

## Data World Record Falls as Computer Scientists Break Terabyte Sort Barrier

July 27, 2010

Daniel Kane

Computer scientists from the University of California, San Diego broke "the terabyte barrier" - and a world record - when they sorted more than one terabyte of data (1,000 gigabytes or 1 million megabytes) in just 60 seconds.

During this 2010 "Sort Benchmark" competition - the "World Cup of data sorting" - the computer scientists from the UC San Diego Jacobs School of Engineering also tied a world record for fastest data sorting rate. They sorted one trillion data records in 172 minutes - and did so using just a quarter of the computing resources of the other record holder.

Companies looking for trends, efficiencies and other competitive advantages have turned to the kind of heavy duty data sorting that requires the hardware muscle typical of data centers. The Internet has also created many scenarios where data sorting is critical. Advertisements on Facebook pages, custom recommendations on Amazon, and up-to-the-second search results on Google all result from sorting data sets as large as multiple petabytes. A petabyte is 1,000 terabytes.

"If a major corporation wants to run a query across all of their page views or products sold, that can require a sort across a multi-petabyte dataset and one that is growing by many gigabytes every day," said UC San Diego computer science professor Amin Vahdat, who led the project. "Companies are pushing the limit on how much data they can sort, and how fast. This is data analytics in real time," explained Vahdat. Better sort technologies are needed, however. In data centers, sorting is often the most pressing bottleneck in many higher-level activities, noted Vahdat who directs the Center for Networked Systems (CNS) at UC San Diego.

The two new world records from UC San Diego are among the 2010 results released recently on <http://sortbenchmark.org> - a site run by the volunteer computer scientists from industry and academia who manage the competitions. The competitions provide benchmarks for data sorting and an interactive forum for researchers working to improve data sorting techniques.

**World Records** The Indy Minute Sort and the Indy Gray Sort are the two data sorting world records the UC San Diego computer scientists won in 2010, the first year they entered the Sort Benchmark competition.

In the Indy Minute Sort, the researchers sorted 1.014 terabytes in one minute - thus breaking the minute barrier for this terabyte sort for the first time. "We've set our research agenda around how to make this better...and also on how to make it more general," said UC San Diego computer science PhD student Alex Rasmussen, the lead graduate student on the team.

The team also tied the world record for the "Indy Gray Sort" which measures sort rate per minute per 100 terabytes of data.

"We used one fourth the number of computers as the previous record holder to achieve that same sort rate performance - and thus one fourth the energy, and one fourth the cooling and data center real estate," said George Porter, a Research Scientist at the Center for Networked Systems at UC San Diego. The Center for Networked Systems is an affiliated Center of the California Institute for Telecommunications and Information Technology (Calit2).

Both world records are in the "Indy" category - meaning that the systems were designed around the specific parameters of the Sort Benchmark competition. The team is looking to generalize their results for the "Daytona" competition and for use in the real world.

"Sorting is also an interesting proxy for a whole bunch of other data processing problems. Generally, sorting is a great way to measure how fast you can read a lot of data off a set of disks, do some basic processing on it, shuffle it around a network and write it to another set of disks," explained Rasmussen. "Sorting puts a lot of stress on the entire input/output subsystem, from the hard drives and the networking hardware to the operating system and application software."

**Balanced Systems** The data sorting challenges the computer scientists took on are quite different from the modest sorting that anyone with off the shelf database software can do by comparing two tables. One of the big differences is that data in terabyte and petabyte sorts is well beyond the memory capacity of the computers doing the sorting.

In creating their heavy duty sorting system, the computer scientists designed for speed and balance. A balanced system is one in which computing resources like memory, storage and network bandwidth are fully utilized and as few resources as possible are wasted.

"Our system shows what's possible if you pay attention to efficiency - and there is still plenty of room for improvement," said Vahdat, holder of the SAIC Chair in Engineering in the Department of Computer Science and Engineering at UC San Diego. "We asked ourselves, 'What does it mean to build a balanced system where we are not wasting any system resources in carrying out high end computation?'" said Vahdat. "If you are idling your processors or not using all your RAM, you're burning energy and losing efficiency." For example, memory often uses as much or more energy than processors, but the energy consumed by memory gets less attention.

To break the terabyte barrier for the Indy Minute Sort, the computer science researchers built a system made up of 52 computer nodes. Each node is a commodity server with two quad-core processors, 24 gigabytes (GB) memory and sixteen 500 GB disks - all inter-connected by a Cisco Nexus 5020 switch. Cisco donated the switches as a part of their research engagement with the UC San Diego Center for Networked Systems. The compute cluster is hosted at Calit2.

To win the Indy Gray Sort, the computer science researchers sorted one trillion records in 10,318 seconds (about 172 minutes), yielding their world-record tying data sorting rate of 0.582 terabytes per minute per 100 terabytes of data. The winning sort system is made up of 47 computer nodes similar to those used in the minute sort.

According to wolframalpha.com, 100 terabytes of data is roughly equivalent to 4,000 single-layer Blu-Ray discs, 21,000 single-layer DVDs, 12,000 dual-layer DVDs or 142,248 CDs (assuming CDs are 703 MB).

Media Contacts: dbkane@ucsd.edu, (858)534-3262

