



COVER STORY December 13, 2007, 5:00PM EST

Google and the Wisdom of Clouds

A lofty new strategy aims to put incredible computing power in the hands of many

by [Stephen Baker](#)

One simple question. That's all it took for Christophe Bisciglia to bewilder confident job applicants at Google ([GOOG](#)). Bisciglia, an angular 27-year-old senior software engineer with long wavy hair, wanted to see if these undergrads were ready to think like Googlers. "Tell me," he'd say, "what would you do if you had 1,000 times more data?"

What a strange idea. If they returned to their school projects and were foolish enough to cram formulas with a thousand times more details about shopping or maps or—heaven forbid—with video files, they'd slow their college servers to a crawl.

At that point in the interview, Bisciglia would explain his question. To thrive at Google, he told them, they would have to learn to work—and to dream—on a vastly larger scale. He described Google's globe-spanning network of computers. Yes, they answered search queries instantly. But together they also blitzed through mountains of data, looking for answers or intelligence faster than any machine on earth. Most of this hardware wasn't on the Google campus. It was just out there, somewhere on earth, whirring away in big refrigerated data centers. Folks at Google called it "the cloud." And one challenge of programming at Google was to leverage that cloud—to push it to do things that would overwhelm lesser machines. New hires at Google, Bisciglia says, usually take a few months to get used to this scale. "Then one day, you see someone suggest a wild job that needs a few thousand machines, and you say: Hey, he gets it."

What recruits needed, Bisciglia eventually decided, was advance training. So one autumn day a year ago, when he ran into Google CEO Eric E. Schmidt between meetings, he floated an idea. He would use his 20% time, the allotment Googlers have for independent projects, to launch a course. It would introduce students at his alma mater, the University of Washington, to programming at the scale of a cloud. Call it Google 101. Schmidt liked the plan. Over the following months, Bisciglia's Google 101 would evolve and grow. It would eventually lead to an ambitious partnership with IBM ([IBM](#)), announced in October, to plug universities around the world into Google-like computing clouds.

As this concept spreads, it promises to expand Google's footprint in industry far beyond search, media, and advertising, leading the giant into scientific research and perhaps into new businesses. In the process Google could become, in a sense, the world's primary computer.

"I had originally thought [Bisciglia] was going to work on education, which was fine," Schmidt says late one recent afternoon at Google headquarters. "Nine months later, he comes out with this new [cloud] strategy, which was completely unexpected." The idea, as it developed, was to deliver to students, researchers, and entrepreneurs the immense power of Google-style computing, either via Google's machines or others offering the same service.

What is Google's cloud? It's a network made of hundreds of thousands, or by some estimates 1 million, cheap servers, each not much more powerful than the PCs we have in our homes. It stores staggering amounts of data, including

numerous copies of the World Wide Web. This makes search faster, helping ferret out answers to billions of queries in a fraction of a second. Unlike many traditional supercomputers, Google's system never ages. When its individual pieces die, usually after about three years, engineers pluck them out and replace them with new, faster boxes. This means the cloud regenerates as it grows, almost like a living thing.

A move towards clouds signals a fundamental shift in how we handle information. At the most basic level, it's the computing equivalent of the evolution in electricity a century ago when farms and businesses shut down their own generators and bought power instead from efficient industrial utilities. Google executives had long envisioned and prepared for this change. Cloud computing, with Google's machinery at the very center, fit neatly into the company's grand vision, established a decade ago by founders Sergey Brin and Larry Page: "to organize the world's information and make it universally accessible." Bisciglia's idea opened a pathway toward this future. "Maybe he had it in his brain and didn't tell me," Schmidt says. "I didn't realize he was going to try to change the way computer scientists thought about computing. That's a much more ambitious goal."

ONE-WAY STREET

For small companies and entrepreneurs, clouds mean opportunity—a leveling of the playing field in the most data-intensive forms of computing. To date, only a select group of cloud-wielding Internet giants has had the resources to scoop up huge masses of information and build businesses upon it. Our words, pictures, clicks, and searches are the raw material for this industry. But it has been largely a one-way street. Humanity emits the data, and a handful of companies—the likes of Google, Yahoo! ([YHOO](#)), or Amazon.com ([AMZN](#))—transform the info into insights, services, and, ultimately, revenue.

This status quo is already starting to change. In the past year, Amazon has opened up its own networks of computers to paying customers, initiating new players, large and small, to cloud computing. Some users simply park their massive databases with Amazon. Others use Amazon's computers to mine data or create Web services. In November, Yahoo opened up a cluster of computers—a small cloud—for researchers at Carnegie Mellon University. And Microsoft ([MSFT](#)) has deepened its ties to communities of scientific researchers by providing them access to its own server farms. As these clouds grow, says Frank Gens, senior analyst at market research firm IDC, "A whole new community of Web startups will have access to these machines. It's like they're planting Google seeds." Many such startups will emerge in science and medicine, as data-crunching laboratories searching for new materials and drugs set up shop in the clouds.

For clouds to reach their potential, they should be nearly as easy to program and navigate as the Web. This, say analysts, should open up growing markets for cloud search and software tools—a natural business for Google and its competitors. Schmidt won't say how much of its own capacity Google will offer to outsiders, or under what conditions or at what prices. "Typically, we like to start with free," he says, adding that power users "should probably bear some of the costs." And how big will these clouds grow? "There's no limit," Schmidt says. As this strategy unfolds, more people are starting to see that Google is poised to become a dominant force in the next stage of computing. "Google aspires to be a large portion of the cloud, or a cloud that you would interact with every day," the CEO says. The business plan? For now, Google remains rooted in its core business, which gushes with advertising revenue. The cloud initiative is barely a blip in terms of investment. It hovers in the distance, large and hazy and still hard to piece together, but bristling with possibilities.

Changing the nature of computing and scientific research wasn't at the top of Bisciglia's agenda the day he collared Schmidt. What he really wanted, he says, was to go back to school. Unlike many of his colleagues at Google, a place teeming with PhDs, Bisciglia was snatched up by the company as soon as he graduated from the University of

Washington, or U-Dub, as nearly everyone calls it. He'd never been a grad student. He ached for a break from his daily routines at Google—the 10-hour workdays building search algorithms in his cube in Building 44, the long commutes on Google buses from the apartment he shared with three roomies in San Francisco's Duboce Triangle. He wanted to return to Seattle, if only for one day a week, and work with his professor and mentor, Ed Lazowska. "I had an itch to teach," he says.

He didn't think twice before vaulting over the org chart and batting around his idea directly with the CEO. Bisciglia and Schmidt had known each other for years. Shortly after landing at Google five years ago as a 22-year-old programmer, Bisciglia worked in a cube across from the CEO's office. He'd wander in, he says, drawn in part by the model airplanes that reminded him of his mother's work as a United Airlines ([UAUA](#)) hostess. Naturally he talked with the soft-spoken, professorial CEO about computing. It was almost like college. And even after Bisciglia moved to other buildings, the two stayed in touch. ("He's never too hard to track down, and he's incredible about returning e-mails," Bisciglia says.)

On the day they first discussed Google 101, Schmidt offered one nugget of advice: Narrow down the project to something Bisciglia could have up and running in two months. "I actually didn't care what he did," Schmidt recalls. But he wanted the young engineer to get feedback in a hurry. Even if Bisciglia failed, he says, "he's smart, and he'd learn from it."

To launch Google 101, Bisciglia had to replicate the dynamics and a bit of the magic of Google's cloud—but without tapping into the cloud itself or revealing its deepest secrets. These secrets fuel endless speculation among computer scientists. But Google keeps much under cover. This immense computer, after all, runs the company. It automatically handles search, places ads, churns through e-mails. The computer does the work, and thousands of Google engineers, including Bisciglia, merely service the machine. They teach the system new tricks or find new markets for it to invade. And they add on new clusters—four new data centers this year alone, at an average cost of \$600 million apiece.

In building this machine, Google, so famous for search, is poised to take on a new role in the computer industry. Not so many years ago scientists and researchers looked to national laboratories for the cutting-edge research on computing. Now, says Daniel Frye, vice-president of open systems development at IBM, "Google is doing the work that 10 years ago would have gone on in a national lab."

How was Bisciglia going to give students access to this machine? The easiest option would have been to plug his class directly into the Google computer. But the company wasn't about to let students loose in a machine loaded with proprietary software, brimming with personal data, and running a \$10.6 billion business. So Bisciglia shopped for an affordable cluster of 40 computers. He placed the order, then set about figuring out how to pay for the servers. While the vendor was wiring the computers together, Bisciglia alerted a couple of Google managers that a bill was coming. Then he "kind of sent the expense report up the chain, and no one said no." He adds one of his favorite sayings: "It's far easier to beg for forgiveness than to ask for permission." ("If you're interested in someone who strictly follows the rules, Christophe's not your guy," says Lazowska, who refers to the cluster as "a gift from heaven.")

A FRENETIC LEARNER

On Nov. 10, 2006, the rack of computers appeared at U-Dub's Computer Science building. Bisciglia and a couple of tech administrators had to figure out how to hoist the 1-ton rack up four stories into the server room. They eventually made it, and then prepared for the start of classes, in January.

Bisciglia's mother, Brenda, says her son seemed marked for an unusual path from the start. He didn't speak until age

2, and then started with sentences. One of his first came as they were driving near their home in Gig Harbor, Wash. A bug flew in the open window, and a voice came from the car seat in back: "Mommy, there's something artificial in my mouth."

At school, the boy's endless questions and frenetic learning pace exasperated teachers. His parents, seeing him sad and frustrated, pulled him out and home-schooled him for three years. Bisciglia says he missed the company of kids during that time but developed as an entrepreneur. He had a passion for Icelandic horses and as an adolescent went into business raising them. Once, says his father, Jim, they drove far north into Manitoba and bought horses, without much idea about how to transport the animals back home. "The whole trip was like a scene from one of Chevy Chase's movies," he says. Christophe learned about computers developing Web pages for his horse sales and his father's luxury-cruise business. And after concluding that computers promised a brighter future than animal husbandry, he went off to U-Dub and signed up for as many math, physics, and computer courses as he could.

In late 2006, as he shuttled between the Googleplex and Seattle preparing for Google 101, Bisciglia used his entrepreneurial skills to piece together a sprawling team of volunteers. He worked with college interns to develop the curriculum, and he dragooned a couple of Google colleagues from the nearby Kirkland (Wash.) facility to use some of their 20% time to help him teach it. Following Schmidt's advice, Bisciglia worked to focus Google 101 on something students could learn quickly. "I was like, what's the one thing I could teach them in two months that would be useful and really important?" he recalls. His answer was "MapReduce."

Bisciglia adores MapReduce, the software at the heart of Google computing. While the company's famous search algorithms provide the intelligence for each search, MapReduce delivers the speed and industrial heft. It divides each task into hundreds, or even thousands, of tasks, and distributes them to legions of computers. In a fraction of a second, as each one comes back with its nugget of information, MapReduce quickly assembles the responses into an answer. Other programs do the same job. But MapReduce is faster and appears able to handle near limitless work. When the subject comes up, Bisciglia rhapsodizes. "I remember graduating, coming to Google, learning about MapReduce, and really just changing the way I thought about computer science and everything," he says. He calls it "a very simple, elegant model." It was developed by another Washington alumnus, Jeffrey Dean. By returning to U-Dub and teaching MapReduce, Bisciglia would be returning this software "and this way of thinking" back to its roots.

There was only one obstacle. MapReduce was anchored securely inside Google's machine—and it was not for outside consumption, even if the subject was Google 101. The company did share some information about it, though, to feed an open-source version of MapReduce called Hadoop. The idea was that, without divulging its crown jewel, Google could push for its standard to become the architecture of cloud computing.

The team that developed Hadoop belonged to a company, Nutch, that got acquired. Oddly, they were now working within the walls of Yahoo, which was counting on the MapReduce offspring to give its own computers a touch of Google magic. Hadoop remained open source, though, which meant the Google team could adapt it and install it for free on the U-Dub cluster.

Students rushed to sign up for Google 101 as soon as it appeared in the winter-semester syllabus. In the beginning, Bisciglia and his Google colleagues tried teaching. But in time they handed over the job to professional educators at U-Dub. "Their delivery is a lot clearer," Bisciglia says. Within weeks the students were learning how to configure their work for Google machines and designing ambitious Web-scale projects, from cataloguing the edits on Wikipedia to crawling the Internet to identify spam. Through the spring of 2007, as word about the course spread to other universities, departments elsewhere started asking for Google 101.

Many were dying for cloud knowhow and computing power—especially for scientific research. In practically every field, scientists were grappling with vast piles of new data issuing from a host of sensors, analytic equipment, and ever-finer measuring tools. Patterns in these troves could point to new medicines and therapies, new forms of clean energy. They could help predict earthquakes. But most scientists lacked the machinery to store and sift through these digital El Dorados. "We're drowning in data," said Jeannette Wing, assistant director of the National Science Foundation.

BIG BLUE LARGESSE

The hunger for Google computing put Bisciglia in a predicament. He had been fortunate to push through the order for the first cluster of computers. Could he do that again and again, eventually installing mini-Google clusters in each computer science department? Surely not. To extend Google 101 to universities around the world, the participants needed to plug into a shared resource. Bisciglia needed a bigger cloud.

That's when luck descended on the Googleplex in the person of IBM Chairman Samuel J. Palmisano. This was "Sam's day at Google," says an IBM researcher. The winter day was a bit chilly for beach volleyball in the center of campus, but Palmisano lunched on some of the fabled free cuisine in a cafeteria. Then he and his team sat down with Schmidt and a handful of Googlers, including Bisciglia. They drew on whiteboards and discussed cloud computing. It was no secret that IBM wanted to deploy clouds to provide data and services to business customers. At the same time, under Palmisano, IBM had been a leading promoter of open-source software, including Linux. This was a key in Big Blue's software battles, especially against Microsoft. If Google and IBM teamed up on a cloud venture, they could construct the future of this type of computing on Google-based standards, including Hadoop.

Google, of course, had a running start on such a project: Bisciglia's Google 101. In the course of that one day, Bisciglia's small venture morphed into a major initiative backed at the CEO level by two tech titans. By the time Palmisano departed that afternoon, it was established that Bisciglia and his IBM counterpart, Dennis Quan, would build a prototype of a joint Google-IBM university cloud.

Over the next three months they worked together at Google headquarters. (It was around this time, Bisciglia says, that the cloud project evolved from 20% into his full-time job.) The work involved integrating IBM's business applications and Google servers, and equipping them with a host of open-source programs, including Hadoop. In February they unveiled the prototype for top brass in Mountain View, Calif., and for others on video from IBM headquarters in Armonk, N.Y. Quan wowed them by downloading data from the cloud to his cell phone. (It wasn't relevant to the core project, Bisciglia says, but a nice piece of theater.)

The Google 101 cloud got the green light. The plan was to spread cloud computing first to a handful of U.S. universities within a year and later to deploy it globally. The universities would develop the clouds, creating tools and applications while producing legions of computer scientists to continue building and managing them.

Those developers should be able to find jobs at a host of Web companies, including Google. Schmidt likes to compare the data centers to the prohibitively expensive particle accelerators known as cyclotrons. "There are only a few cyclotrons in physics," he says. "And every one of them is important, because if you're a top-flight physicist you need to be at the lab where that cyclotron is being run. That's where history's going to be made; that's where the inventions are going to come. So my idea is that if you think of these as supercomputers that happen to be assembled from smaller computers, we have the most attractive supercomputers, from a science perspective, for people to come work on."

As the sea of business and scientific data rises, computing power turns into a strategic resource, a form of capital. "In a sense," says Yahoo Research Chief Prabhakar Raghavan, "there are only five computers on earth." He lists Google,

Yahoo, Microsoft, IBM, and Amazon. Few others, he says, can turn electricity into computing power with comparable efficiency.

All sorts of business models are sure to evolve. Google and its rivals could team up with customers, perhaps exchanging computing power for access to their data. They could recruit partners into their clouds for pet projects, such as the company's clean energy initiative, announced in November. With the electric bills at jumbo data centers running upwards of \$20 million a year, according to industry analysts, it's only natural for Google to commit both brains and server capacity to the search for game-changing energy breakthroughs.

What will research clouds look like? Tony Hey, vice-president for external research at Microsoft, says they'll function as huge virtual laboratories, with a new generation of librarians—some of them human—"curating" troves of data, opening them to researchers with the right credentials. Authorized users, he says, will build new tools, haul in data, and share it with far-flung colleagues. In these new labs, he predicts, "you may win the Nobel prize by analyzing data assembled by someone else." Mark Dean, head of IBM's research operation in Almaden, Calif., says that the mixture of business and science will lead, in a few short years, to networks of clouds that will tax our imagination. "Compared to this," he says, "the Web is tiny. We'll be laughing at how small the Web is." And yet, if this "tiny" Web was big enough to spawn Google and its empire, there's no telling what opportunities could open up in the giant clouds.

It's a mid-November day at the Googleplex. A jetlagged Christophe Bisciglia is just back from China, where he has been talking to universities about Google 101. He's had a busy time, not only setting up the cloud with IBM but also working out deals with six universities—U-Dub, Berkeley, Stanford, MIT, Carnegie Mellon, and the University of Maryland—to launch it. Now he's got a camera crew in a conference room, with wires and lights spilling over a table. This is for a promotional video about cloud education that they'll release, at some point, on YouTube ([GOOG](#)).

Eric Schmidt comes in. At 52, he is nearly twice Bisciglia's age, and his body looks a bit padded next to his protégé's willowy frame. Bisciglia guides him to a chair across from the camera and explains the plan. They'll tape the audio from the interview and then set up Schmidt for some stand-alone face shots. "B-footage," Bisciglia calls it. Schmidt nods and sits down. Then he thinks better of it. He tells the cameramen to film the whole thing and skip stand-alone shots. He and Bisciglia are far too busy to stand around for B footage.

[Baker](#) is a senior writer for *BusinessWeek* in New York .

Xerox Color. It makes business sense.

Copyright 2000-2009 by The McGraw-Hill Companies Inc. All rights reserved.

The McGraw-Hill Companies