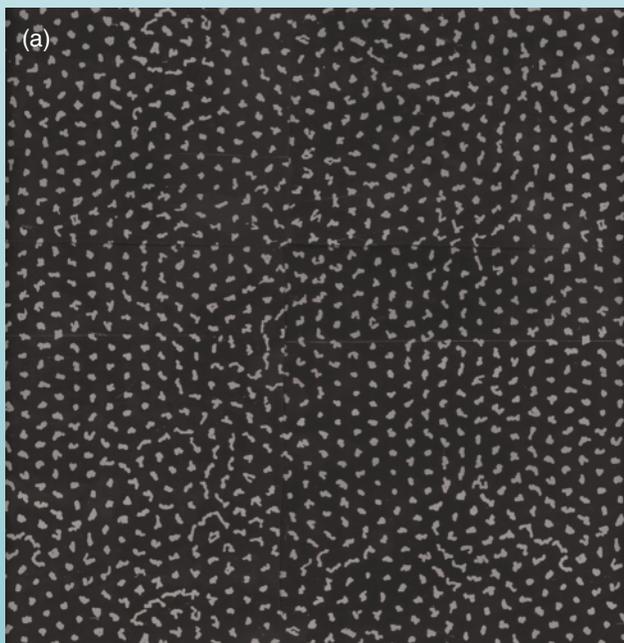
Particle tracking past and present contributes to a better understanding of the fundamental properties of materials. (a) In this example of Livermore physicist Berni Alder's pioneering computer simulation work, published in *Physics Review* in 1962, a simulation performed on the Livermore Advanced Research Computer supercomputer tracked 870 particles over time. (b) Recent work on the ASCI Blue Pacific includes this quantum-level simulation of a mixture of hydrogen fluoride and water molecules at high temperatures and pressures. The simulation tracked hundreds of atoms and thousands of electrons extremely accurately.

vistas in scientific simulation. "For instance," says Seager, "you begin to approach what you need to model complex biological systems. Having BlueGene/L would be like having an electron microscope when everyone else has optical microscopes, it's that much of a leap forward."

And after that? "Perhaps there will be computers that align DNA to do processing, or Josephson junction machines, or all-optical machines. Who knows what will happen in hardware, software, and information technology in the next 50 years," says Crawford. "Whatever innovation ends up driving the next era in computing will probably explode on the scene, much like the Internet did."

Fifty years ago, the birth of the electronic scientific computer ushered in a new era. Rather than having to accept crude approximations because the more exact equations were too difficult to solve, scientists could use the great speed and high accuracy of computers to simulate the phenomena they were trying to understand. Livermore researchers pushed the limits of each advanced machine, from using crude one-dimensional codes on the Univac and early IBM machines to complex three-dimensional codes on the current ASCI machines. Through ASCI and the coming generations of

supercomputing machines, another era appears on the horizon, an era in which enormously fast and powerful supercomputers will allow computer simulation to come into its own as a predictive science along with theory and experiment.

—Ann Parker

Key Words: Advanced Simulation and Computing (ASCI), ASCI BlueGene/L, ASCI Purple, ASCI White, computation history, Cray, IBM, Livermore Advanced Research Computer (LARC), supercomputer, Univac.

For further information, see the following Web sites on computation, past and present:

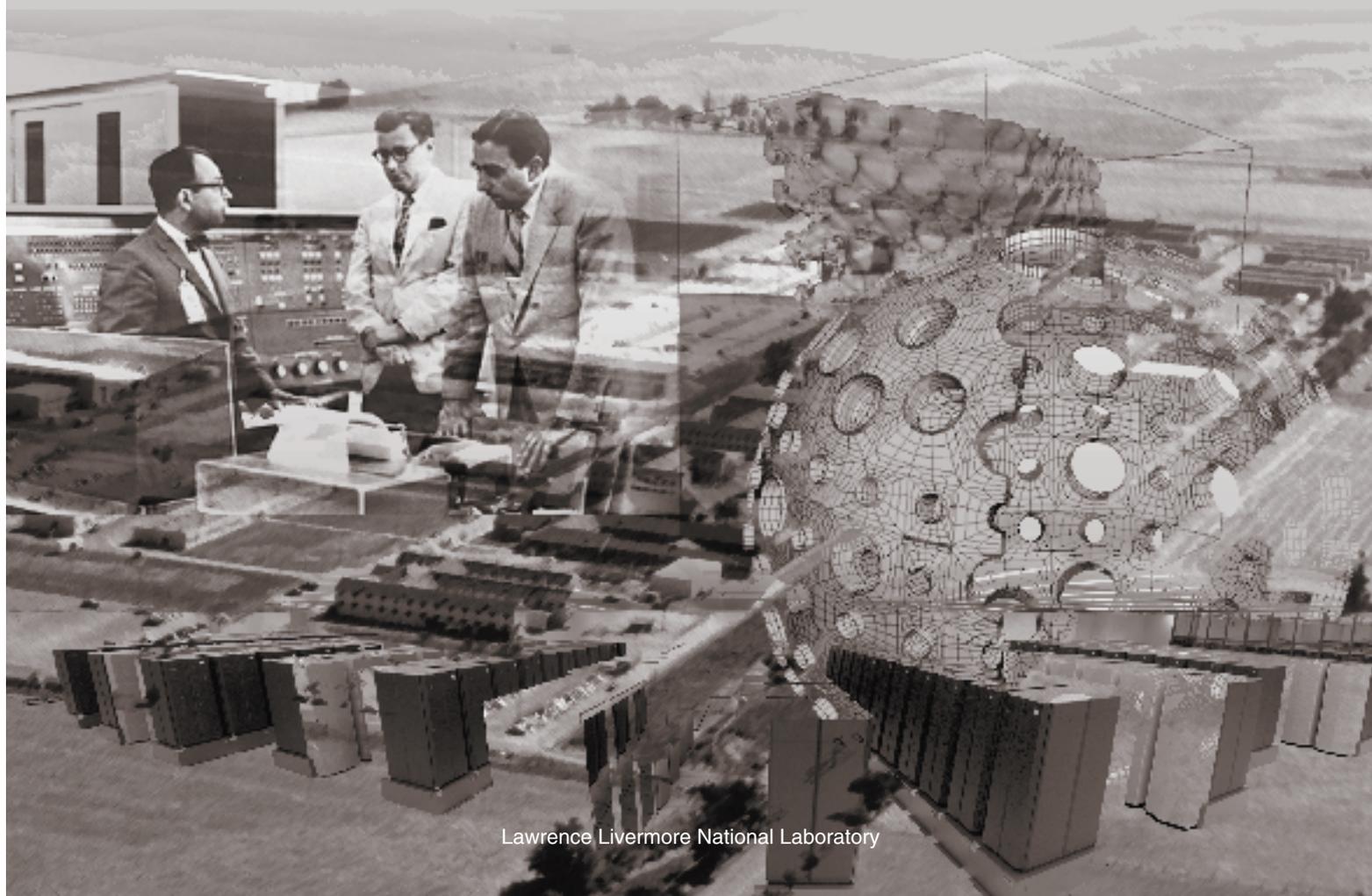**Computation at LLNL:**
www.llnl.gov/comp/

ASCI at LLNL:
www.llnl.gov/asci/

Oral History of Computation at LLNL:
www.nersc.gov/~deboni/Computer.history/

For further information about the Laboratory's 50th anniversary celebrations, see the following Web site:

www.llnl.gov/50th_anniv/