

# Cloud Computing for the Sciences



**W**hat exactly is cloud computing, and how is it important to science and engineering? This simple-looking question turns out to be hard to answer. According to Wikipedia, “Cloud computing is a style of computing

in which dynamically scalable and often virtualized resources are provided as a service over the Internet. Users need not have knowledge of, expertise in, or control over the technology infrastructure ‘in the cloud’ that supports them.” In case you find this definition to be less than satisfying, Wikipedia goes on to add that “cloud computing services usually provide common business applications online that are accessed from a web browser, while the software and data are stored on the servers.” In nuts-and-bolts terms, I’d say that

1521-9615/09/\$25.00 © 2009 IEEE  
COPUBLISHED BY THE IEEE CS AND THE AIP

FRANCIS SULLIVAN  
*IDA Center for Computing Sciences*

the cloud is central computing services supplied via a distributed computing platform making use of technologies and concepts that have emerged over the past 10 or so years.

Clearly, advances in technology—especially as regards network bandwidth and CPU cost—have made the idea of a distributed computing model very attractive. Equally clear are the issues raised by the cost of electrical power and the need for communication security and data integrity. But none of these is in itself the definition of cloud computing; the big question for this magazine's readers isn't so much the definition but rather what is the role of the cloud in scientific computing.

Accordingly, this special issue of *CiSE* isn't intended to provide a definition of cloud computing—it isn't even a status report on the current state of cloud computing. Our aim is to provide some background on cloud computing and begin to address the issues related to scientific computing. I'm quite confident that we'll see other articles on this topic in future issues of this magazine.

The three articles in the current issue discuss three important aspects of cloud computing. James L. Johnson's article, "SQL in the Clouds," takes us behind the scenes as it examines and generalizes MapReduce, one of the fundamental cloud computing algorithms. Said simply, the "map" step of MapReduce breaks a problem into parts and distributes those pieces over many others that either do the work or distribute again; the "reduce" step takes the answers to all the subproblems and combines them to get the output. The article "Graph Twiddling in a MapReduce World" by Jonathan Cohen begins an examination of the possibility of decomposing useful graph operations in terms of MapReduce cycles. The challenge for scientific computing is to carry out a similar effort for other algorithms important to scientific computing. Thomas Sterling and Dylan Stark raise some of the important issues needing attention if the cloud model is to continue to advance in their article, "A High-Performance Computing Forecast: Partly Cloudy." In particular, they note that the cloud concept doesn't address and can't satisfy the needs of other workflow classes requiring extreme-scale, tightly coupled capability computing, large sensitive datasets, and optimized algorithms.

There is a very good reason for the ambiguity and vagueness in the definition of cloud computing. Because of the current very high level of activity concerning cloud computing and its possibilities as a revenue generator, a generally agreed upon definition doesn't yet exist. Clearly, each vendor would like its definition to be *the* definition. A

few months ago, I attended one of the many, many cloud computing conferences that have begun springing up like dandelions after a spring rain. As often happens at such events, the talks ranged from "okay" all the way up to "pretty good," while the snacks and the conversations during break-out sessions were truly great. I was on the fringes of one particularly interesting conversation that quickly evolved into a vigorous debate. The apparent topic was about the best ways to implement basic functions in a cloud architecture. The participants got so deeply engaged that they continued to talk after the end of the break, missed the next actual conference presentation, and probably would have missed the rest of the morning session had they not discovered that when they said cloud computing, they were talking about three different things. One thought of the cloud as a medium for scientific collaboration, another thought of it as a way to offer IT services such as email and database search, and the third was convinced that the cloud's sole real function was the storage and protection of data.

**I**n 2019, will we do all of our computing via laptops costing less than US\$100 and never suffer from any problems related to service or security levels? I'm sure we won't, but I'm equally sure that we'll use the cloud in a very big way.

*Francis Sullivan is the director of the IDA Center for Computing Sciences, Bowie, Maryland. He's also the former editor in chief of this magazine. Contact him at [fran@super.org](mailto:fran@super.org).*

## Contact *CiSE*

**Writers:** Visit our "Write for Us" section at [www.computer.org/cise/author.htm](http://www.computer.org/cise/author.htm).

**Letters to the Editors or Article Reprints:** Email [cise@computer.org](mailto:cise@computer.org). Provide an email address or daytime phone number.

**Subscribe:** Visit [https://www.aip.org/forms/journal\\_catalog/order\\_form\\_fs.html](https://www.aip.org/forms/journal_catalog/order_form_fs.html) or [www.computer.org/subscribe/](http://www.computer.org/subscribe/).

**Subscription Change of Address:** For the IEEE or IEEE Computer Society, email [address.change@ieee.org](mailto:address.change@ieee.org). For the AIP, email [subs@aip.org](mailto:subs@aip.org). Specify *CiSE*.

**Missing or Damaged Copies:** For IEEE Computer Society subscribers, email [help@computer.org](mailto:help@computer.org). For AIP subscribers, email [claims@aip.org](mailto:claims@aip.org).

**Reprint Permission:** Email William Hagen, Copyrights & Trademarks Manager, at [copyrights@ieee.org](mailto:copyrights@ieee.org).

[www.computer.org/cise](http://www.computer.org/cise) or <http://cise.aip.org>