From the Editor in Chief...

Staring at Clouds

Fred Douglis • *f.douglis@computer.org*

Later this year, *IC* will be running a special theme issue on *cloud computing*, but after recently attending the Cloud Computing Expo in New York, I decided to jump the gun: I expect to focus on cloud computing for the next few columns because it's a fertile topic that's of great personal interest.

30,000-Foot Overview

Many of *IC*'s readers are no doubt already familiar with cloud computing, but I'll start with a quick overview for those who aren't. Additional background is available from many sources, including Wikipedia (http://en.wikipedia.org/wiki/Cloud_computing).

Generally, cloud computing's premise is to lower computing costs (both in money and time) by providing computational resources in a shared infrastructure. This sharing lets end users avoid capital expenditures and simply pay for resources as they use them. It also greatly reduces deployment times: companies can sign up for these resources and be using them in a matter of minutes or hours rather than the weeks or months it might take to order and install new hardware.

One technology that's greatly facilitated the cloud computing paradigm is *virtualization*. By letting users execute in a virtual environment, the cloud provider can isolate them from each other and support a consistent execution environment regardless of heterogeneity in the underlying hardware or software.

Virtualization is commonplace, though not a requirement for supporting cloud computing. Similarly, most platforms use shared resources, but some explicitly devote entire computers to single clients. Many meter the service and charge for processor time, storage, and network bandwidth, though flat-rate subscriptions over set time periods are also possible.

Paradigm Shift?

In the early 1990s, Ian Foster and Carl Kesselman coined the term "grid computing" as a way of conveying the availability of computation the same way that electric grids make power available to the masses. But over time, the term "grid" became strongly associated with high-performance scientific computing, in which institutions donated or exchanged computational resources but didn't pay for them incrementally. When I first started hearing about "cloud" computing rather than "grid" computing, I had the impression that cloud computing was simply repackaging the old technology. Now, I believe these are truly distinct (though related) technologies, and that cloud computing advocates were merely beaten to the more appropriate name! This is unfortunate because it leads to confusion, but such is life.

Early last year, Nicholas Carr published *The Big Switch: Rewiring the World, from Edison to Google* (W.W. Norton & Company, 2008). This book discusses the electric power grid's early days, and how large companies went from producing their own electricity to buying it from utilities, virtually overnight. It then relates this transformation to the availability of public computational resources. The analogy is a powerful one, though the targeted consumer is different. Large companies might run their own internal clouds to amortize infrastructure costs, but they're reluctant to leverage external platforms due to the control and security they must cede to service providers. Smaller companies, however, are finding the cloud a godsend because it lets them

- start building, testing, and deploying their applications remarkably quickly and
- react quickly to bursts in load.

For example, at his keynote at the expo, Amazon's Werner Vogels talked about Animoto's (www.animoto.com) use of Amazon Web Services to develop its system to merge photos into

 1089-7801/09/$25.00 © 2009 IEEE

music videos and how it reacted to being "slashdotted" and having its load increase by orders of magnitude when word got out.

To some extent, moving to the cloud requires a company to shift from provisioning its hardware to provisioning its software. It needs to be able to react to bursts in its load, or perhaps work with a company such as RightScale (www.rightscale.com) to monitor the system and manage that adaptation on its behalf. Companies must also determine whether to use solely open source software or deal with the licensing issues pertaining to running other applications on arbitrary numbers of compute nodes (see, for example, "Utility Computing's 'Dirty Little Secret'" at www.theregister.co.uk/2008/06/25/att_cloud_computing_dirty_secret/).

Mixing and Matching

One very popular topic at the expo was the question of interoperability and vendor lock-in. Customers of various cloud providers want to be able to move from one provider to another without dramatic reimplementation. In addition, they want to use multiple clouds at once — most commonly, a hybrid between an internal corporate cloud and an external provider. The keynote from David Douglas at Sun about multiple clouds is one example of how companies are looking at this space. The now-infamous "Cloud Computing Manifesto" (http://wiki.cloudcommunity.org/wiki/Cloud_Computing_Manifesto) became public just as the expo started, and generated a great deal of controversy based on which companies endorsed or declined to endorse its goals of open standards, interoperability, and so on.

Interoperability isn't important only to end users; it's extremely useful to the cloud ecosystem that's evolving around the basic infrastructure. A standard API would enable an external entity to build something

once, then use it to monitor and control a variety of platforms.

At the same time, heterogeneity has its benefits. I think back to the time that the original Internet Worm was unleashed, just over 20 years ago. I was working on a Unix-like operating system called Sprite, but Sprite wasn't binary-compatible with Unix; the worm managed to get in the door of our system but not replicate itself. One cloud computing mailing list has recently had an extended discussion about whether a cloud provider would be a juicy target for a hacker, given that someone who could break into the underlying system might somehow get access to applications from many customers in a single blow. To the extent that clouds today are each a bit different, if not dramatically so, they provide a more diffuse target. But as the industry matures — to the extent that it migrates toward a single interface — it might also become easier for attackers to threaten.

Today's news (in early April) is about hackers threatening our electrical grid; it will require our full attention to ensure that tomorrow's news isn't about the threats to our computational grid.

Acknowledgments

The opinions expressed in this column are my personal opinions. I speak neither for my employer nor for *IEEE Internet Computing* in this regard, and any errors or omissions are my own.

IEEE Internet Computing

Staff
Lead Editor: Rebecca Deuel-Gallegos, rdeuel-gallegos@computer.org
Director, Products & Services: Evan Butterfield
Senior Editorial Services Manager: Crystal R. Shif
Magazine Editorial Manager: Steve Woods
Staff Editors: James Figueroa, Jennifer Gardelle, and Jenny Stout
Production Editor/Webmaster: Monette Velasco
Publications Coordinator: Hazel Kosky, internet@computer.org
Contributing Editors: Cheryl Baltes, Greg Goth, Rita Scanlan, Keri Schreiner, and Joan Taylor
Graphic Artist: Alex Torres
Senior Business Development Manager: Sandy Brown, sbrown@computer.org
Membership/Circulation Marketing Manager: Georgann Carter
Senior Advertising Supervisor: Marian Anderson, manderson@computer.org

IEEE Computer Society Publications Office
10662 Los Vaqueros Circle
Los Alamitos, CA 90720 USA

Technical cosponsor:
IEEE COMMUNICATIONS SOCIETY

IEEE Communications Society Liaison
G.S. Kuo • gskuo@mail.com

Writers: Access www.computer.org/internet/author.htm. Articles are peer reviewed for technical merit and copy edited for clarity, style, and space. Unless otherwise stated, bylined articles and departments, as well as product and service descriptions, reflect the authors' or firms' opinion; inclusion in this publication does not necessarily constitute endorsement by the IEEE or the IEEE Computer Society.
Letters to the Editors: Email lead editor Rebecca Deuel-Gallegos at rdeuel-gallegos@computer.org.
On the Web: Access www.computer.org/internet/.
Subscribe: Visit www.computer.org/subscribe/.
Subscription Change of Address: Send requests to address.change@ieee.org.
Missing or Damaged Copies: Contact help@computer.org.
To Order Article Reprints: Email internet@computer.org or fax +1 714 821 4010.