



COMPUTING AND THE CLOUD

The computing cloud continues to envelop daily life, from the way we do business at every level to how we entertain ourselves with videos, music, books and games. Why should companies, large and small, bear the expense of maintaining their own IT resources any longer? As a growing volume of advertising asserts, it's much better to be in the cloud.

There's also a great deal of money to be made in the cloud. Worldwide spending on public IT cloud services is expected to exceed \$107 billion in 2017 and to experience a compound annual growth rate of 23.5 percent between 2013–2017, five times that of the IT industry as a whole, according to the International Data Corporation, a market intelligence firm.

Cristian Borcea likens the rise of cloud computing in the 21st century to the transformation of the country's energy infrastructure more than a hundred years before, when home-based and local generation began to evolve into the modern grid, dominated by gigantic utilities pumping electricity to houses and offices hundreds of miles away.

“In the computing world of today, companies such as Google, Amazon and Microsoft are quickly becoming those utilities, leasing storage and computing services at their huge data centers,” says Borcea, associate professor and associate chair of the Department of Computer Sciences, who in 2010 offered one of the first college courses in the nation on cloud computing.

“Even though cloud storage systems are designed to offer high availability and reliability, outages and data loss incidents are still possible due to software and hardware malfunction, natural disasters, or malicious attacks.”

Reza Curtmola, associate professor of computer science

IBM is a striking illustration of the trend. Synonymous with the computer, the company sold its PC business to Lenovo in 2004, and earlier this year followed up by selling the Chinese technology company its Intel-based servers as well. “It’s just a service company now, and it offers private and hybrid cloud solutions to customers,” he notes.

“But a major difference between the two sectors,” Borcea adds, “is that the energy industry is highly regulated. Right now the cloud is not regulated in any way. Cloud providers are not liable for losses or breaches. The only laws that apply are general privacy laws and the only guarantees that exist on either reliability or data security are service agreements.”

DATA IN DANGER

Data in the cloud can be lost, stolen or corrupted. It can be mistakenly deleted when a server is updating. Hardware can fail. Software illegally distributed through the cloud can unleash embedded viruses. Service can become unavailable during downtimes. There is no legal redress for any of these mishaps. Under the current system, it’s the responsibility of users to back up their own data.

“Technology is always ahead of the law,” Borcea says, adding, however, that he expects the federal government to develop regulations over the next several years to protect the growing stream of data and computation migrating to the cloud. In the meantime, he and other NJIT colleagues are devising ways to bring transparency, accountability, and improvements in functionality to a system in transition.

Reza Curtmola, an associate professor of computer science, won a five-year “CAREER” award from the National Science Foundation in 2011 to devise methods for assessing the security of data storage through a remote testing mechanism he has developed.

“When people store data in the cloud they effectively lose control over it, while current storage models require them to trust cloud service providers,” he says. “You cannot sue if you lose data. Even though cloud storage

systems are designed to offer high availability and reliability, outages and data loss incidents are still possible due to software and hardware malfunction, natural disasters, or malicious attacks. This makes it difficult to assess the risk of outsourcing data and therefore unsuitable for applications that require long-term integrity and reliability.”

“What I’m trying to do,” he adds, “is to tackle the problem from a technical perspective, to come up with ways to provide additional guarantees by making sure that providers are doing what they’re expected to do. Right now, cloud users don’t have the means to check how their data is managed and cloud service providers are not transparent about how they go about it.”

Curtmola’s assessment mechanism would allow cloud clients to check whether their data is being stored correctly and whether the provider can produce it on demand. If adopted by cloud storage providers, it would offer additional guarantees, extending the range of applications that could benefit from cloud-based storage.

“I store a set of data in the cloud and then issue a challenge to the storage server to show me that the data is still there and uncorrupted. The goal is to detect if something goes wrong, because knowing is the important first step,” he says. “It’s important to detect early – and if you find that one part of your data is corrupted, you can take steps towards repairing.”

To restore data that has been lost or corrupted, Curtmola replaces it with healthy data stored with other providers. “Redundancy is important,” he says. “I recommend storing it in multiple places, maybe even in other clouds

at different providers. When the data is valuable, the additional cost of multiple replicas is worth it.”

But despite the system’s uncertainties, economics continues to drive business to the cloud.

WAY TOO MUCH STORAGE

Cloud computing began to emerge about a decade ago as a boon to large web-based companies such as Amazon and Google who looked at their operations and realized they had “way too much storage” on most days in order to accommodate periodic surges in demand, Borcea notes.

“Websites such as CNN.com are a good analogy. If there is no big news, comparatively few people are watching. But if there is news, everybody’s watching and so it makes economic sense to overprovision,” Borcea says. “Cloud providers, too, need to be able to prepare for the peak load, while on the average day they are nowhere near that, using maybe 10 to 15 percent of their total capacity.”

Amazon was the first web-based company to lease storage and it added computing services a year or so later. Leasing unused space made economic sense to Amazon, while its huge data centers, provisioned by hardware and Internet connections purchased at cheaper bulk rates, allowed the company to sell storage and computing at favorable rates.

“It’s a commodity – people are buying and selling as much as they need, and the price of storage is getting cheaper and cheaper as the technology advances,” notes Curtmola.

GREAT FOR START-UPS

Companies operating in the cloud are essentially leasing virtual computers on demand, and they can connect to these virtual computers from their own computing infrastructure. They save money on hardware as well as system administration and receive free or cheaper services such as e-mail and software packages.

“It’s a great deal for start-ups,” Borcea says. “If I’m a company at NJIT’s Enterprise Development Center, I may have a lot of customers

all of a sudden before I've earned much revenue. I may not be able to buy the hardware it would take to serve them so it would make sense to rent in the cloud, where you can scale up inexpensively and ultra-fast."

Increasingly, large companies are moving at least some of their data and operations to the cloud. *The New York Times* was in the vanguard when the company moved its digital archive to the cloud in 2006. Scientific research is also moving to the cloud. Hedge funds run some of their computationally-intensive models there as well. Big companies distribute software through the cloud.

"This will increase. Going forward, it's likely we'll have our own computers mostly for private use," Curtmola says.

For the time being, however, data sensitivity and long-term integrity still hinder broader use of the cloud and storage of more sensitive data," he adds. "Healthcare providers would not want to store medical data in the cloud, for example, and financial companies wouldn't put algorithms in the cloud."

SECURITY VS. ECONOMICS

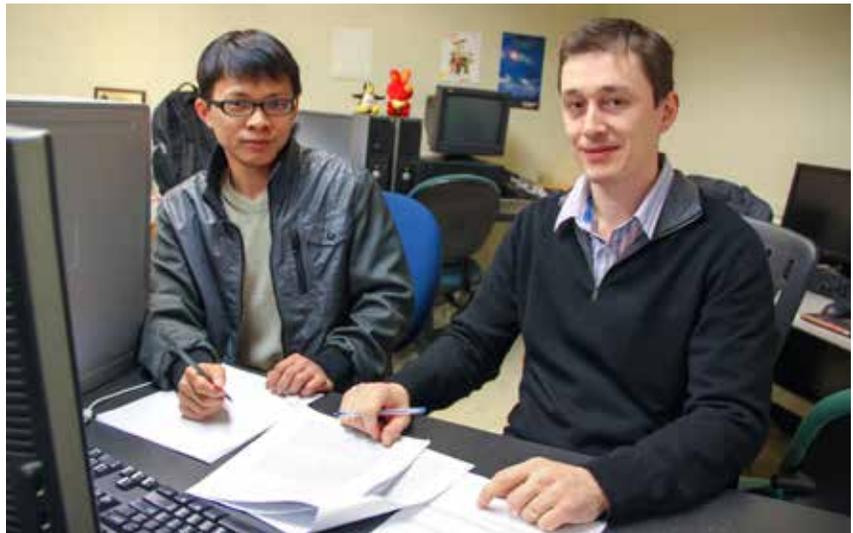
Disclosures of NSA surveillance have also sent shivers through the cloud, and Google says it is now encrypting e-mail when it is in transit as a result. "It's easier to tap into a wire than to gain access to a computer," Curtmola says.

But even security concerns are tempered by economic realities.

"How much security does it pay to add? Companies don't want to make handling data slower," Curtmola says. "Some data that is stored for safekeeping, but rarely used, can be encrypted. But if you want to search data or actually use it, it looks like gibberish and can't be read."

"For their part, cloud providers like Google and Amazon don't want you to encrypt, because they data mine. That's why services like Gmail are free. They mine it to advertise," Curtmola says, adding with a laugh, "That's one of the reasons why security measures proposed by the academic community are slow to be adopted."

Xiaoning Ding, an assistant professor of



computing sciences, works on ensuring application functionality and performance in the cloud.

"If there are problems with running a program, it can be difficult to determine if it's your program or the system from the service provider. It's important to be able to identify the root causes," says Ding, who studies application testing and troubleshooting with IBM and others.

"We run a program several times on a local machine and in the cloud," he says. "By monitoring and analyzing its behavior on different systems, we can identify functionality problems and reveal many insights about the root causes."

Ding also researches ways to improve the speed of applications and programs in the cloud by using multi-core processors.

"Some programs may be slowed down by more than ten times in the cloud," he says. "We want to make sure the cloud provides high performance for every program in it. We're working on designing a better infrastructure with researchers at Intel Labs, so programs will perform better. The main idea is to allocate computing resources in the cloud to programs in a way that better fits their real demand."

But as they try to make the cloud a more secure and high-functioning environment,

Associate Professor Reza Curtmola (right) is among the NJIT researchers working to make the cloud safer data-storage territory. In 2013, a paper co-authored with doctoral candidate Bo Chen, proposing better ways to ensure the integrity of data entrusted to cloud-service providers, received the Outstanding Paper Award from the Third Conference on Data and Application Security and Privacy sponsored by the Association for Computing Machinery.

NJIT researchers will also seek to exploit its current potential.

Borcea, who will be teaching a new graduate-level course on mobile cloud computing, is working on mobile phone technology that will leverage the cloud as an energy- and storage-saving platform.

"The main limitation of a mobile phone is battery capacity, but we can use the cloud to back up data and run apps. What I'm working on is an avatar, which is a software surrogate of the phone that would live in the cloud and would synchronize with the phone. A program that needed a lot of energy to run would operate in the cloud and it would perform faster." ■

<http://ccs.njit.edu>

Author: Tracey Regan is an NJIT Magazine contributing writer.