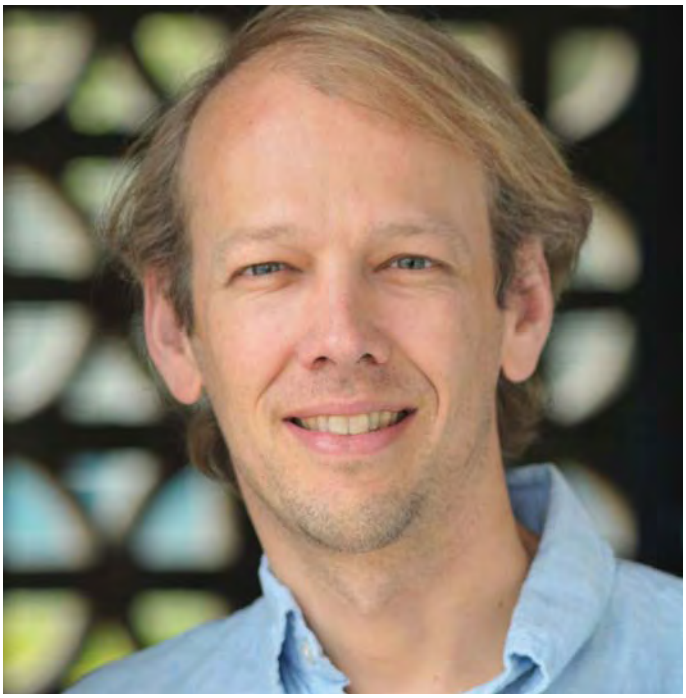
# THE CLOUD AT WORK

## *Interviews with Pete Beckman of Argonne National Lab and Bradley Horowitz of Google*



**Pete Beckman,**
**director of the Argonne Leadership Computing Facility, Argonne National Laboratory, interviewed by Sumit Narayan**

*Pete Beckman is the director of the Argonne Leadership Computing Facility at Argonne National Laboratory (ANL) in Illinois. Argonne National Lab is the United State's first science and engineering research laboratory as well as home to one of the world's fastest supercomputers.*

*Beckman explains cloud computing from a scientist's perspective, and speculates where it might be headed next. (He also notes that Argonne has a well-developed student internship program, but not enough candidates!)*

—Sumit Narayan

**Sumit Narayan**: "Cloud computing" is the new buzz word among computer scientists and technologists. It's used in different ways to define a variety of things. Tell us a little about the origins of cloud computing. What does cloud computing mean to you?

**Pete Beckman**: Distributed computing, which people have often referred as "computing out there," as opposed to "computing on your local machine," has been around for a very long time, somewhere around 20 years.

We went from a period of distributed computing, to meta-computing, to grid computing, and now to cloud computing. They're all a little different, but the notion is that the services, either compute or data, are located remotely, and scientists have to then work out protocols, policies, and security measures necessary to run stuffs remotely or to get to the data that is remote.

So, grid computing was focused primarily on sharing the resources among the providers and the genesis of cloud computing came from some technologies that allowed folks to do this in a very clear and sandboxed way. They provide a very definitive and well-described interface and also allow you to run a piece of code or software remotely.

Virtual machine technology in particular has made this much easier than in the past. Rather than the complicated nature of deciding which software can be run and which packages are available, now you are able to ship the entire virtual machine to other sites, thereby allowing for utility computing, another way we refer to cloud computing. The challenge that still remains is data, how to best share the data.

**SN**: What are they key new technologies behind cloud computing? Can you talk a little about the different services available through cloud computing?

**PB**: From a technical standpoint, the only one that has a strong root in technology is the virtual machine-based "infrastructure as a service" (IaaS). That's the only one that has a technological breakthrough. All the others are just a model breakthrough. In other words, the idea that I could store my data locally or remotely has been around for a long time.

The idea that I can create a web application that makes it look like I'm running something locally when it is really remote—these are model differences in terms of APIs and providing a per-user capacity and so forth. The technology, the one that is really a technological breakthrough, is using and shipping around virtual machines.

An example of model breakthrough is what people are doing when they say they are going to run their email on the cloud. To an organization, it looks like they have their emails present locally like they used to. This is because they never really had it close to them on a server down the hall. They were probably POP'ping, or IMAP'ping to a server that was within their infrastructure, but probably in another building. When people move their email to the cloud, they are now getting that as a service remotely and are being charged an incremental fee, like a per-user fee.

There is no provisioning that the site has to do for hosting and running their own machines. To the user, it all looks the same, except that the IMAP server is now some other IMAP server. The people are doing the same for calendars, HR, travel, etc.

So all these systems which used to reside locally in an IT organization on-site are getting cloud versions, which essentially, only requires a network connection.

Now the virtual machine part, that's really a technological breakthrough that allows me to run anything, not just what they provide in a package like POP or IMAP, but anything I want. That is unique and the new thing over the last couple of years.

**SN**: Do you think the era of personal computers is coming to a close? Will we increasingly rely on cloud services like Google for search, computation, on Dropbox/S3 for storage, or Skype for communication?
**PB**: It is changing. Whether or not it comes to a close, probably not. But it is changing dramatically.

Let me give you a couple of examples. There already are netbooks. People are migrating into using a really cheap portable device. We have had lightweight Linux for quite some time now. You may also know about Google's Chrome OS.

I have one of the initial versions of a netbook, and it is really cool. You can do simple student stuff: web, email, PDFs, Skype, basic spreadsheet, and Word documents. But now, a lot of people are using Google Docs for that. If you look at the future of computing, if we really get to a ubiquitous network and all these things rely on network, then these lightweight portal devices will become the way many people access their machine or their servers, or cloud services.

Some things still need intense local hacking, for example, image or media editing. But, even those are making their way out into the cloud. There is a great Photoshop-as-a-service app, that lets you upload a picture and then using a web interface, change the color, scratch it, shrink it, crop it, etc. And again, you don't really require a high-end notebook for that. An inexpensive Atom processor on your netbook is enough.

With respect to media editing, Animoto is an example one that is happening in the cloud. However, for most home users, there will still be a gap. Uploading photos or videos of my kids to the cloud for editing is probably still out of reach. But not for long.

**SN**: Cloud computing is well suited for business needs. As a scientist, why do you think cloud computing is important? What opportunities is cloud computing opening in the scientific community that weren't supported by big computing clusters like Blue Gene?
**PB**: Scientists have been using computers and hosting data remotely for a long time. They are used to this. It's kind of funny when people who are not from the scientific community visit Argonne, because they imagine scientists here come to use our supercomputer. The fact is, they don't. They are able to do [their computing] using SSH on their laptop from a coffee shop in Italy. The scientific community has been doing its work remotely for a long time, either on a supercomputer or mid-range machines.

But now instead of the science community provisioning all these servers and mid-range computers, we will be able to allow the cloud to do that for us.

Of course, supercomputers are still super. They are different from cloud resources that are primarily useful for mid-range computing.

Argonne has a project for cloud computing, Magellan, funded by the U.S. Department of Energy for research into the scientific community and the cloud as a way to free up scientists so that they are not provisioning, or setting up two or three servers down the hall.

The other thing that's changing is, in the past, supercomputers had a very well-defined set of software stacks—the package is either in the stack or not, in terms of support for that package. But with IaaS cloud architecture, scientists can customize and make their own special stacks. We see this a lot in metagenomics and biology, where scientists have a very complex workflow of 10 different tools and they want to create a web interface for that. They want it all together in a package so that they can run their genomics application. Doing it in a virtual machine means they don't have to worry whether their package is supported, or if they have the right version of Perl, or if they added the Java bindings to MySQL, they can just put everything together in a virtual machine and ship it around and run it wherever.

**SN**: Is there anything about cloud computing that worries you, like security, infrastructure? What are the risks?
**PB**: Security is very complex in a virtualized environment. It becomes a very big challenge and there are a couple of things we want to be able to do. We really want to give people a total virtual machine. That means we would be giving them root access on a virtual machine.

In the past, the language we have used to describe security has been to say that there is a "user," and an "escalated privileged user," like a root. All the documentation, discussion, cyber-security plans differentiate those two very clearly. Users and escalated privileged user—administrator, or a root.

In a virtualized environment, when you give someone a virtual machine, they have the root access on that virtual machine, not on the complete infrastructure. As an analogy, you can think of it as someone being able to turn a mobile phone on and off. That's administrator privileges on the phone. But you don't get to control the cell towers. They're still controlled by someone else: the mobile phone service providers. So this notion of security really has to change. We have a lot yet to explore and change, and there will be a lot of research in that space.

Another thing of course will be, if you do something that you are not supposed to do when you are using a virtual machine, who is to blame? Is it the users to whom you handed the virtual machine? Are they responsible for all the security? If they upload their virtual machine that has a bunch of security holes in it and someone gets into that—how do we stop that? How do we manage that risk? How do we scan our own virtual machine? So that's a pretty big research area, and one we will be exploring at Argonne.

**SN**: How do you think cloud computing would impact education?
**PB**: Oh I think cloud computing is just amazingly fun and fantastic for education, largely because of its low barrier to entry. If you look at the Beowulf, the cluster mailing lists, there are people who have set up their own clusters at various places. These are school-aged kids who have enormous amounts of energy, and they get a couple of old or new computers and wire them together.

Occasionally, you'll see stories about folks who have a 16-node cluster bought from the cheapest parts possible on the planet. These machines have no cases! They're just sitting on a table with their motherboards, and they work just fine.

There certainly is value in that sort of hacking, but a lot of colleges do not have that sort of expertise. Yet, they want to teach parallel processing, MPI programming and scientific calculations, MATLAB. Cloud computing offers promise here.

I can imagine in the future a student just being handed 100 credit-hours on the cloud. The professor would say, "We are going to do a homework assignment. We are going to write a code that calculates the surface area to volume ratio of this system with each particle moving around." Now, each student has his or her own credit hours that he or she uses in the cloud to do the computation.

That's the sort of thing where we are likely to be headed, and it's fantastic for universities! More and more students can get access to

resources. You can do a simple MATLAB thing or parallel processing, or write a new ray-tracer and do some visualization. However, it will still be mid-range computing. It won't have the impact of a 5,000-core supercomputer that has Infiniband, but it's a great way to get students started and tinkering.

I can imagine five years from now universities routinely handing out cloud credit hours to students.

**SN**: What are the opportunities for students to consider, both at ANL and outside?
**PB**: Argonne has a lot of student interns. We have a fantastic student program. Usually, our problem is that we cannot find enough students! It's not that we don't have enough slots for undergraduates or graduates who are in computational science or computer science—we don't have enough students!

Argonne has a catalog (**www.dep.anl.gov**) that lists all its projects. Students can essentially apply to a project for a summer position.

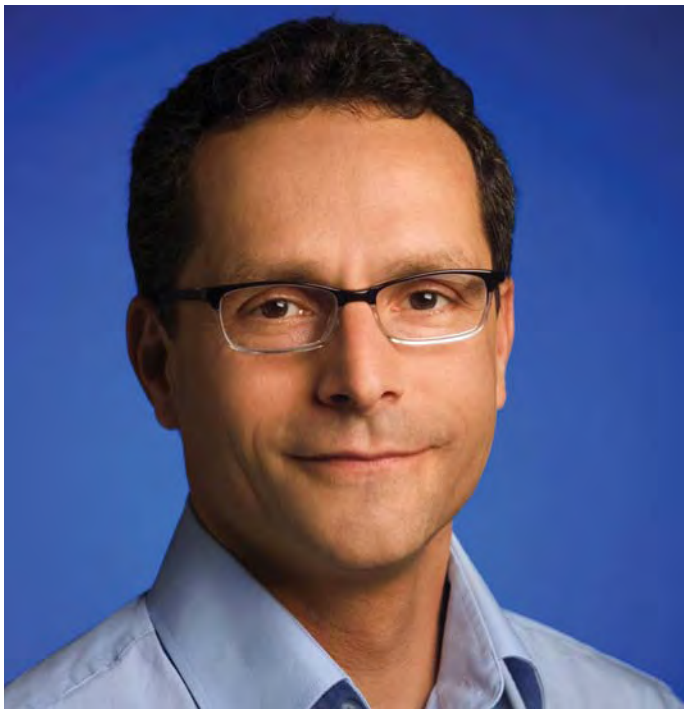
But with respect to the Magellan project, we have a web site that we are still working on: **www.megallen.alcf.anl.gov**. There, we will have a place to apply for cycles or time on the cloud machine. And if a student has a fantastic idea for exploring cloud computing in some way that benefits the lab and is in line with the mission of this lab in understanding cloud computing, then they can get time on the machine.

**SN**: What is your vision for the future, with respect to cloud computing? Where do you see cloud computing in ten years? Fifty years?
**PB**: We are slowly moving to faster and faster networks, even with respect to our homes. We can imagine that as we improve that last mile, more things will be stored in the cloud, everything from your home network and 802.11x, all the way to businesses relying on cloud services that will be spread out around multiple data servers. You're never going to be without email, your pictures, your data because it is replicated somewhere in the cloud. We are rapidly moving towards that.

Now, providing for that mid-range computing is where the science will go: We will be hosting climate data, earth data, and so forth, and we'll allow scientists to slice and dice and explore the data in the cloud here at Argonne and elsewhere. There will always be the need for a supercomputer that goes beyond commodity inexpensive computing to provide a capability that you can do in a high-end environment.

You probably have read stories about Google's data centers and how cheap they are. Google carefully calculates every penny spent, in terms of motherboards and CPUs and hard disks. For high-performance computing, we do that too, but on a different level, optimizing different variables. We're not optimizing the cost to run queries per second that can force an embarrassingly parallel search, but instead, we're trying to figure out how fast we can simulate 100 years of climate change. We need specialized architecture for that, and we will always need high-end architectures like Blue Gene that are low-powered but still massively parallel.

In the future, we will probably see a move toward cloud computing for mid-range capabilities like email, calendaring, and other services that businesses now sometimes host on site. And, we will also have space where we will solve the world's most challenging problems in climate, materials, biology or genomics on very high-performance machines like exa-scale machines.



## Bradley Horowitz, vice president of Product Management, Google, interview by Chris Heiden

*Bradley Horowitz oversees product management for Google Apps, including Gmail, Calendar, Google Talk, Google Voice, Google Docs, Blogger, and Picasa. Before joining Google, he led Yahoo!'s advanced development division, which developed new products such as Yahoo! Pipes, and drove the acquisition of products such as Flickr and MyBlogLog.*

*Previously, he was co-founder and CTO of Virage, where he oversaw the technical direction of the company from its founding through its IPO and eventual acquisition by Autonomy. Horowitz holds a bachelor's degree in computer science from the University of Michigan, and a master's degree from the MIT Media Lab and was pursuing his PhD there when he co-founded Virage. Here, he discusses the issues affecting cloud computing and how we must address them going forward. The ubiquitous nature of computing today is a perfect fit for computing in the cloud.*
—Chris Heiden

**Chris Heiden**: Describe your background and how it brought you into cloud computing.
**Bradley Horowitz**: Much of my research has involved looking at what computers do well and what people do well, and figuring out how to marry the two most effectively. Over time, that line has shifted, and computers now do many things that used to require a lot of manual effort. Much of it requires very large networked resources, and very large data sets. A good example is face recognition. These kinds of tasks are much more easily accomplished in the cloud, both in terms of computing power and data sets. And Google runs one of the largest—if not the largest—"cloud computers" on the planet. So it's great to be able to build applications that run on this massively scaled architecture.

More broadly, I've always been interested in the way people socialize and collaborate online. The most interesting part of cloud computing is in the network and the interaction between computers—not in individual computers themselves.

**CH**: What do you see as the single most important issue affecting cloud computing today?
**BH**: The network is getting exponentially faster. It's Moore's Law meeting Nielsen's Law. Not only can you pack tremendous computing power into a mobile phone to access the cloud anywhere, but bandwidth is growing exponentially as well, roughly 50 percent per year.

And it's starting to be available everywhere. At this point you expect the network to be fast and always-on. Airplanes used to be off-limits to cloud computing, but not anymore. WiFi is becoming standard on planes. People now buy phones expecting 3G data connections—talking has become the "oh yeah" feature at the bottom of a list of a dozen web and data features.

**CH**: What are some of the aspects of cloud computing that you are working on that will revolutionize the field?
**BH**: Everything "cloud" is really about collaboration. It's not just a larger computer—cloud computing gives birth to a new way of working, where the stuff you produce lives in the center of a social circle and can be collaborated on by everyone in that circle. We're trying to make that collaboration process more seamless, so sharing is easy and you can work on stuff together without worrying about the mechanics of version control and invites and the like.

> **"Lock-in and closed formats are a losing strategy. They make you lazy as a provider. And on the web, laziness is deadly, since competitors are a click away."**
>
> ***—Bradley Horowitz***

I'm not sure people have quite grasped how much of computing can and will shift to the cloud. It's not that there isn't ever-growing computing power on the desktop or in beefy all-purpose servers. But the part of computing people really care about—that developers are developing for, and that yields the most interesting real-world applications—is the networked part. As an example, you don't use Google Docs to write solo reports and print them out to stick on a shelf somewhere. You use it to get 10 people to hack out a plan everyone's happy with in one hour. It's the networked process that's revolutionary, not the app used in isolation.

**CH**: What about concerns people have voiced about trusting cloud computing? Do you see these concerns as slowing adoption?
**BH**: I'd flip that on its head. The providers actually can't keep up with user and customer demand. They can't build this stuff fast enough for what people want to do in the cloud. At universities, for example, it's students demanding their administrators switch to Gmail. Universities like Arizona State and Notre Dame have switched to cloud-based email, and we're seeing big businesses like Motorola making the switch, too.

It helps that the stats are finally starting to get out comparing apples to apples. Desktop and on-premises applications break down far more often than web apps, even if you don't hear about it as often since it happens behind closed doors. The nice thing about cloud computing is that all these service providers are so publicly accountable for even the slightest glitch. There's no tolerance for downtime, which is great for users.

Google's cloud is made up of a highly resilient network of thousands and thousands of disposable computers, with apps and data spread out on them across geographies. Gmail, for example, is replicated live across multiple datacenters, so if a meteor hits one, the app and data keeps on running smoothly on another.

And efforts like the Data Liberation Front and Google Dashboard make it clear that users maintain control of their data, even when they're using the cloud. They can take their stuff to go anytime they like, and they can always see how the data is being used. The value to you if you're a cloud provider is not in locking in users' data; rather, it's in the service, the flow of data, transactions, choices. Lock-in and closed formats are a losing strategy. They make you lazy as a provider. And on the web, laziness is deadly, since competitors are a click away.

**CH**: Describe how much of an impact cloud computing will have in the next evolution in computing. How will it affect the everyday computer user?
**BH**: This shift to cloud computing has already started invisibly for most users. Most people use webmail without realizing it's essentially a "cloud" app. That may be how they start using other cloud apps. They get invited to collaborate on a document online, and they use that doc every day, and then another, and one day they wake up and realize it's been months since they've opened up their desktop software.

But it's going to start getting more obvious as people switch to netbooks and smartphones, and as they stop having to worry about all the mechanics of backing of their discs, worrying about where they stored something, or hassling with document versions. It's already making daily life more mobile and more fluid. Even the word developer is becoming more and more synonymous with web developer. The web is now the primary platform for building apps, not an afterthought you hook into an app after it's built.

Cloud computing is already changing the way businesses, governments, and universities run. The City of Los Angeles just switched to the cloud using Google Apps. Philadelphia International Airport cuts down on delays and costs by coordinating their operations using Google Docs, and they're looking at technologies like Google Wave too.

And it's individuals, too. We hear about soldiers in the desert in Iraq keeping in touch with their families using video chat in Gmail, and checking their voicemail on Google Voice.

The movement away from the desktop and on-premise solutions has been sluggish for software makers with entrenched interests on the desktop. They're moving so slowly compared to what users want. These fast, lightweight, social, mobile apps are what people are actually using to communicate with each other and get work done. They sprouted in the cloud, they're built to run in the cloud, and they're now "growing down" into big enterprises, driven by user demand.