

Researchers work to take the errors out of the cloud

October 14 2011, By Bill Steele

Cloud computing, which taps the resources of a network of remote computers, offers tremendous potential for storing and processing vast amounts of data quickly and cheaply. The catch: As cloud computing applications become larger, the potential for errors also grows. So a Cornell team of computer scientists plans to develop methods for improving the reliability of cloud computing, testing their work in a piece of the real cloud.

The effort by Ken Birman, professor of computer science; Hakim Weatherspoon, assistant professor of computer science; and research associate Robbert van Renesse, is supported by a two-year, \$370,000 grant from the National Science Foundation (NSF). In a unique partnership between NSF and Microsoft Corp., Microsoft will give the researchers access to Windows Azure -- a cloud-computing platform that provides on-demand computing and storage through Microsoft data centers -- to test their ideas in a challenging environment. The team also will work with the massive [storage systems](#) maintained by the [Internet Archive](#).

Cloud computing refers to farming out [data storage](#) and computing onto huge data centers accessed over the Internet. Economies of scale allow cloud providers to offer services at a fraction of what it would cost users to set up and maintain their own server farms. And by locating near power sources like hydroelectric plants, [cloud computing](#) often promises to be eco-friendly.

[Cloud systems](#) make many copies of data or applications to give multiple users quick results. But replication can introduce errors, and avoiding those errors slows down the system. If a user is changing a bit of data, others have to be locked out of all copies of that data until the change is completed; and then the change has to be propagated to all the duplicates. So rather than making customers wait, many managers simply accept the risks of stale data in return for speed and low cost.

At [Amazon](#), for example, speed is a priority -- so customers may see information about a product in a fraction of a second, but the accuracy of that information is less reliable. Similarly, YouTube may occasionally serve the wrong video -- but chances are you'll just click again.

The result, the Cornell researchers say, is a system that is "inconsistent by design." But consistency is vital in managing medical records, the electric power grid or air traffic control.

The new research will be more about engineering than discovery. "It's not that we don't have a pretty good idea how to do it," Birman said. But testing a system on a large scale will demonstrate to the industry that such an approach is practical, he said.

One way to speed replication is by multicasting, in which a computer sends a data change to many server addresses at once, rather than reeling off the addresses one after another. Cloud services have avoided multicasting because it can confuse routers, sending the same message to every computer in the system instead of just those meant to receive it. Birman gets around this by consolidating addresses into groups, somewhat the way an email program combines several addresses into a single alias. To avoid errors, Birman locks replicas with out-of-date data out of a working group to be restarted and reinitialized later.

Meanwhile, Weatherspoon is working to speed data transmission over

the fiber-optic lines that connect servers, and to design better ways to organize stored data. Since using the cloud is partly about saving energy, he hopes to distribute data in ways that make it possible to spin down large groups of disk storage units when they're not needed, saving both the power to drive the disks and the air conditioning to cool them.

Van Renesse focuses on security. He believes that so-called Byzantine security systems that are absolutely impenetrable can be scaled up to cloud size and still work -- perhaps not as fast as insecure systems, but fast enough to be useful.

"We can offer much stronger guarantees with pretty comparable speed," Birman concluded. "Not a radical breakthrough, just good engineering. We'll be giving out free software."

Provided by Cornell University

Citation: Researchers work to take the errors out of the cloud (2011, October 14) retrieved 23 April 2024 from <https://phys.org/news/2011-10-errors-cloud.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.