MIT Big Data Storage System Dramatically Speeds Information Access

Helen Knight, MIT

CAMBRIDGE, MA -- As computers enter ever more areas of our daily lives, the amount of data they produce has grown enormously.

But for this “big data” to be useful it must first be analyzed, meaning it needs to be stored in such a way that it can be accessed quickly when required.

Previously, any data that needed to be accessed in a hurry would be stored in a computer’s main memory, or dynamic random access memory (DRAM) — but the size of the datasets now being produced makes this impossible.

So instead, information tends to be stored on multiple hard disks on a number of machines across an Ethernet network. However, this storage architecture considerably increases the time it takes to access the information, according to Sang-Woo Jun, a graduate student in the Computer Science and Artificial Intelligence Laboratory (CSAIL) at MIT.

“Storing data over a network is slow because there is a significant additional time delay in managing data access across multiple machines in both software and hardware,” Jun says. “And if the data does not fit in DRAM, you have to go to secondary storage — hard disks, possibly connected over a network — which is very slow indeed.”

Now Jun, fellow CSAIL graduate student Ming Liu, and Arvind, the Charles W. and Jennifer C. Johnson Professor of Electrical Engineering and Computer Science, have developed a storage system for big-data analytics that can dramatically speed up the time it takes to access information.

The system, which will be presented in February 2014 at the International Symposium on Field-Programmable Gate Arrays in Monterey, CA, is based on a network of flash storage devices.

Flash storage systems perform better at tasks that involve finding random pieces of information from within a large dataset than other technologies. They can typically be randomly accessed in microseconds. This compares to the data “seek time” of hard disks, which is typically four to 12 milliseconds when accessing data from unpredictable locations on demand.

Flash systems also are nonvolatile, meaning they do not lose any of the information they hold if the computer is switched off.

In the storage system, known as BlueDBM — or Blue Database Machine — each flash device is connected to a field-programmable gate array (FPGA) chip to create an individual node. The FPGAs are used not only to control the flash device, but are also capable of performing processing operations on the data itself, Jun says.

“This means we can do some processing close to where the data is [being stored], so we don’t always have to move all of the data to the machine to work on it,” he says.

What’s more, FPGA chips can be linked together using a high-performance serial network, which has a very low latency, or time delay, meaning information from any of the nodes can be accessed within a few nanoseconds. “So if we connect all of our machines using this network, it means any node can access data from any other node with very little performance degradation, [and] it will feel as if the remote data were sitting here locally,” Jun says.

Using multiple nodes allows the team to get the same bandwidth and performance from their storage network as far more expensive machines, he adds.

The team has already built a four-node prototype network. However, this was built using five-year-old parts, and as a result is quite slow.

So, they are now building a much faster 16-node prototype network, in which each node will operate at 3 gigabytes per second. The network will have a capacity of 16 to 32 terabytes.

Using the new hardware, Liu is also building a database system designed for use in big-data analytics. The system will use the FPGA chips to perform computation on the data as it is accessed by the host computer, to speed up the process of analyzing the information, Liu says.

“If we’re fast enough, if we add the right number of nodes to give us enough bandwidth, we can analyze high-volume scientific data at around 30 frames per second, allowing us to answer user queries at very low latencies, making the system seem real-time,” he says. “That would give us an interactive database.”

As an example of the type of information the system could be used on, the team has been working with data from a simulation of the universe generated by researchers at the University of Washington. The simulation contains data on all the particles in the universe, across different points in time.

“Scientists need to query this rather enormous dataset to track which particles are interacting with which other particles, but running those kind of queries is time-consuming,” Jun says. “We hope to provide a real-time interface that scientists can use to look at the information more easily.”

Source URL (retrieved on *03/10/2014 - 2:39pm*):
http://www.scientificcomputing.com/news/2014/01/mit-big-data-storage-system-dramatically-speeds-information-access