5

DESIGNING DATA-CENTRIC COMPUTERS

All most people know about Watson's hardware is what they saw on TV: the machine's cool, purple façade, the avatar that stood between the human contestants on *Jeopardy!* and Watson's quirky robotic voice. But there's more to the machine than its TV persona. If you were to take a tour inside the original Watson computer at the IBM Research lab in Yorktown Heights, N.Y., you'd better understand what made its performance on *Jeopardy!* possible—and why computers must be redesigned for the new era of cognitive computing.

The computer takes up a corner of a datacenter on the lab's second floor. Two parallel rows of refrigerator-sized metal racks are loaded with ninety-two servers. A door at one end turns the space between the racks into a room within a room. At the time of the *Jeopardy!* contest, the microprocessors were running twice as fast as the next-fastest general purpose computer on the planet.[1] Air-conditioning fans and smaller fans within the servers create a roar that's nearly deafening. The microprocessors give off an immense amount of heat. And that's not a good thing. All the heat and wind signals just how much energy Watson consumes—85,000 watts of electricity when it's running at full throttle. That's enough to keep the lights on in a small town. In comparison, the human brain consumes just 20 watts of energy. Unless we can make computers many orders of magnitude more energy efficient, we're not going to be able to use them extensively as our intelligent assistants. Computing intelligence will be too costly to be practical. Scientists at IBM Research believe that to make computing sustainable in the era of big data, we will need a different kind of machine—the data-centric computer.

Today's computers are processor-centric. The microprocessor, which is the central processing unit in the von Neumann architecture, is where much of the action happens in computing. Working hand in hand with the operating system, the microprocessor sends out instructions to various components within the computer, requesting data from where it's stored, including memory chips and disk drives. If the computer is part of a larger network, the processor fetches data from storage systems located out on the network. A new system design is needed to greatly reduce the amount of movement required. The new design puts the data at the center, rather than the microprocessor. This design will accomplish several beneficial things at once, according to Bijan Davari, an IBM fellow who operates at the cutting edge of system design. Machines will perform computations faster, make sense of large amounts of data, and be more energy efficient.

The race to make ever more capable computers began in earnest in the 1970s, when engineers began packing more and more transistors on a single chip. In the computer industry, the term "scaling" refers to ability to add more resources, for example, more electronic circuitry and storage capacity. The process of packing more transistors on a chip is called "scaling down." In addition, over the years, as businesses and governments needed ever more data-processing power, computer companies built server computers that were designed specifically to handle large and complex tasks. They pursued two strategies. One was to combine more processing power and other resources within a single machine—called "scaling up." The IBM mainframe is an example of this approach. Another is "scaling out," which is stringing many servers together so they behave like a single large system. Examples of this are today's supercomputers and Google's vast server farms.[2]

IBM scientists believe an essential element of data-centric computers will be a new kind of scaling: "scaling in." Engineers will integrate memory with logic and messaging in dense, three-dimensional packages. Today's memory chips are single layers of silicon with large amounts of data-storing circuitry packed into them. They're used to store data temporarily and serve them up to the microprocessor when requested. The chips take up space in a computer and require a lot of data movement. But what if you were to stack memory chips like so many pancakes? They'd take up much less space, and much less data movement would be required. That's exactly what

IBM and other computer industry leaders are doing with a concept called the hybrid memory cube.

This will be an amazing little device. In a single cube, you will be able to stack many layers of memory chips. The layers are cemented together and connected electronically by tiny holes that extend vertically through the stack and are filled with copper. A layer of logic circuitry on the bottom of the cube exposes the memory resources of the entire cube for use by the microprocessor. The design will require 90 percent less space than today's memory chips and 70 percent less energy.[3] In time, memory cubes will likely have their own processors attached to them. Logic and memory will be one, greatly reducing the von Neumann bottleneck.

This technology will have a profound effect on the way computing is done. Giant supercomputers, used for everything from oil exploration to crash tests in the auto industry, will shrink in size and will use much less energy. The cubes will be used in next-generation server farms to greatly reduce the amount of space and electricity required. Further down the road, the technology will make it possible to pack an immense amount of computing resources into smartphones, tablets, and other portable devices.[4]

A second key element of a data-centric computer will be the way processing is done within the system. Today, a computer is equipped with a microprocessor, or CPU, that performs all or most of the computation. All of the data that the CPU acts upon must be transported from where they are stored to the microprocessor then back again. In some cases, data are transported hundreds or even thousands of miles. A data-centric computer would distribute processor chips throughout the system to greatly reduce the movement of data.

This concept is being tried out today on specialized computing systems that are used to analyze huge volumes of data. Murali Ramanathan, a neuroscientist at the State University of New York at Buffalo, is in charge of one such project. His research focuses on spelling out the role of genetic and environmental factors in multiple sclerosis. MS is a cruel disease that most often strikes young adults. The body's own immune system attacks the brain and spinal cord, causing physical disabilities and cognitive problems. The cause isn't known, and there is no cure. Murali's goal is to find a cure for the disease—or discover a way to prevent it in the first place. What makes his analysis so challenging is the fact that there are so many potential combinations of genes that could contribute to MS and, when genetic information is combined with environmental factors such as diet and smoking, the number of possible causes increases exponentially.

This is a data-intensive problem. Problems like these can't be addressed effectively with conventional supercomputing, where massive parallelism is the key concept. A machine breaks the problem into small chunks and farms them out to thousands of microprocessors to address in parallel, then combines the results of those calculations to derive an answer. To deal with tasks like Murali's, it makes sense to meld parallelism with the strategy of moving some of the processing closer to the data.[5]

Murali and his colleagues acquired a new machine specifically designed for data-intensive analysis. The refrigerator-size machine uses specialized processor chips, called field-programmable gate arrays, to filter the data on storage disks before passing along only the relevant pieces of information to the main processors. The chips can be custom programmed after they're in the machine to handle certain kinds of queries, and, as a result, they reject as much as 90 percent of the data. This implies that the system performs some of the analysis as the data are moving off the disks, rather than handling all of it on the main processor or, in some cases, shipping it off to specialized analytics computers. As a result, the data do not travel as far, and less energy is required to move bits around. The work can be done faster and more efficiently.[6]

To get an idea of how effective this kind of computing can be, consider the results that Murali and his colleagues are achieving. Shortly after they installed the new computer, they ran a problem that contained 100,000 genetic and environmental variables. That means there were nearly 5 billion two-predictor combinations and 100 trillion three-predictor combinations to be evaluated. Yet it took just eleven minutes to do the analysis—compared to twenty-seven hours on their conventional computer. Murali was sitting at his desk on the SUNY campus when he received an e-mail containing the summary results. "I was very, very excited. This can open the door for us," he says. "We're on the cusp of being able to solve problems that were previously unsolvable."

No matter where the data are crunched in a computing system, it will be increasingly essential to move data around more quickly and efficiently. This is the third key dimension of designing data-centric computers. The most promising means of moving data faster is by harnessing photonics, the generation, transmission, and processing of light waves. Already, telephone companies and broadband communications providers use pulses of light traveling via fiber-optic cables to speed the transmission of data, voice communications, and multimedia content such as movies and TV shows over long distances. In addition, fiber-optic technologies move data within buildings and between racks of machines in computer datacenters. Next, photonics will be used to connect servers in a single rack and chips on a circuit board. Ultimately, the technology will be used within the chips themselves.

In a decade-long initiative, a team of scientists at IBM Research has achieved breakthroughs in nanophotonics that could greatly improve the performance of all kinds of computers. Their newest technology, called CMOS integrated nanophotonics, combines traditional electrical components with optical components on a single silicon chip. In today's advanced computers, optical elements are placed on separate chips, which impedes the performance of the computers and adds to costs. By combining optical and electrical elements on a single chip, the researchers hope to make it possible for manufacturers to produce large quantities of such devices through production processes that are already in use. Their work has now advanced beyond the pure science stage into product development, with a goal of reaching the market within the next few years. This melding of light and electricity will help computers move data around with a speed and energy efficiency that is impossible using today's technologies.

SCENARIO: ADDING INTELLIGENCE TO THE ELECTRIC GRID

Today's "smart grids" aren't actually very bright. Most electric utility systems that claim the "smart" label use advanced metering systems that enable customers to monitor their electricity use and enable two-way communications between the meter and the utility provider. The systems make it possible for utility companies to set pricing that varies by the season and the time of day, providing discounts to customers who shift energy use to off-peak periods. Useful, yet there's so much more that can be done to make utility grids smarter. One of the most important elements will be distributing intelligence throughout the grid by moving a lot of analysis to where the data are generated, that is, moving the data processing to the data.

A utility grid is a complex system that connects sources of energy generation through a distribution network with users of energy, both commercial and residential. Grids are becoming more complex as more sources of energy are added to the mix, including wind and solar; as more consumers produce their own energy; and as new uses of electricity spread, such as the charging of electric vehicles. Sophisticated information-technology networks need to be built to monitor and manage the grids, better balancing supply and demand, coordinating transactions among the diverse parties involved, and optimizing the grid so it's both efficient and resilient.

Potentially, electrical grids could be managed by large, integrated computing systems—cognitive systems—that stretch from the home appliances to wind farms off the coast. Today's grids are equipped with many sensors that measure not just generation, transmission, and consumption but more detailed information such as problems with particular transformers or power lines. Data collected from equipment scattered throughout the system are transmitted to central computers to be recorded and analyzed. But, in the future, with more data pouring in from sophisticated sensors across the entire network, this approach will require more energy, bandwidth, and storage than is practical. Some of the analysis will have to move to the data.

In a decade or so, intelligent devices capable of data analysis will be embedded in every piece of equipment and appliance connected to the grid. These devices will collect detailed information on energy consumption and help us manage our energy use. Intelligent software agents installed in the devices will control how appliances function, based initially on preset parameters. For example, in a family home, an air-conditioning system might be set up to reduce the temperature beginning one hour before the family members arrive home at the end of the day. But the system might learn from motion sensors that one resident is coming back earlier and working in the home office. It may begin to cool earlier—but just in the part of the house that person frequents. The intelligent agent could also poll family members, perhaps via their smartphones, to see how satisfied they are with the heating and cooling setup, and, based on their reactions, make modifications.

In another scenario, several families in a neighborhood might switch to electric vehicles. Standard practice would be to charge the vehicles' batteries at night, but if multiple households charge their vehicles at the same time it will overtax the neighborhood transformer. To avoid power outages, intelligent devices in the transformer and homes will detect sudden increases in demand for vehicle charging and will coordinate with one another to schedule charging sessions for the various EVs at different times during the night.[7]

Putting intelligence in the network has the potential to transform the way energy is generated and used. This will make it easier for new sources of energy to be integrated into the system. It will provide strong incentives for conservation, hasten repairs after power outages, and make it possible for societies to expand electricity services without needing to make massive investments in new generation plants. When all of this comes about, smart grids will begin to truly deserve the name.

JOURNEY OF DISCOVERY: RETHINKING HOW COMPUTERS ARE DESIGNED

How much does it cost to move a single bit of data from point of origin to point of computation? Apparently, until recently, nobody bothered to try to answer that question—perhaps because it didn't seem consequential. But now, in the era of big data, computer scientists see that when you move a massive amount of data over long distances—or even over short distances many times—the costs quickly add up. So that's why David Turek, IBM's vice president of exascale computing, asked the question of a roomful of IBM computing systems experts in 2011. "Fortunately, they were willing to listen to a stupid question," he says. "In the data world, it's becoming a big issue, and we have to come up with a way of quantifying it."[8]

That simple question helped focus an initiative at IBM that's aimed at fundamentally re-architecting computing systems. It's the most radical rethinking of how computers should be designed since IBM's introduction of its System 360 family of computers in the early 1960s. The project is an example of the new culture of innovation, demonstrating a willingness to abandon old ways of seeing things and start clean. The initiative is called Data Centric Deep Computing, or DC^2.

IBM fellow Jim Kahle is responsible for piloting DC^2—and for answering Dave Turek's question. He has played leading roles in some of IBM's most significant advances in chip design over the past three decades. They include the Cell processor, a chip that was employed in the Sony PlayStation game console, the IBM Roadrunner supercomputer and advanced digital TV sets. Yet the Cell chip did not become as widely used throughout computing as Jim had hoped. With DC^2, he dreams of creating a revolutionary design that, ultimately, permeates the entire computing world, starting with high-performance computing and filtering out to everything else.[9]

Jim was born in Venezuela, where his father worked in the oil business, and the family later moved to Indonesia before Jim came back to the United States to study computer science at Rice University. Jim credits his childhood experience with diverse cultures with preparing him to run large and complex engineering projects where close collaboration is essential. (At its peak, the Cell development team was 450 people strong and included engineers not just from IBM but from Sony and Toshiba.) Jim reached out to people with a wide range of expertise to make sure new ideas came into the discussion. He's doing the same thing now with DC^2.

In the early months of the DC^2 project, Jim convened large gatherings of experts on silicon semiconductors, photonics, memory technologies, computer architecture, software, mathematics, and analytics. They began by describing to one another the new uses for computing that they see coming in the next decade and then explained how they're preparing for them. Based on that input, Jim is now mapping a new computer systems architecture for the cognitive era and a process for achieving it. "We have to figure out how to reorganize ourselves, which is not easy culturally. We have to steer the boat in another direction," Jim says.

He and his crew haven't yet finished the discovery phase, but, along the way, they have developed some firm ideas about the principles that will guide them to the end of the journey. The cornerstone of their thinking is the belief that there must be a shift from processor-centric computing to data-centric computing. Jim remembers the day in late 2011 when the idea of data-centric computing first came to light. He and about a dozen scientists and engineers had gathered in conference room 36-002 of the IBM Research lab in Yorktown Heights—the DC^2 war room. They were stumped. They had a lot of great technology ideas but no unifying principle. Then team member Ravi Nair, arriving late to the meeting, rushed into the room and showed his colleagues a simple drawing he had made. On the left of the page was a graphic depiction of traditional computing showing how data move across all the levels of the data-storage hierarchy, from disk and tape storage to various types of memory and finally to the processor. On the right, he showed how processing could take place at each level of the hierarchy and said energy use would decrease as a result. Jim suddenly saw the light. The team would turn computing inside out. The brains in a computer wouldn't be in one place anymore. They'd be all over the place. The processing would move to the data.

They had their big insight, but how would it play out in the detailed design of the computer? To get to that, Jim had to answer the more basic question Dave Turek had asked him a few weeks earlier: How much *does* it cost to move a bit of data? Jim went off and did his homework and came up with some useful comparisons. For instance, it costs anywhere from 800 to 6,000 picocents (one picocent is a billionth of a penny) to transfer a single bit through a network, depending on how far the bit travels. In contrast, it costs roughly from 6 to 27 picocents to move data within a single computer.[10] So it's 200 to 1,000 times more costly to move data a long distance than it is to perform the processing next to the place the data are stored.

The fact that the costs of data transfer are so variable in different situations means that Jim and his team had to carefully evaluate each kind of transfer with a cost-benefit analysis. That led them to some interesting conclusions. They decided that specialized processors attached to components of the computing system would be assigned to perform particular tasks. They'd do a slice of the processing—the part that is optimally done next to the data in memory, on disks, or on tape drives or at the point where data flow into the system. In other cases, depending on the task at hand, most of the data may be moved to the main processors.

Now the specialists on Jim's team are going off to engineer new components, and the software architects have to design the enabling software. All this has to work together. Designing a data-centric computer will take several years. Ultimately, Jim and his team hope to produce computing systems that recognize the tasks that are being assigned to them and automatically parse out the various processing tasks to different chips similar to the way a music composer assigns notes and percussive sounds to different instruments in an orchestra.

It will take a tremendous effort to redesign computers to take on big data, but the potential payoff is great. It's possible that a decade or so from now the power of today's fastest supercomputer will fit in a box on your desk and tomorrow's fastest computers will be performing tasks we can't even imagine today.