by Kurt Riesselmann

Computing the Quarks

A piece of steel may look cold and lifeless. But like any other piece of matter, it is bursting with activity deep inside. Electrons whiz around inside atoms, and a sea of never-resting quarks and gluons populates the nucleons that make up the atomic core.

For more than 30 years, researchers have tried to get a handle on how the quarks clump under the influence of the strong force to form protons and other constituents of matter. Although theorists discovered quantum chromodynamics (QCD)—the set of equations that describe the strong force—in the early 1970s, solving the equations has always been a struggle. The calculation of particle processes at low energy seemed impenetrable, and results produced by crude approximations didn't match experimenters' observations.

That has changed. During the last two years, the perseverance in taming the equations of QCD with supercomputing power has finally begun to pay off. Using ever-more-sophisticated computers and algorithms, theorists have begun to reproduce experimental results for QCD phenomena such as the decay of a quark-antiquark pair at rest. Moreover, in the last twelve months theorists took the lead and predicted physical quantities. To the theorists' delight, new experimental results then matched the theoretical numbers with equal precision.

"To do the necessary calculations has proved very challenging, and the only technique that has really succeeded is to directly 'discretize' the equations, feed them into a computer, and let the computer work very hard," explained Frank Wilczek when receiving the Nobel Prize in Physics in 2004, which we shared with David Gross and David Politzer. "In fact, these kinds of calculations have pushed the frontiers of massively parallel computer processing."

Quantum rubber bands

Gross, Politzer, and Wilczek discovered in 1973 that the strong force between two quarks is rather weak when the quarks are close together—a phenomenon called asymptotic freedom. But as the distance increases, the strong force—mediated by gluons—acts like a rubber band strong enough to prevent, say, a single quark from escaping the interior of a proton. Even when the rubber band "snaps," no single quark is produced. Instead, following Einstein's equation $E=mc^2$, the energy released creates two extra quarks that take the place of the free ends of the broken rubber band halves.

"Experimentalists don't observe free quarks," says Fermilab's Andreas Kronfeld, who is an expert in QCD calculations using computers. "To measure and understand quark properties, you need theorists to relate quarks to what is observed in experiments. The tool to use to get to the quark interactions is lattice QCD."

In quantum theories such as QCD, particles are represented by fields. To simulate the quark and gluon activities inside matter on a computer, theorists calculate the evolution of the fields on a four-dimensional lattice representing space and time. Using today's computers, a typical lattice simulation that approximates a volume containing a proton might use a grid of 24x24x24 points in space evaluated over a sequence of 48 points in time. The values at the intersections of the lattice approximate the local strength of quark fields. The links between the points simulate the rubber bands—the strength of the gluon fields that carry energy and other properties of the strong force through space and time, manipulating the quark fields.

At each step in time, the computer recalculates the field strengths at each point and link in space. In its simplest form, the algorithm for a single point takes into account the changing fields at the eight nearest-neighbor points, representing the exchange of gluons in three directions of space—up and down; left and right; front and back—and the change of the fields over time—past and future. Starting with random quark and gluon fields, theorists

The wiggly lines of Feynman diagrams illustrate the QCD processes between quarks and gluons that take place inside particles such as protons. A new generation of supercomputers begins to provide theorists with enough computing power to solve the equations of QCD with high precision.
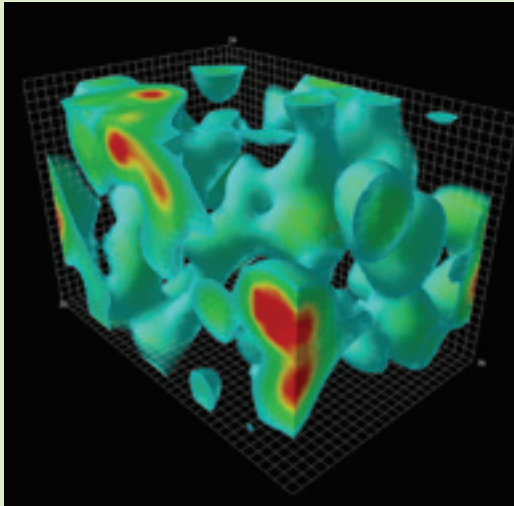
"This is what's happening all the time within us, we have these little lava lamps," said Frank Wilczek in his Nobel lecture in 2004 when he showed this QCD animation created by Derek Leinweber. The animation illustrates the fluctuations of the quark and gluon fields over time, revealing a lumpy structure that Leinweber dubbed the QCD lava lamp. A video is available at www.physics.adelaide.edu.au/theory/staff/leinweber/VisualQCD/Nobel/

produce hundreds of configurations that are the basis for calculating properties such as the masses of particles composed of quarks.

"Lattice QCD is a solid, predictive theory," says Kronfeld. "However, you have a finite but huge number of integrals to do. And that's what has taken us so long."

Teraflops for QCD

The computing power needed for these calculations is immense. Since the invention of lattice QCD by Kenneth Wilson in 1974, theoretical physicists have sought access to the most powerful computers, and physicists have actively contributed to the development of high-performance parallel computers. In the 1980s, the Cosmic Cube machines at Caltech, the Columbia University computers, the GF11 project at IBM, the Italian APE computers, the Fermilab ACPMAPS installation and the PACS machines in Japan were all designed and built out of the desire to simulate lattice QCD.

While the performance of the early machines was still measured in megaflops (one million floating point operations per second), the computing power reached tens of gigaflops in the early 1990s. Today's top-performance computers for the lattice QCD community—the QCD-on-a-chip (QCDOC) machines developed by a US-Japan-UK collaboration, and the next-generation machines of the Array Processor Experiment (apeNEXT) in Europe—operate in the multi-teraflops (millions of megaflops) range. A 1.2-teraflop machine, the SR8000F1 supercomputer at KEK, will be replaced with a more powerful machine in 2006. Using an alternative approach, scientists at Fermilab and Thomas Jefferson National Laboratory (JLab) hope to achieve over two teraflops for lattice QCD applications with large clusters of off-the-shelf PCs this spring.

The US lattice QCD program, which includes approximately 200 scientists, is supported by the Department of Energy and the National Science Foundation. The scientific activities of the program are coordinated by a national lattice QCD executive committee.

"Our job is to lead the efforts to construct the infrastructure," says committee chairman Bob Sugar, of the University of California, Santa Barbara. "We determine what hardware and software best meets the needs of the field. We draw plans, raise money, and see that the plans are properly implemented."

Three factors play the primary roles in determining the performance of a massively parallel computer: the rate at which the central processing unit (CPU) performs floating point operations (additions and multiplications); the

rate at which data is transferred between the processor's memory and its CPU; and the rate at which data is transferred between processors. The optimal balance among these three rates depends strongly on the problem being studied.

Commercial supercomputers, which must perform well on a wide variety of problems, are generally not optimal for any particular one. However, a computer designed for a single problem balances the CPU speed and data movement rates to optimize total performance and price-to-performance ratio. For this reason, lattice QCD scientists around the world increasingly build their own computers, obtaining significantly better price/performance than with commercial supercomputers.

The QCDOC and apeNEXT collaborations have taken the custom design of QCD machines to the next level. Scientists in these groups designed custom microchips that integrate a microprocessor unit, memory, and network communication interfaces as well as several megabytes of on-chip memory, an approach known as System-on-a-Chip (SoC) design.

"The communication of a microprocessor with memory is terribly complicated," says apeNEXT collaborator Hubert Simma, a physicist at the German laboratory DESY Zeuthen. "Normally, you use a separate microchip to handle this communication, which increases the time it takes to access data. By integrating the memory and networking interfaces on the same chip as the processor, you achieve much faster access."

The compact design of the QCDOC and apeNEXT machines reduces the physical size of these supercomputers and their consumption of electrical power, thereby reducing the need for cooling—a major concern for most commercial supercomputers. A typical supercomputer might consume many thousand kilowatts of power, producing the same amount of heat as tens of thousands of light bulbs. In contrast, a QCDOC machine, which has 12,288 microprocessors and a peak performance of 10 teraflops, requires only about 100 kilowatts of power, greatly reducing its cost of operation. The apeNEXT computers have racks with 512 microprocessors and a peak performance of 0.66 teraflops, and seven of 24 racks have been built so far. The racks will be distributed among INFN Rome (12), DESY Zeuthen (4), Universität Bielefeld (6) and Université de Paris-Sud (2).

Cooperating with industry

Designing and building QCD supercomputers often happens in close connection with industry. Japanese lattice QCD scientists, for example, have worked with the supercomputer company Hitachi in the past. And the microchips of the apeNEXT machines "represent one of the few cases of complete computing processors developed in Europe," says Raffaele Tripiccione, coordinator of the apeNEXT project. Its chip design is the basis for the 1-gigaflops-plus mAgic processor, produced by Atmel, a company with headquarters in California. In contrast to other high-performance chips, the mAgic architecture delivers its performance at a low clock frequency of 100 MHz, lowering power consumption and heat dissipation in the same way as the apeNEXT and QCDOC machines.

The QCDOC collaboration built on more than 20 years of experience in QCD supercomputer design at Columbia University, under the visionary leadership of Norman Christ; the collaboration worked closely with experts from IBM's Thomas J. Watson Research Center, which developed its own QCD machine, GF11, in the late 80s and early 90s. Other collaboration members came from Brookhaven National Laboratory, the RIKEN BNL Research Center, and the University of Edinburgh. The $15-million cost for the construction of three QCDOC machines was shared equally among the US Department of

Energy, RIKEN (the Institute Physical and Chemical Research in Japan), and the Particle Physics and Astrophysics Research Council (PPARC) in the United Kingdom. For the DOE computer, assigning computer time to various QCD projects is the task of the Lattice QCD Executive Committee.

No.1 in the TOP500

The exceptional price-to-performance ratio of the QCDOC project, which cost approximately $1 per megaflop, as well as the low power consumption of the QCDOC machines, make it the choice for the foundation of a new generation of IBM supercomputers known as BlueGene/L. To make the new super-computer more attractive for a broad class of applications, IBM generalized the massively parallel QCDOC architecture while retaining its cost and power-consumption advantages. The modest performance of the BlueGene SoC processors is offset by the fact that more than one thousand processors fit into a single rack without exceeding standard cooling capabilities.

The largest BlueGene/L machine built to date is a supercomputer with 65536 processors, which is currently listed as No.1 in the TOP500 list of supercomputers. It achieves a performance of 136.8 teraflops on the bench-mark used for the TOP500 listing, but the performance on QCD code would be significantly lower.

"BlueGene is the son of QCDOC," says Nicholas Samios with pride. Samios is the director of the RIKEN BNL Research Center, a RIKEN sister institute located at BNL. "IBM was involved in the R&D [of the QCDOC machine] from the very beginning."

In fact, the BlueGene supercomputer has retained enough of the QCDOC characteristics that the machine is still a viable QCD machine. According to Shoji Hashimoto, scientist at the Japanese particle physics laboratory KEK, his lab will begin operating a Blue Gene machine with a peak performance of 57.3 teraflops in March 2006, using it for lattice QCD and other applications. For QCD code, the machine is expected to sustain 12 to 15 teraflops.

The demand for more powerful QCD machines also persists in the United States. "We are having scheduling problems," says Samios about the RIKEN machine. "I think ten teraflops does not satisfy the appetite of these people. Our scientists are very imaginative. When they have more computing power, they find things to do that they couldn't do before."

PC clusters

The Department of Energy is also funding an alternative approach to providing the lattice QCD community with enough computing power at low cost. Instead of building special-purpose machines, scientists at Fermilab and JLab are taking advantage of powerful and low-cost off-the-shelf computer components, building large, sophisticated PC clusters dedicated to QCD simulations.

The idea goes back to the mid-90s, when Fermilab began using PC clusters for the computing-intensive analysis of large amounts of experimental data. In late 1997, Don Holmgren, who helped to set up the first PC clusters, was looking for additional applications. He approached Fermilab theorist Paul Mackenzie, who now serves on the Lattice QCD Executive Committee.

"We were trying to do this revolution," says Holmgren. "At the time, PCs and Linux machines were much more cost effective for the processing of the huge amount of data collected by the particle physics experiments. So we asked around and looked for other applications."

To test the idea, Holmgren and Mackenzie built the first small cluster for QCD applications in 1999. In 2001, the DOE Scientific Discovery through Advanced Computing (SciDAC) program began to provide funding to Fermilab and JLab to build clusters for the lattice QCD community. The clusters have

grown in size, and every year the labs buy the best PCs and networking technology, eventually discarding the old systems. Today, Fermilab operates three clusters with a total of 512 PCs with Intel Xeon and Pentium processors, and Myrinet and Infiniband networking technologies. JLab owns three clusters with a total of 768 Intel Xeon processors using Myrinet and gigabit ethernet mesh networks. Together, the Fermilab and JLab clusters sustain about 1.3 teraflops on QCD code. In connection with the purpose-built QCD machines, PC clusters provide theorists with a range of computing power for different types of QCD calculations.

For PC clusters, communication among processors is as important as for QCD supercomputers. For a prototype system built in 2000, Fermilab spent as much money on the communication network as on the 80 PCs themselves. Although PC clusters are limited in size due to power consumption, space constraints, and scalability of off-the-shelf components, they are a cost-effective solution that allows for annual upgrades with the latest PC technology. Over the next four years, the DOE Lattice Computing project will provide $9.2 million, most of which will be used to expand the clusters.

"For many years the Holy Grail was one dollar per megaflop," says Holmgren. "In 2000, our cluster was operating at a cost of about 15 dollars per megaflop. This year, we are about there."

Progress in software

Having enough computing power is only one ingredient in advancing QCD calculations. To efficiently harvest the enormous amounts of computing power now available to the lattice QCD community, scientists also need to optimize the software and the scientific algorithms used in their QCD calculations.

"One tends to emphasize the hardware," says Sugar. "But the software is really important for these difficult-to-handle machines. The code has to be adapted to the machines."

Ideally, a lattice QCD application should be able to run both on a cluster system and a QCDOC machine. To achieve this goal, the DOE SciDAC program has provided funding for software development since 2001. The lattice QCD community has designed and implemented a standard QCD Applications Program Interface that provides a uniform programming environment on a variety of computer systems, from PC clusters to supercomputers.

The interface has allowed theorists to focus on optimizing their scientific algorithms and methods. With more computing power now available, they have done away with their crudest approximations and have begun to produce attention-grabbing results.

"With today's theoretical methods, we still need more computing power," Sugar says. "The two things go hand in hand. That's always going to be the case."

A QCDOC daughter card with two processing nodes and two 128 MB memory modules.

Photo: Norman Christ, Columbia University