Cornell collaboration with IBM to speed up 'the cloud'

March 16 2010, By Bill Steele

(PhysOrg.com) -- More and more of today's computing is happening in "the cloud" -- not just on the desktop or even on the big servers in the basement but all over the Net at once. Government agencies, banks and companies like Google, Amazon and Microsoft maintain dozens of huge data centers all over the country and the world, all sharing data back and forth over high-speed fiber-optic lines.

But the data sharing doesn't always go smoothly; sometimes the short streams of ones and zeros known as data packets get distorted, delayed or dropped altogether. "It is not unusual for network packets travel 10,000 miles just to be dropped by the end-host," explained Hakim Weatherspoon, assistant professor of computer science, "but it is frustrating." Lost packets have to be resent, slowing down the overall data transfer.

Weatherspoon is collaborating with Hani Jamjoom, M.Eng. '97, a research manager at IBM's Watson Research Center, to study the causes of these distortions and develop ways for [cloud computing](#) applications to deal with them. Their work will be partly funded by a $20,000 IBM faculty award.

To find out what happens to data packets in their travels, Weatherspoon operates a testbed called the Cornell NLR Rings, which sends data on loops up to 16,000 miles round-trip around the National Lambda Rail high-speed fiber-optic research network. Network packets can be sent out and back through New York, Chicago, Denver or, in the largest loop,

on a complete circuit around the country.

In order to understand packet loss experienced by the computer at the end of the chain and the associated reduction in performance, Weathersoon's research group, collaborating with physics post-doctoral researcher Daniel Freedman, has developed an apparatus that uses a very precisely modulated laser to generate packets of optical signals, then analyzes what comes back with sub-picosecond accuracy. "The instrument measures 'ground truth' on the wire," Weatherspoon explained. In other words, it shows what's really happening, not just a measurement from a distance. Software to operate the testing device was developed by Freedman and Weatherspoon's graduate student Tudor Marian.

A surprise in early testing was that transmission problems show up on the uncongested LambdaRail network, meaning they also may appear on private networks used by businesses and institutions. "I have discovered that contrary to the widely held supposition that such networks are largely stable, lossless and jitter free, these networks can be rather unstable, prone to loss and sources of significant jitter," Weatherspoon reported in his proposal to IBM. The longer the path, the worse packet loss experienced by the end-host, he added.

A key problem, the research shows, is that packets tend to bunch up en route and arrive in rapid-fire streams that the receiving computers can't process fast enough.

Other aspects of the IBM collaboration include developing software that parallelizes the incoming signals to make better use of "multicore" parallel processors, which may help in dealing with rapid-fire incoming packets, and finding ways to reduce energy consumption in large data centers by telling the systems when they can turn off disks that are not needed.

Overall, Weatherspoon said, there are "mismatches between abstraction and physical realities" in today's cloud computing. By removing these bottlenecks, he said, he hopes to facilitate the next generation of data centers.

Provided by Cornell University