

# Data Travels Six Times Faster in the Clouds

26 February 2009

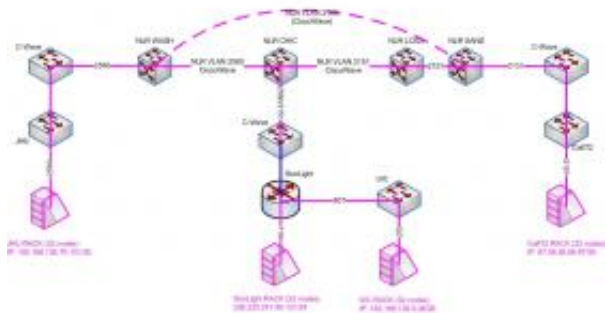


Sector is cloud computing system designed for data intensive computing. Sector is designed to run on racks of commodity computers (such as pictured). The racks may be located within a single data center or across several geographically distributed data centers. Credit: Michal Sabala, NCDM, University of Illinois at Chicago

(PhysOrg.com) -- The National Center for Data Mining (NCDM) at the University of Chicago at Illinois established a cloud computing system that can quickly compile data from widely geographically distributed data centers across high performance networks. NCDM used the Open Cloud Testbed, managed by the Open Cloud Consortium, to demonstrate the "Sector System" at the annual meeting of the American Association for the Advancement of Science conference earlier this month in Chicago.

"We demonstrated that our system is six times faster than competing technology," said Robert Grossman, NCDM director and Open Data Group managing partner. "Without the requirement of costly and cumbersome data transfer from various locations to one central location, this opens the way to exciting collaborative scientific discovery."

THE OPEN CLOUD TESTBED



This is a diagram of Phase 1 of the Open Cloud Testbed. Phase 1 of the Open Cloud Testbed consists of 4 racks located at the University of Illinois at Chicago, the StarLight Facility in Chicago, Johns Hopkins University in Baltimore, Maryland, and Calit2 at the University of California at San Diego. The 4 racks are connected by a 10 Gb/s network, provided by the Cisco C-Wave and regional high performance networks at each of the locations. The Open Cloud Testbed is managed by the Open Cloud Consortium. Credit: Open Cloud Consortium

Grossman and his team demonstrated using a common benchmark called Terasort. They found there was less than a 5 percent performance penalty when Terasort was run across the four data centers distributed across the country compared to running the entire computation within one data center. Prior to the Sector System, such computations were rarely done, as performance penalties were as high as 30 percent.

"With the Sector System, data intensive computing can scale not only to a data center, but for the first time, across data centers," said Grossman. "This enables locating data centers in areas in which power and cooling is cost-effective."

The Open Cloud Testbed consists of racks of computers located at the University of Illinois at Chicago, the StarLight facility in Chicago, Johns Hopkins University in Baltimore, Maryland, and the University of California at San Diego, all connected

by a wide area 10 Gb/s network, and all running a variety of cloud computing services, including cloud storage services and cloud computing services. The technology that makes this possible uses an open architecture design, specifically the open source sector system developed by the NCDM (sector.sf.net).

Although cloud computing is becoming common, processing data by clouds today is almost always done within a single data center. Generally, data intensive computing across geographically distributed data centers is avoided due to the difficulties and cost of moving large amounts of data over long distances. Sector employs an alternative network protocol called UDT designed to swiftly and smoothly transfer data.

According to Joe Mambretti, director of the International Center of Advanced Internet Research at Northwestern University and co-director of the Open Cloud Testbed, "These innovative technologies provide unique capabilities that will enable new generations of applications that can make discoveries involving large volumes of highly distributed data."

Provided by National Science Foundation

APA citation: Data Travels Six Times Faster in the Clouds (2009, February 26) retrieved 19 May 2019 from <https://phys.org/news/2009-02-faster-clouds.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.*