

# Research brings cloud costs back to earth

July 13 2011

---



(PhysOrg.com) -- Researchers from Swinburne University of Technology are looking for ways to reduce the high cost of internet data storage and retrieval in cloud computing.

While [cloud computing](#) – which relies on remote, rather than local servers – offers almost unlimited capacity for data storage and processing, current usage charges mean the costs are expanding at the same near-limitless rate.

Social media such as Facebook and Flickr are simple examples of cloud computing, but the drain on resources from these sites doesn't compare to the volumes of high-end data generated by the world's research institutions, healthcare systems and industries.

Government agencies such as the Australian Taxation Office, Bureau of Statistics, and Treasury are all potential heavy users of cloud computing services, and the costs to them are high and rising. An estimated \$1 billion could be saved if the Australian government develops a data centre strategy – the core for cloud computing – for the next 15 years.

This is why, using funding from an Australian Research Council Discovery Project Grant, researchers from Swinburne's Centre for Computing and Engineering Software Systems (SUCCESS), are developing more cost-effective models for cloud computing's heavy users.

Professor Yun Yang and Professor John Grundy (from Swinburne) and Dr. Jinjun Chen (now with the University of Technology, Sydney) have been exploring the management of raw data and intermediate data sets, which are generated from processing this initial data.

“The trade-off is going to be between storage cost and computation cost,” Professor Grundy said. “Finding this balance is complex, and there are currently no decision-making tools to advise on whether to store or delete intermediate datasets, and if to store, which ones.”

To overcome this, the researchers have developed a mathematical model which factors in the size of the initial datasets, the rates charged by the service provider and the amount of intermediate data stored in the specified time.

“The formula can be used to find the best deals for storing data in the

cloud,” Professor Yang said.

They have also developed an Intermediate Data-dependency Graph (IDG) which helps users decide whether they are better off spending money on storage or computation for intermediate datasets.

“IDG records how each intermediate dataset is generated from the one before it and shows the generation relationship between them. This means if a deleted intermediate dataset needs to be regenerated, the IDG could find the nearest predecessor of the dataset. This can save computation cost, time and electricity consumption,” Professor Grundy said.

The researchers have been evaluating these solutions by simulating a pulsar survey used to crunch information from radio telescopes.

“Searching for pulsars – rapidly spinning stars that beam light – is a typical scientific application,” Professor Yang said. “It generates vast amounts of data – typically at one gigabyte per second. That data will be processed and may be reanalysed by astronomers all over the world for years to come.

“We used the prices offered by Amazon cloud’s cost model for this evaluation. For example, 15 cents per gigabyte per month for storage, and 10 cents per hour for computation.”

From one set of raw beam data collected by the telescope, the pulsar application generated six milestone intermediate datasets. The model generated three different cost scenarios. The minimum cost for one hour of observation data from the telescope and storing intermediate data for 30 days was \$200; for storing no data and regenerating when needed, \$1000; and for storing all intermediate data, \$390.

This gave the researchers options for which data to keep, and which to delete. “We could delete the intermediate datasets that were large in size but with lower generation expenses, and save the ones that were costly to generate, even though small in size,” Professor Yang said.

These are only a few of the solutions the researchers have come up with so far. To cater to different sectors, the group is also working on models that will allow users to determine the minimum cost on-the-fly, and as frequently as they wish.

Provided by Swinburne University of Technology

Citation: Research brings cloud costs back to earth (2011, July 13) retrieved 17 July 2024 from <https://phys.org/news/2011-07-cloud-earth.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.