no. 2-0033

TUCK SCHOOL OF BUSINESS AT DARTMOUTH
WILLIAM F. ACHTMEYER
CENTER FOR
GLOBAL LEADERSHIP

Computing at the Speed of Life

IBM stumbled in the early 1990s and Silicon Valley rose soon thereafter. The upheaval created a forceful perception: that everything innovative in information technology was spawned in the small, high-growth firms in the Valley, while IBM, on the other hand, was oversized and decrepit, an icon of a past age.

Thousands of IBMers knew that it was untrue. Nonetheless, the perception made it difficult for the company to attract the most talented computer scientists and for IBM salespeople to do their job. Selling information technology to corporations was a lengthy and complex process, and companies in the industry were notorious for making exaggerated claims and obfuscating their wares' real capabilities. Baffled by complexity and cloudiness, buyers were swayed by perceptions. Many believed the buzz: that IBM, while reliable and secure, was not cutting edge. Reversing this perception became a priority at IBM's highest levels.

One of many angles the company explored for addressing this priority was dominating the esoteric but periodically high-profile arena of high-performance computing. A few scientists—meteorologists, experts in global warming, security and intelligence analysts, particle physicists, geneticists—relied on so-called supercomputers. In their efforts to push the frontiers of their disciplines, each was constrained by the speed of the world's fastest computers.

The domain of high-performance computing was remote to most people. Nonetheless, if IBM could play a role in pushing the frontiers of science, it would be in a position to make strong, symbolic statements about its prowess at innovation.

Lou Gerstner, then CEO, understood the power of symbols. In 1997, IBM garnered tremendous public attention by pitting man versus machine in a chess match. Deep Blue, an IBM computer, defeated reigning world chess champion Gary Kasparov. By 1999, it was time to make a new, even more powerful statement.

Biotechnology was by far the most compelling field of scientific inquiry that year. Scientists were closing in on a complete map of the human genome, and in the process they unleashed tremendous volumes of new knowledge about how the body worked at the molecular level. Scientists anticipated revolutionary new medical therapies.

This case was written by Professor Chris Trimble of the Tuck School of Business at Dartmouth. The case was based on research sponsored by the William F. Achtmeyer Center for Global Leadership. It was written for class discussion and not to illustrate effective or ineffective management practices.

To order additional copies please call: 603-646-4055

Embedded in human DNA are recipes for tens of thousands of proteins, the large and complex molecules that are at the core of biological functioning. Proteins are constructed from smaller molecules, known as amino acids, of which there are 20 varieties. The code in DNA prescribes the *sequence* of amino acids in a protein. The complete mapping of the human genome in 2000 was a momentous milestone in biotechnology. However, amino acid sequence does not unlock all of a protein's secrets. A full description of a protein also includes the three-dimensional *structure* of the protein. Therefore, scientists were eager to also understand the process by which proteins took or altered shape, known as *protein folding*. Misfolded proteins presaged biological malfunction.[1] In fact, scientists believed protein misfolding was an important step in the biological mechanisms that led to such diseases as Alzheimer's, cystic fibrosis, and mad cow.[2]

Scientists had "discovered" the structure of less than 10 percent of human proteins—the simpler ones—through complex laboratory techniques, among them X-ray crystallography and nuclear magnetic resonance spectroscopy. The discovery process was difficult and time consuming; it could take up to a full year to work out just one structure. Thus, scientists looked for faster alternatives, such as computer simulation of protein folding.[3] As of 1999, however, such a simulation was beyond the capabilities of even the world's fastest computers.

Researchers at IBM were well aware of the problem. They entered an international event[4] in which participants built simulations to predict protein structure using only the amino acid sequence as an input. Outcomes were judged by comparing simulation results to actual structures of known proteins. IBM frequently sought such opportunities to benchmark its capabilities. Dr. William Pulleyblank, then head of mathematical sciences at IBM Research, explained,

> We always want to be on the supercomputing frontier, and that requires being tied in with users of supercomputers and their problems. Molecular biology was clearly going to go in a direction that would require tremendous computational intensity, and we wanted to be there.

There were multiple methods for simulating protein folding. The most direct approach involved simulating the behavior of each individual atom, but that approach most quickly exceeded the practical limits of the supercomputers available at the time.

[1] In fact, a change in protein shape can alter its physical properties dramatically. Hard-boiling an egg is a process of unfolding and then refolding proteins into a different shape.
[2] Alzheimer's patients often have a hard, gunky substance in the affected areas of their brains, formed when proteins misfold.
[3] Another computationally intensive approach was estimating protein structure by correlating the sequence of the protein in question with that of proteins with known structures. This approach was less computer-resources intensive than the simulation approach, but many believed the simulation technique would prove more accurate.
[4] The biennial event was known as CASP, the Critical Assessment of Techniques for Protein Structure Prediction.

This would all change if the speed of supercomputers could be increased by a factor of 10, 100, or more. That desire fit nicely with the ambitions of Dr. Ambuj Goyal, head of computer science within IBM Research. Dr. Goyal was a high-performance computing veteran who had been heavily involved in designing the IBM SP series, IBM's most powerful machines at that time. He felt it was time to put a stop to research that churned out reliable, incremental improvements in the speed of supercomputers without changing their fundamental design. In his view, IBM Research should be dedicated to fundamental science, not product development.

Dr. Goyal started talking about building the first "petaflop" computer. Computer speeds were measured in "flops": floating point operations per second. In 1999, the world's fastest computers operated at speeds of a few teraflops—that is, one trillion flops. A petaflop, one million billion flops, was 1,000 times faster than a teraflop. The IBM Research team estimated a petaflop computer theoretically would be capable of simulating the folding of a protein in about one year, roughly the same amount of time as it took scientists in laboratories. (What was not known was whether the simulations would result in accurate predictions of the protein structure. Nobody had ever had the computing power necessary to perform the experiment.)

There was no way to build a petaflop computer following the existing supercomputer design paradigm. For years, advances in computational power came primarily through increases in processor speed and sophistication. Unfortunately, better processors consumed more power. According to Dr. Goyal, a petaflop computer based on existing design concepts would not only occupy the space of several football fields, it would require a dedicated power plant to supply its electricity. The electric bill would quickly exceed the cost of the computer. It just was not plausible.

IBM Research pursued a different design concept, massive parallelism, believing that the company could build the petaflop computer while actually *reducing* the speed and complexity of the microprocessors. The major leap in performance would come from linking tens of thousands of microprocessors to work in concert. At best, what could be created was a plausibility argument as to why this might work.

Because a plausibility argument was a long way from a working design, the research staff was hesitant to make their ambition public. Nonetheless, Mr. Gerstner and his PR staff seized the opportunity to make a major announcement. IBM held a press conference in November 1999 to proclaim the breakthrough development of a new supercomputer capable of advancing biology and medicine. The company announced it would invest $100 million over five years in a new machine, to be dubbed BlueGene. *The New York Times* published an article under the headline "IBM Plans a Supercomputer that Works at the Speed of Life." The article described BlueGene's ambition as one that dwarfed Deep Blue, announced a goal of increasing the speed of supercomputers by a factor of 500, and noted that only a company of IBM's scale could assemble the necessary expertise in computers, mathematics, biology, chemistry, and physics to build such a machine.

The $100 million investment required commitment from Director of Research Dr. Paul Horn and augmentation by corporate funds and outside funding. Dr. Goyal explained how the company estimated BlueGene's cost and development time:

> We had some experience to get us in the right ballpark. A $20 million investment seemed uninteresting, and $200 million was too much. Five years seemed about right. We knew we could do it in 50 and that 2 was impossible.

IBM expected the investment would signal a redoubled commitment to research and, in doing so, would help recruit the world's greatest minds. Furthermore, the company anticipated the effort to build BlueGene would lead to new fundamental insights in computing that would steadily trickle down to computer designs with broader commercial uses; that had been the company's experience in the past. The Deep Blue project had paved the way for new concepts in networking, databases, and scalable computing.

To a group of researchers at IBM, the BlueGene project was suddenly very real.

Building BlueGene

Dr. Paul Coteus was one of the first IBM researchers to become involved in the BlueGene effort. At the time of the announcement, Dr. Coteus had been working on developing methods for higher-speed communications between microprocessors. Seeking funding, he engaged contacts at a U.S. government laboratory, the Lawrence Livermore National Laboratory (LLNL), in a conversation about his work. The lab was a heavy consumer of supercomputers. One of its responsibilities was certifying the country's nuclear weapons stockpile. Nuclear weapons decayed over a period of years. Advances in supercomputing enabled the lab to simulate the decay process as well as nuclear explosions and, ultimately, to certify the operability of the weapons without test-firing them.

Dr. Coteus's pitch to LLNL was based on the expectation that an increase in communications capacity could lead to a discontinuous leap in supercomputing performance. The laboratory declined initially, but a few weeks after the November 1999 BlueGene announcement, Dr. Coteus received a surprise call. Not only was the lab interested in financing the communications project, it was interested in funding the development of the entire new supercomputer. Speculating on the motivations behind the move, Dr. Coteus noted that government labs like LLNL competed with each other for prestige. Involvement in a project to design BlueGene could elevate the lab's importance on the national research scene.

During the lengthy and complex negotiations between IBM and the laboratory, IBM acquired some funding to support the earliest stages of the BlueGene effort through shorter term contracts with other government entities. Ultimately, LLNL agreed to fund more than half of BlueGene's anticipated development costs and received the right to purchase the first BlueGene, while IBM retained the patents. To get the contract, IBM had to agree to hit a series of technical targets and milestones along the five-year development path, including a target for computer speed at a time when the company had not even

designed the chip that would reside inside BlueGene. Dr. Pulleyblank recalled the moment IBM made the commitment:

> IBM built all kinds of exit clauses into the contract, but that did not help us relax. After all, what would happen if the president of the United States called and asked for his computer? We were nervous, but it is often when you make these commitments that you do your best work.

Dave Turek, another leader in the BlueGene development effort, commented on the contract:

> The seriousness of what the lab was up to, certifying the nuclear weapons stockpile, left little room for error. We had worked with the lab for several years and always treated the nature of their work with paramount seriousness. BlueGene had to be flawless.

That seemed anything but guaranteed at the time. One of the IBM researchers' biggest concerns was the reliability of microprocessors, which periodically failed. In single-processor machines, this amounted to an occasional unexpected need to restart. BlueGene, however, was to be a machine composed of tens of thousands of microprocessors. On the basis of typical microprocessor failure rates, BlueGene could be expected to fail as often as once per day. Given the nature of the applications that would run on BlueGene, so frequent a need to restart would render BlueGene impractical. As of the signing of the contract, IBM had a few ideas about how to build a "self-healing" machine, one that would isolate failed microprocessors without interrupting the operation of the remaining chips. But those ideas had not been proven or tested.

Another concern was whether it was possible to build networks connecting the chips that operated at a sufficiently rapid pace. Theoretically, it was possible, but it had never been done before. Early design steps involved building simulations of the computer itself, but there remained doubt about what would happen when the machine was turned on for the first time. According to one computer developer on the project, beyond just a few hundred processors, IBM would be moving into unknown territory.

Assembling the Team

A global team of IBM researchers and developers tackled the BlueGene challenge. The nucleus of the team was a group of research scientists located in IBM's Thomas J. Watson Research Center in Yorktown, N.Y., about 35 miles north of New York City. The research team included experts in materials science, semiconductors, computer architecture, and computer science—some of IBM's most respected scientists. The lure of working on a breakthrough supercomputer design, the size of the budget, and the project's high public profile proved irresistible for many.

During 2000, while the company was finalizing the contract with LLNL, IBM reorganized its research division. Dr. Mark Dean was appointed vice president of systems and assumed

responsibility for the project. Dr. Dean subsequently asked Dr. Pulleyblank to lead the project day to day, nearly a full-time responsibility.

The scope of the project was so great that people from many parts of IBM Research contributed, and not all of them formally reported to Dr. Pulleyblank. Individual researchers at IBM had substantial discretion over which projects they dedicated their energies to. Their accomplishments in the categories of papers, patents, and commercial projects were evaluated over three-year periods on the basis of peer recognition, including recognition from external peers at other research institutions. In selecting which research proposals to pursue, IBM sought to find the right balance between long-term fundamental research and nearer-term work on practical applications. Another mechanism for ensuring researchers' work did not stray too far from practical applications was hiring select teams of scientists who were experts not in computing but in the problems that computers were used to solve. To many such scientists, the opportunity to work for IBM was intriguing. Dr. Coteus elaborated:

> I'm a particle physicist by training, and that's hardly unusual around here. Lots of people in computationally intensive sciences end up as interested in the computing as they were in their initial scientific discipline.

IBM typically hired talented researchers just out of their doctoral or master's programs and attempted to retain them for life. The researchers were employed at seven research centers in different countries so that the company could attract the best talent anywhere in the world.

In Dr. Goyal's recollection, a tremendous number of ideas were submitted in the early stages. Several dozen researchers, each with the hope of presenting the strongest design ideas and then leading some aspect of the project, immersed themselves in the challenge day and night. There was substantial competition and conflict.

To ensure the project could be completed in the tight five-year time frame, IBM took the unusual step of tackling many of the steps of the project in parallel, designing hardware, software, and applications concurrently. Because of the size of the team, that demanded rich communication and coordination. However, according to Dr. Pulleyblank, it had the potential to speed up the process. Dr. Coteus elaborated:

> In a traditional development project with hundreds of engineers involved, none of them really knows how the whole thing works. With BlueGene, we started with a clean sheet, and the core design team understood the entire project.

The biggest expense within the research group was the researchers' salaries. Unusual for a research project, the BlueGene effort also required a major capital expenditure of several million dollars to modify the Watson Research Center facilities to accommodate the electrical equipment and chilled water piping that would power and cool the computer. "It seemed like a huge sum to spend all at once," recalled Dr. Pulleyblank.

A second critical group, which got involved toward the end of 2002, was IBM's Engineering & Technology Services division, located in Rochester, N.Y. Steve Lewis, director of engineering services, coordinated a far-flung team that pushed the project forward from the research team's initial designs to fully developed designs. The team also coordinated the massive process of manufacturing the first BlueGene and delivering it to LLNL. Mr. Lewis reflected,

> BlueGene was built across the IBM matrix. I had fewer than 10 direct reports, but there were more than 100 people involved from different parts of IBM. In a way, our division served as the general contractors that found all of the skilled and experienced people within the company who put BlueGene together. They were in different locations for each component of the machine, from Canada to India to New York and more. The fact that we were working on such a high-profile project created genuine excitement. People were eager to work with us.

However, the BlueGene project did pull many people away from their day-to-day responsibilities for existing brands, inevitably leading to some conflicts. Business units that gave up some of their talented people's time were compensated through the BlueGene budget, but that did not always mean business leaders were willing to give BlueGene priority, especially toward the end of quarters when the urgency to hit short-term targets was highest. At times, the direct support of some of IBM's most senior leaders, including Dr. Horn and Nick Donofrio, executive vice president of innovation and technology, was critical.

IBM evaluated progress on the project primarily on the basis of budget versus actual expenditures and against the milestones in the LLNL contract. Measuring progress against budget was not an exact science because so many of the researchers involved were also involved in other projects. According to one research leader, "Anyone can hit a budget. It's not the money, it's what you accomplish that matters."

BlueGene's design involved creating two specialized, application-specific integrated circuits (chips), one for computing and the other for linking adjacent processors. Designing and building the chips was the longest lead-time piece of any computer design process. It took until June 2003, almost four years into the five-year window, before the team had a test chip up and running.[5] The chip had a unique design, but its speed was noteworthy for being unremarkable. The chip ran at only 700 MHz, a processor speed available on reasonably priced laptop computers that same year.

It was clear by that point that a petaflop computer was not within reach in the five-year window. The team would defer that goal until the second-generation BlueGene design.

[5] The chip design was a so-called "system on a chip" design. Each node in BlueGene was self-contained and had its own memory registers embedded in it. The technology for including processing and memory on the same chip was only a few years old. The processor also operated on a reduced instruction set, which minimized functionality but increased speed.

There was another target, however, that motivated the team. In June 2002, Japanese multinational NEC Corporation had overthrown IBM as the builder of the world's fastest computer. Its so-called Earth Simulator zipped along at just shy of 36 teraflops. Some politicians in Washington, D.C., were deeply concerned that the U.S. had given up its edge in advanced computing.

Design plans for the first BlueGene called for a 64,000-chip machine, enough to deliver roughly one-third of a petaflop of computing power, some 10 times the speed of the Earth Simulator *if* BlueGene worked as designed. But IBM also knew that several rivals, including Silicon Graphics, Inc. in California, were working on new supercomputers of their own. In November 2004, the end of IBM's five-year window, an industry organization[6] would release a new list of the world's fastest supercomputers. The updated list would be announced at Supercomputing 2004, a major industry conference. Would IBM lead the list? Would Silicon Graphics? Would NEC retain its position? The race was on.

IBM proceeded to build a rapid succession of prototypes, doubling the number of processors each time. Dr. Pulleyblank recalled the atmosphere of the next year:

> We got through 1, 2, 4, 8—no problem. And when we got to 512, it felt like we'd reached a tremendous milestone. Some of our networking experts did not think we could get even that far. And then someone said to me, "Bill, you're only at 512. How are you possibly going to get to 64,000?" Nobody had ever built so large a machine.

With each doubling of the number of chips, there was a risk that the machine would produce far less than a doubling in its total speed. If the communications network bogged down, like a highway during rush hour, BlueGene would spend all its time trying to send messages to itself, and calculations would grind to a halt.

The team pushed straight through 512 chips to 1,000, 2,000, 4,000, and then, in September 2004, BlueGene was running with over 8,000 chips and just *over* 36 teraflops. BlueGene was the fastest computer in the world.

IBM debated how to publicize the achievement. Some people at the company wanted to send out a press release immediately; others wanted to keep quiet until the new list came out in November.

IBM decided to go public right away. With the ensuing press releases, IBM gained recognition for being the first U.S. firm to outpace NEC, if only by fewer than one teraflop. Meanwhile, the IBM team anticipated that *if* Silicon Graphics was also on the cusp of dethroning the Earth Simulator, the announcement might actually be cause for celebration for the Silicon Graphics team since they might anticipate they'd take the top

[6] Twice per year, in June and November, the "official" list of the 500 fastest computers was updated and published at www.top500.org.

spot in November. Indeed, at Supercomputing 2004, the firm announced its Columbia machine had achieved 52 teraflops.

IBM stole the show, however. Just in time, the BlueGene team doubled capacity again. BlueGene was operational with 16,000 chips blazing at more than 70 teraflops and consumed remarkably little electrical power. It was a tremendously satisfying moment for the IBM team and a second PR coup in only three months.

Commercializing BlueGene

At the same time, IBM decided to up the ante by preparing BlueGene for commercial launch. Commercialization was hardly a forgone conclusion. Because IBM had made giant strides in supercomputer design while developing BlueGene, some involved in the project felt that it would have been acceptable if LLNL's BlueGene was the only one ever built. Commercialization would require major new investments in marketing, selling, and software. It would also require an investment in ongoing support, and because BlueGene was such new technology, the cost of that support was difficult to estimate. Mr. Turek, who had a long history in high-performance computing and would be closely involved in the commercialization effort, elaborated:

> For a while, we discussed ways to write sales terms and conditions that would limit our exposure to servicing costs. But I had learned in an earlier product launch that there was no way for IBM to do that. Customers don't read terms and conditions. They look at the IBM logo, and that stands for reliability. When customers call, we have to respond.

Another consideration was the likelihood that BlueGene would dampen sales of IBM's existing lines of high-performance computers. Overall, commercialization was a big risk.

On the other hand, IBM's newly acquired position leading the top 500 list, plus its overall dominance of the list, put the company in a position of influence. IBM anticipated serving on more policy committees in Washington, giving the company a head start on government sales. Furthermore, there was plenty of advance interest. In fact, Argonne National Laboratory, a U.S. Department of Energy research facility, had convened a BlueGene future-users conference in 2002, before the machine even existed. Scientists representing a wide range of disciplines attended, eager to gain access to the new machine.

Leo Suarez, IBM's head of deep computing, led the effort to commercialize, coordinating the efforts of developers, marketers, and sales and support teams, many of whom had dual reporting lines to both Mr. Suarez and their functional organizations. Mr. Lewis and IBM's Engineering & Technology Services division would have a renewed role in the effort, this time to build not one machine but a network of suppliers and a system for routine manufacture and installation of BlueGene computers.

The high-performance computing market was well known to IBM. BlueGene was intended for the very highest end of the market. Mr. Suarez could easily count the

organizations that would be interested in BlueGene. The early potential buyers—government and research institutions—numbered no more than a few dozen.

The market was not homogeneous, however. Users would run a wide variety of applications. Thus, one of the first steps in the commercialization effort was to explore the types of applications that BlueGene could run effectively. Not every application was suited for a massively parallel machine.

IBM started by approaching the most likely users. There was no such thing as a quick supercomputer sale, however. Potential buyers needed to be convinced that the new machine could solve their problems, and that took time.

Typically, on any team of scientists in a computationally intensive field, there were experts in supercomputers who could figure out how to structure their scientific problems in the most efficient way for a given supercomputer. Because BlueGene had a revolutionary architecture, the task was not routine. Potential BlueGene users would have to figure out how their mathematical models and simulations could be broken into small pieces and parceled out to thousands of microprocessors. Conceptually, writing a BlueGene application was like writing small programs for each of thousands of computers and specifying the information to be passed between them.

Potential BlueGene purchasers typically had made tremendous investments in developing software applications, and they were very interested in the level of effort required to "port" their existing software to BlueGene. Some applications could be reconfigured more readily than others to take advantage of a massively parallel computer like BlueGene. But in every case, porting involved risks. There was never a guarantee that an application would work well on a new machine, so buyers always had an incentive to simply stick with their existing computers.

To reduce the strain of porting, the IBM design team had decided early that BlueGene would be compatible with the Linux operating system and would use a standard protocol for passing messages between processors.[7] Linux was commonly used, and the underlying source code was freely available. (After decades of building closed-standard systems, IBM had fully embraced open standards by the late 1990s, and BlueGene was no exception.) Although some rewriting of code was generally necessary, it would be much less intense for users of Linux machines. Often, to convince potential buyers that BlueGene would work on their applications, IBM helped them port one of their applications and "test drive" it in IBM's "on-demand center." This could take several months for complex applications.

IBM also had to convince potential and existing customers that the company intended to commit to the new line of supercomputers for the long haul. Customers wanted to see product road maps to be sure there would still be a staff of BlueGene experts within IBM

[7] Argonne National Laboratory had already developed an open-standard model, known as Message Passing Interface (MPI), for creating parallel programs and was already using it in multiprocessor machines, such as IBM's own SP series.

10 years down the road. The Linux operating system and the fact that BlueGene was designed with commonplace components wherever possible helped build customers' confidence.

Once convinced of BlueGene's value, scientists would have to sell the investment to their organization. Buyers in government and academic institutions typically faced thorny, bureaucratic purchasing processes. They were required to float open requests for proposals and often specified in advance the amount they would pay for a certain computer speed.

Even after IBM made the sale, installing a BlueGene was anything but an overnight affair; it could take a few months. Thus, depending on the complexity and size of the system being ordered, the amount of time from initiation of a sales conversation to recognition of revenue on IBM's income statement could exceed a year. Mr. Suarez reflected,

> This is no business for anyone wanting to get rich quick. It is a very difficult business to make a return in unless you have a lot of science and research behind you. There are very few companies in the game at all.

By 2007, there were nearly 30 BlueGene machines in operation in government and research institutions. The owners, the most sophisticated supercomputer users in the world, were self-sufficient.

IBM had seen that users of less powerful supercomputers generally required more support and more help porting applications. Nonetheless, Mr. Suarez, now with a track record and a supportive user base behind him, was ready to develop new customer segments, particularly in the corporate sector.

He believed an investment in BlueGene could pay a quick return in many industries—in finance and insurance, to run risk analytics; in nuclear power, to prove the potential to extend the life of commercial reactors; in manufacturing, to optimize supply chains and inventory management; in transportation, to optimize scheduling; in petroleum, to identify opportunities to extend the life of existing wells; in entertainment, to improve digital media; in aviation, to model airflows around aircraft; in healthcare, to improve medical imaging; in automotive safety, to simulate crashes. Mr. Suarez hoped that by demonstrating the economics of owning a BlueGene to these users, he'd succeed in raising prices and margins, escaping from the early BlueGene users' mind-set that you paid a certain amount for each teraflop.

IBM had extensive relationships with IT leaders in many of the corporations that potentially could benefit from BlueGene. With academic and government research institutions, a supercomputer was the flagship technology purchase. Within corporations, it was a much smaller piece of a wide-ranging inventory of computing technologies and applications. Dave Jursik was the IBM sales executive responsible for high-performance computing sales, including BlueGene and older IBM supercomputer lines. He described the general approach to serving corporations:

> We have a client-centric model. The sales teams with the greatest seniority are business generalists, focused on customer sets. For example, we have a group focused on the automotive industry that understands the full range of information technology applications used by the industry. Sales specialists who focus on a specific IBM hardware platform or software application support these client-focused teams. BlueGene is now one more tool that client-facing teams can call upon to best serve clients.

Despite the wide range of possibilities, expanding the market was difficult. To evaluate the use of BlueGene or to develop or redevelop applications for it, corporations had to assemble the right team of experts, including those with expertise in business, supercomputers, and scientific or mathematical fields relevant to particular industries. It was outside the scope of normal operations. Furthermore, some applications, such as car crash simulations, could not be adapted easily to a massively parallel architecture; it was a lot of work. Mr. Jursik and the sales team emphasized that massively parallel designs were going to become much more commonplace as the space and power consumption of traditionally designed high-performance computing platforms made them impractical. In other words, the cost of adapting to the new architecture could be delayed but not avoided.

To accelerate adoptions, IBM endeavored to reposition its offering, in part by eliminating use of the term "high-performance computing." The phrase seemed to convey "exotic, difficult, and really expensive," according to Mr. Turek. IBM replaced the phrase with "business-performance computing" and "deep computing," harkening back to Deep Blue and conveying the notion of enabling deep insight. New marketing materials focused more on applications and value to buyers than on the machines' specifications.

Mr. Suarez partnered with Mr. Jursik's team for BlueGene sales. Because BlueGene was not part of an existing brand structure, IBM developed a new compensation plan for BlueGene specialists. It was attractive to the sales force; a single BlueGene sale could make a salesperson's quarter. IBM did not want salespeople driven solely by quick commissions, however. Mr. Suarez and Mr. Turek sought patient salespeople whose focus was helping customers solve their problems, not pushing products on them.

By 2007, IBM had installed BlueGene in two corporate settings. An electric utility was using the machine to better estimate the lifespan of its nuclear reactors, and a supplier of technology to the aerospace industry was using BlueGene in computational fluid dynamics applications. Mr. Jursik expected a petroleum company to be the next corporate adopter.

Results

Mr. Suarez evaluated his progress based on traditional financial metrics—gross margins, operating income, internal rate of return, etc.—against projections and against a standard hardware hurdle rate of return that the company applied to all hardware systems. Each BlueGene sale offered the promise of lucrative follow-on sales. IBM's services group offered ongoing maintenance and service contracts, and other IBM product divisions

offered more storage, more front-end servers, and so forth. However, these follow-on sales had no impact on Mr. Suarez's P&L; they accrued to other IBM divisions. IBM also made standard cost allocations for shared services across every division.

At the initiation of the commercialization effort, Mr. Suarez and his team reviewed all possible customers and assigned probabilities that they would buy. That resulted in projections that Mr. Suarez felt could be reasonably assigned an error band of 30 to 40 percent. While progress was steady, results through the end of 2006 fell short of projections. Mr. Turek reflected,

> Everything we see in the market is understandable. A lot of interested parties have a wait-and-see attitude, perhaps waiting for version two.

The company was pleased with the way BlueGene operated. According to Mr. Suarez, customers loved the stability—the machine seemed to run forever—and urged IBM to keep scaling the existing design so they could easily port their applications to ever-faster machines. Mr. Jursik elaborated:

> We have implemented BlueGenes only in environments where our clients will be successful. As a result, our clients are satisfied.

When pushed on the question of BlueGene's return on investment, Dr. Goyal responded with a question of his own:

> BlueGene has attracted a number of the world's most talented graduate students to IBM. What's the return on investment of that?

BlueGene dominated the top 500 list through June 2007, and IBM planned to stay on top. Mr. Suarez added,

> The business has iconic value. It also drives basic research and development, and the knowledge gained trickles down to other development efforts. It's the same reason that automotive companies compete in the Indy 500.

Not every idea explored during the design effort succeeded. An idea to simultaneously send and receive data between two silicon chips over the same wire failed, for example. However, many ideas did succeed. IBM researchers believed the advances the company had made in developing low-power, extremely reliable "self-healing" supercomputers were crucial breakthroughs for the future of parallel computing. Mr. Lewis elaborated:

> Parallelization is becoming the dominant question in computer design because the rate of improvement of single chips is slowing down quickly. IBM is in the lead because of BlueGene. Lots of organizations are developing computer language enhancements that are better suited to parallel computing.

Despite the challenges of the market, Mr. Suarez savored his position:

> This is one of the hardest and yet most interesting jobs I've had in my career at IBM, with the richest opportunities to do true human-benefits science.

Mr. Jursik felt a similar motivation:

> It is a thrill to work on really important problems—"holy grail" type problems—helping to get results that not long before were impossible to get.

BlueGene was being put to work on a wide range of problems. Ironically, however, scientists had made only limited progress on the problem that inspired the name of the world's fastest computer: protein folding. The teams working on that problem still struggled, both for want of new theoretical insights[8] and for lack of sufficient computational power.

In June 2007, IBM unveiled its second-generation BlueGene design[9] and confidently predicted that the first customer installation of a petaflop computer would take place in 2008.

[8] The protein-folding problem was theoretically complex. There were situations wherein proteins did not fold into the most thermodynamically ideal state, and that limited the effectiveness of atom-by-atom simulation. Also, the actions of "chaperone proteins," which helped other proteins fold, were not fully understood.
[9] The new machine doubled the number of processors per chip to four, operated at a high frequency, had a higher memory capacity in each chip, and had a faster network.