

BY ALAIN LAEDERACH, PhD, STANFORD UNIVERSITY, DEPARTMENT OF GENETICS

Elastic Computing, Virtualization, and the Demise of the Common Cluster

After a hard day in “Lab,” it is rare to come home with the feeling of actually having accomplished anything tangible. Sure, the software I am developing has some new feature (with the accompanying 10 new bugs), and perhaps an algorithm is slightly more likely to yield biological insight, but all in all, I’ve mostly pushed and pulled electrons around. I often envy my experimental collaborators, who at the end of the day have some DNA in their Eppendorf, or a new plasmid that will express a protein. Those are tangible, day-to-day results.

The other day, however, I came home having the feeling of true accomplishment. We had spent the entire day in the server room reorganizing our clusters and servers: We pulled power and Ethernet cords, tightened screws, and inserted RAM chips. But this might be the last time I ever upgrade RAM on a rack mounted computer because Amazon (yes, the online book store where you bought the latest Harry Potter) now offers another option.

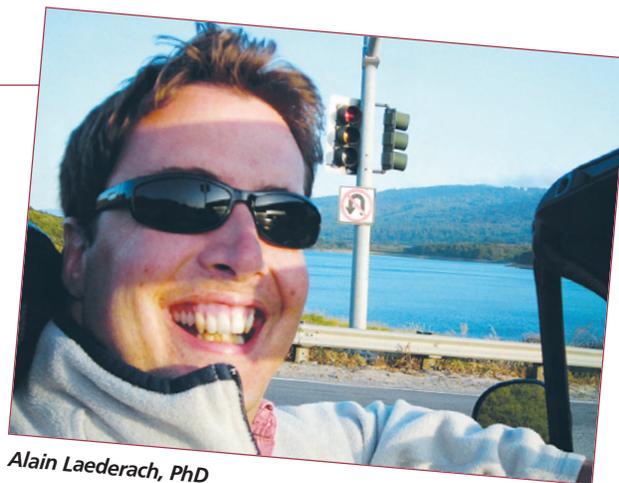
Until now, computing has been about machines—a fixed cost. Labs buy a certain number of computers with a new grant, and that is it. The number of jobs queued on the hardware is highly correlated with conference deadlines, but most of the time, the hardware keeps itself busy running daily cron jobs.

Now there is a new player in the game, it is called EC2 (for Elastic Cloud 2) and it is available at <http://www.amazon.com/ec2> along with more Harry Potter paraphernalia than I ever imagined. Amazon is the first company to sell computing as a true commodity independent of hardware. Large companies like Google, Amazon, Oracle and Microsoft are building data centers all over the country to meet their own huge CPU needs. But Amazon is the first to realize that their in-house technology (cheap commodity computing) can make them money—probably a lot of money.

DETAILS

Alain Laederach, PhD, is a post-doc in Russ Altman’s lab at Stanford University. He recently accepted a faculty position at the Wadsworth Center in Albany, NY, and he is not now and has never been associated with Amazon in any way!

You can reach him at alain@helix.stanford.edu



Alain Laederach, PhD

EC2 offers cheap and simple pricing (10 cents per CPU hour on a 3 GHz equivalent processor with 1.7 GB of RAM, and 160 GB of storage). But perhaps more important for computational research, it will mean no more queues before conference deadlines. Amazon will worry about load balancing the world’s computing resources; we’ll just pay for computation as we go.

To those of you who are skeptical about scales and complexity, consider this: Overnight on July 21, 2007, Amazon shipped 2.2 million pre-orders for the latest (and final) Harry Potter novel. This is a company that has some experience with load balancing. Some will still say that the loss in performance with the added layer of virtualization is unacceptable, and that interconnects between virtual machines will never be fast enough for their highly parallel problems. But I’m betting that, eventually, they’ll be buying CPU time too. The EC2 cloud already allows purchases of “large” instances (15 GB of memory, 8 CPUs, 1690 GB of instance storage, 64-bit platform) for 80 cents an hour. The price may sound steep, but consider the fact that you can create 1000 such instances almost instantaneously for \$800. That’s some cheap super-computing, IMHO. □

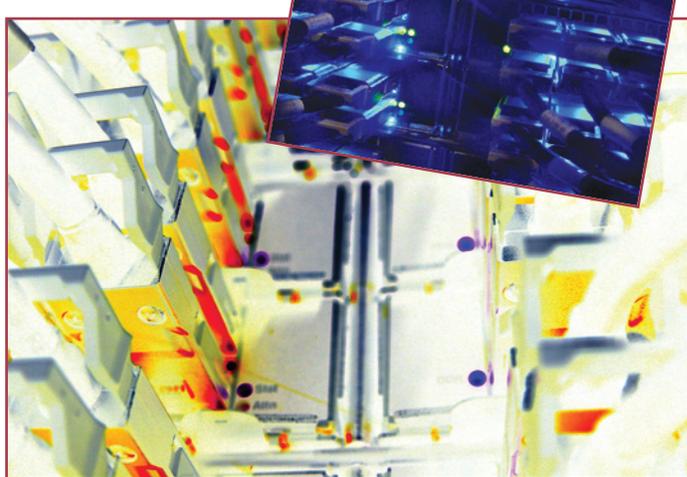


Photo credit: Shirley Wu