

Solution to genomic analysis may be in the clouds

December 5 2014, by Rogini Moorthi



"At the moment, most of the analyses involving genomic data or DNA sequencing are outsourced to service laboratories, where the data is then sent back to the researchers," Dr Hewitt says. Credit: Polygon Medical Animation

Cloud computing is a more efficient and cheaper alternative for researchers wanting to access and analyse large amounts of human genomic data, a local study has found.

The research compared the performance of two [cloud computing](#) service providers—Google Compute Engine (GCE) and Amazon Web Services Elastic MapReduce (EMR).

The scientists investigated the impact of data size on the two cloud platforms using publicly available genomic data (one small *Escherichia coli* CC102 strain and one large Han Chinese male genome).

In both EMR and GCE, CPU utilisation, memory usage and network speeds were monitored between April and June 2014 using Ganglia (open-source monitoring system).

The findings highlighted that analysing genomic data using Google Compute Engine is cheaper and more efficient than its counterpart.

Lions Eye Institute researcher Dr Alex Hewitt says cloud computing refers to accessing data and programs over the internet instead of a computer's hard drive.

"Amazon currently dominates the cloud computing market," he says.

"Most of the bioinformatics tools [software programs designed for extracting meaningful information from genomic data] to date have also been optimised for Amazon."

"We compared the performances of EMR to GCE because we wanted to showcase that other cloud computing services could be cheaper and possibly more efficient.

"This would in turn drive the developers of bioinformatics tools to ensure that they can be used across a variety of cloud providers."

Dr Hewitt says a standard genomic alignment was faster in GCE as Google uses a 2.6 GHz Intel Xeon Sandy Bridge CPU in comparison to Amazon, which uses a 2.0 GHz CPU.

"Furthermore, we found that assessing the genomic data using GCE

involved a cheaper cost of USD\$0.352 per hour in comparison to EMR, which charged \$0.640 per hour."

Cloud computing will become more popular

Dr Hewitt says more and more human genomic analysis will be performed using cloud computing over the next few years.

"The human genome contains close to 25,000 genes and takes up about 100 gigabytes of storage," he says.

"Storing, much less analysing and sharing, all that information is far beyond the capacity of most universities and research institutions.

"At the moment, most of the analyses involving [genomic data](#) or DNA sequencing are outsourced to service laboratories, where the data is then sent back to the researchers."

Dr Hewitt says cloud computing is becoming increasingly important for researchers as it could facilitate sharing and comparing genomes and using large amounts of genetic data to make medical discoveries.

More information: "Benchmarking Undedicated Cloud Computing Providers for Analysis of Genomic Datasets." *PLoS ONE* 9(9): e108490. [DOI: 10.1371/journal.pone.0108490](https://doi.org/10.1371/journal.pone.0108490)

Provided by Science Network WA

Citation: Solution to genomic analysis may be in the clouds (2014, December 5) retrieved 23 April 2024 from <https://phys.org/news/2014-12-solution-genomic-analysis-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.