UC SANTA BARBARA

THE Current

May 21, 2009
Andrea Estrada

UCSB Receives NSF Grant for Cloud Computing Research

UC Santa Barbara is one of 14 universities sharing nearly $5 million in grants from the National Science Foundation (NSF) to participate in the IBM/Google Cloud Computing University Initiative. The grants were awarded through the NSF's Cluster Exploratory (CLuE) program.

With cloud computing, users log into a Web-based service that hosts the applications they need rather than maintaining software on their own computers. Remote machines run everything, from email and word processing to complex data analysis. The term "cloud computing" refers to the cloud symbol that often represents the Internet on diagrams and flow charts.

IBM and Google established the university initiative in 2007 to help computer science students gain the skills necessary to build cloud applications. Universities will use software and services running on an IBM/Google Cloud to explore innovative research ideas in data-intensive computing. These projects cover a range of activities that could lead not only to advances in computing research, but also to significant contributions in science and engineering more broadly.

The UCSB group will explore many of today's data-intensive application domains, including searches on social networks such as Facebook, and protein matching in bioinformatics, all of which require answers to complex queries on highly connected data. The UCSB Massive Graphs in Clusters (MAGIC) project is focused on developing

software infrastructure that can efficiently answer queries on extremely large graph data sets. The MAGIC software will provide an easy-to-use interface for searching and analyzing data, and manage the processing of queries to efficiently take advantage of computing resources, such as large data centers.

"What's interesting about cloud computing is that most people don't realize they use it every day," said Ben Zhao, assistant professor of computer science and one of the four principal investigators. "Amazon, Facebook, and Gmail, for example, have hundreds of thousands of machines that are providing all the application functionality in a reliable and scalable way." The applications don't run on the user's machine, they run in the cloud, he said.

Because of the highly connected nature of these large data sets, existing management tools such as Google's MapReduce or Microsoft's Dryad, which are designed for use with independent data, would get bogged down with communication and data transfers between machines and would not be able to compute at full speed.

"With this grant we're studying the fundamental ideas behind how to efficiently figure out which relationships matter and which don't so we can minimize the need for communication and allow the machines to do their real work, which is computing," Zhao said.

While Zhao's work focuses on social networking, the other PI's at UCSB are working on different angles. Xifeng Yan, assistant professor of computer science and an expert in machine learning, studies biological networks and data mining. Amr El Abbadi, chair and professor of computer science, and Divyakant Agrawal, professor of computer science, are examining how to make databases more intelligent and allow them to support queries about graph data sets.

"One interesting aspect of this proposal is that it comes from the NSF with the support of IBM and Google," Zhao said. "We're working on really large-scale systems, the kind that only large companies like IBM and Google have. We'll have access to 1,600 machines from IBM and Google data centers, and we'll be able to run extremely large- scale experiments on them."

Other institutions participating in the IBM/Google Cloud Computing University Initiative include UC Irvine, UC San Diego, Carnegie-Mellon University, Florida International University, M.I.T., Purdue University, University of Maryland, University

of Massachusetts, University of Virginia, University of Washington, University of Wisconsin, University of Utah, and Yale University.

About UC Santa Barbara

The University of California, Santa Barbara is a leading research institution that also provides a comprehensive liberal arts learning experience. Our academic community of faculty, students, and staff is characterized by a culture of interdisciplinary collaboration that is responsive to the needs of our multicultural and global society. All of this takes place within a living and learning environment like no other, as we draw inspiration from the beauty and resources of our extraordinary location at the edge of the Pacific Ocean.