

# Storage Pods



Search:  Search

Members login

Between the Lines™

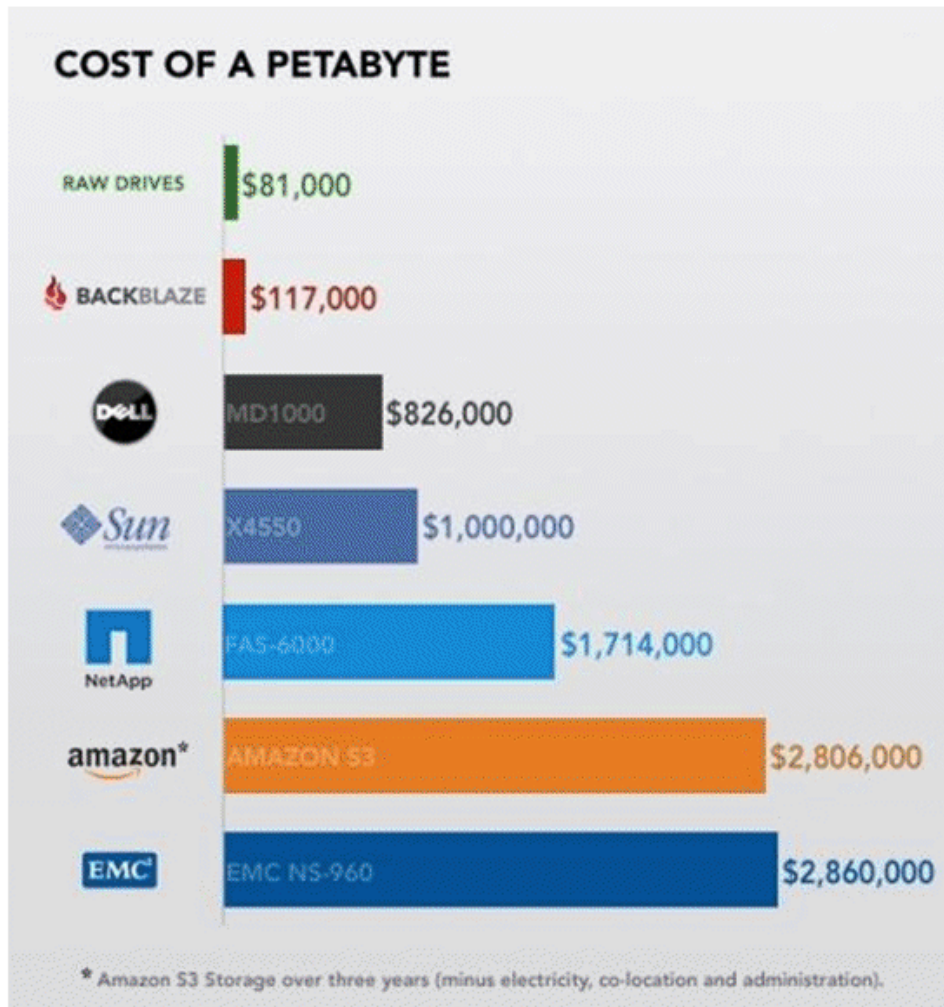
# Larry Dignan, Sam Diaz, Andrew Nusca

September 2nd, 2009

## BackBlaze reveals how to build petabyte storage arrays for a 10th of the cost

Posted by Jason Hiner @ 2:21 pm

Online storage provider [BackBlaze](#) has opened up its kimono and offered detailed specs on how it builds high-density 67TB storage arrays for less than \$8000 each. For IT departments willing to follow the BackBlaze formula, they can get massive storage systems for a fraction of what they'd pay Dell, EMC, Netapp, or Amazon AWS for the same amount of storage (see chart below).



The secret to the BackBlaze formula is that it uses a custom case, off-the-shelf hardware parts, and open source software. In a [blog post](#) on its site, the company provides a complete parts lists, diagrams that show exactly how to build one of its "Red Pods," and details of the software used to make it run. The 3D diagram below shows what a Red Pod looks like. Click play to watch the 19-second animation.



BackBlaze made this revelation in order to show how it can offer reliable service at such a cheap price (\$5/month for unlimited storage). While this blueprint can help IT departments build their own high-density storage arrays, it's doubtful that it will cannibalize any of BackBlaze's business, since BackBlaze is aimed mostly at making backup-to-the-cloud easy for home users and small businesses.

Thus, this is mostly just a publicity stunt by BackBlaze. However, I have to admit that it's a really good one. It can offer some real value to IT departments and create a positive brand impression of BackBlaze with IT pros, who tend to serve as the personal tech support department for all of their friends, family, and neighbors, and that often includes recommending services such as backup.

I'll admit that when I get asked about online backup providers by small businesses and individuals, the first two I tend to think of are [Mozy](#) and [Carbonite](#). With the Red Pod, BackBlaze has now shown up on the radar.



[Jason Hiner](#) is the Editor in Chief of [TechRepublic](#), ZDNet's sister site. Read his blog [Tech Sanity Check](#) at [hiner.techrepublic.com](#). You can also find him on [Twitter](#), [LinkedIn](#), and [JasonHiner.com](#).



# Mashable

## The Social Media Guide

Mashable Lists

Twitter

How To

Music

Travel

WordPress

Jobs

News Channels

Twitter

YouTube

Facebook

Google

MySpace

Video

Advertise

Jobs

Contact ▾

Partners ▾

Network ▾

☀ Spark of Genius Series

 Blippr

 Twitter Guide



## How It Works: The Continually Declining Cost of Unlimited Data Backup

September 1st, 2009 | by Christina Warren

 11 Comments

19  
tweets

retweet

  
post

We all know how important it is to backup our computers, but studies consistently show that most computer users — even those that know better — don't.

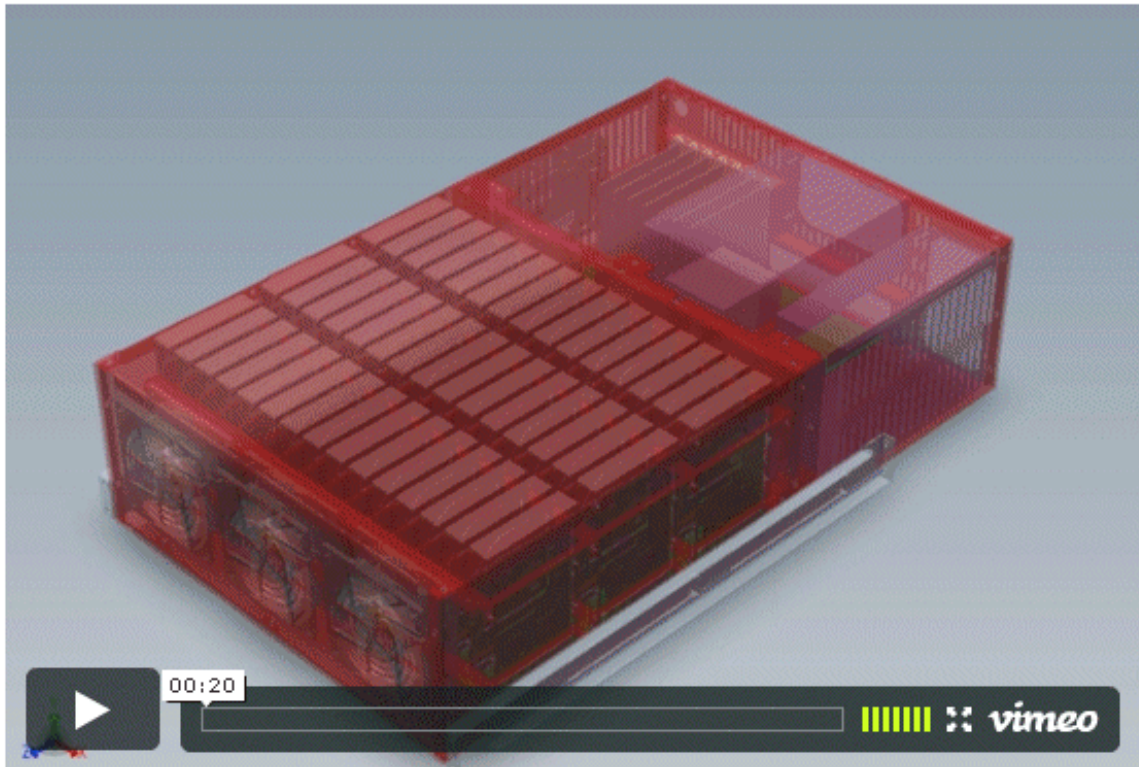
As Internet connections get faster and storage becomes less expensive, online backup services like Backblaze, Mozy and Carbonite have stepped-in to help consumers back up their data.

Not only are these services useful, they are also very inexpensive. Most of the services offer unlimited (truly) storage for your online backups for around \$5 a month. Now, granted, storage prices are constantly decreasing, but how can companies make those sorts of claims?

Online backup service [Backblaze](#) has decided to shed some light on how it manages to offer unlimited backup storage to its customers, while also easily expanding its own cloud-based storage platform. Today, Backblaze posted an entry in its [company blog](#) that explains what tools and hardware it uses to maintain cheap, efficient storage for its customers. Backblaze hopes that other businesses who are looking at building cloud-based storage solutions can use this information as inspiration or as a blueprint for themselves.



<http://mashable.com/2009/09/01/backblaze-online-backup/>



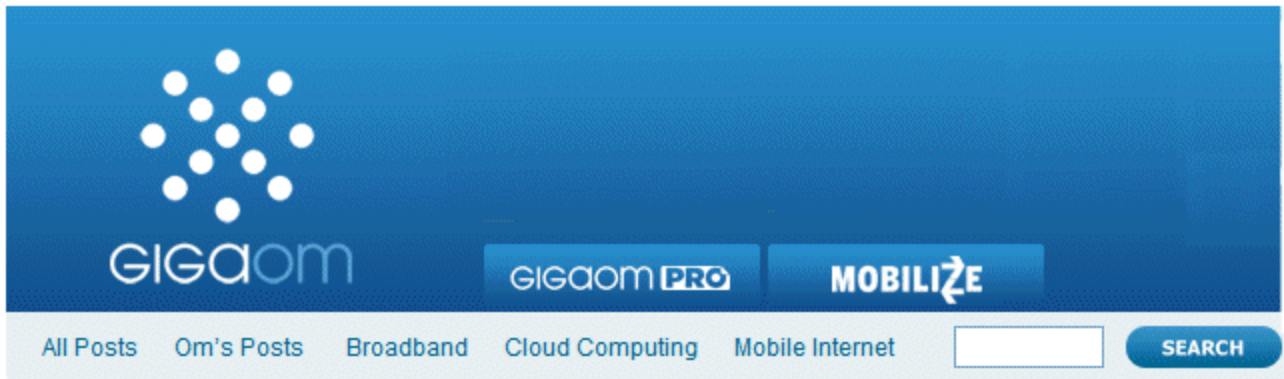
Backblaze has posted its special formula for how its pods work. Commodity hard drives and free software are the secret sauce that keeps prices low. I talked to Gleb Budman, Backblaze's co-founder and CEO, and he told me that Backblaze currently has about a petabyte of storage (that's 1000 terabytes or one million gigabytes). To put that in perspective, [Flickr's](#) entire library of photographs is about 1.5 petabytes in size. Gleb looked at using a company like [Amazon](#) or EMC, but for that kind of storage, Backblaze would have had to pay over \$2 million. By building their own storage pods, Backblaze's costs are under \$120,000.

As more and more businesses look to using the cloud for storage, cost-effective mass storage is going to become increasingly important. I think it's great that companies like Backblaze are offering some insight into how their systems work. As consumers, it's also nice to see how online storage systems work, and put the low-cost into perspective.

Do you back-up your computer? What methods or services do you use?

---

Tags: [backblaze](#), [cloud storage](#), [online backup](#), [the cloud](#)



The header features the Gigaom logo (a cluster of white dots) on a blue background. Below the logo are two buttons: "GIGAOM PRO" and "MOBILIZE". A navigation menu includes "All Posts", "Om's Posts", "Broadband", "Cloud Computing", and "Mobile Internet". A search bar with a "SEARCH" button is also present.

## Are You Ready for Open-Source Hardware?

By Om Malik | Tuesday, September 1, 2009 | 6:00 AM PT | 1 comment

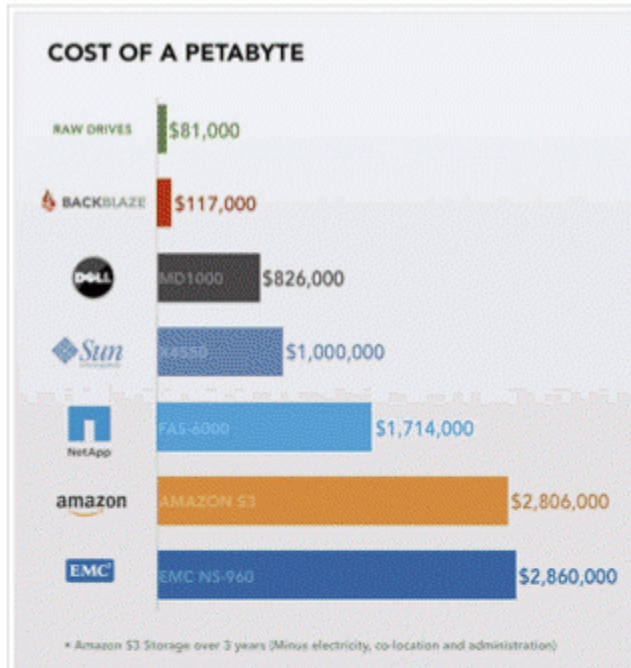


According to [the Chaos Theory](#), in a giant system that has lots of interconnections, even the smallest action can have a massive impact. It's more simply described by the [butterfly effect](#). This theory has taken its toll on the software business, thanks to the rise of open-source software platforms. Today, I learned about a move made by Backblaze, a small San Francisco-based online back-up service that can cause a similar disruption in the storage industry.

The company, whose primary business is selling online storage to consumers for a small monthly fee today, announced that it's giving away the design of its storage cluster for anyone to use, modify and build upon. The design allows anyone to build large storage clusters -- from a few terabytes to over a petabyte. What's so disruptive about this? What if I told you that you could build a petabyte-sized cluster for around \$120,000?

Now compare that to a couple of million dollars via a storage company like EMC Corp. or a server maker such as Sun Microsystems. The image below actually does a much better job of making a comparison between the Backblaze solution and other commercial storage options.





Actually if this works, companies like NetApp and EMC could be in trouble. Just like Linux slowly eroded away the premiums charged by the likes of Sun, these storage giants could see their business be negatively impacted. As the IT world transitions to cloud-based computing, the need for web-scale storage systems is going to increase. Google, for instance, has shown that you can build gigantic storage systems out of commodity parts and smart software.

In their GigaOM Pro report (subscription required), "Will Storage Go the Way of The Server?," analysts Juergen Urbanski and George Gilbert pointed out that: "The long-term future of storage is about smart

software that manages a large pool of cheap interchangeable hardware. Despite being one of the fastest growing technology sectors in terms of capacity, the economics for many participants are deteriorating."

Looks like Backblaze just wants to accelerate that by giving away its designs. "We are hoping that people who are in the hardware business will take our design and build devices by improving on our design and in turn selling large amounts of storage at an affordable price," Gleb Budman, CEO and co-founder of the company, told me during a conversation earlier this morning.

"At a fundamental level, we are a software company and we don't want to build hardware," he said. The company had to build its own mousetrap because it didn't have much of a choice. The Amazon S3 offering wasn't a feasible option, and the high-end systems from Netapp and EMC were way too expensive. Today the company has about petabyte and a half of storage space, putting it in the league of Facebook's photo storage system.



At Backblaze, we provide unlimited storage to our customers for only \$5 per month, so we had to figure out how to store hundreds of petabytes of customer data in a reliable, scalable way—and keep our costs low. After looking at several overpriced commercial solutions, we decided to build our own custom Backblaze Storage Pods: 67 terabyte 4U servers for \$7,867.

A Backblaze Storage Pod is a self-contained unit that puts storage online. It's made up of a custom metal case with commodity hardware inside. Specifically, one pod contains one Intel Motherboard with four

SATA cards plugged into it. The nine SATA cables run from the cards to nine port multiplier backplanes that each have five hard drives. [[The Backblaze Blog](#)]

Is this the perfect solution for everyone? Who knows? [Gary Orenstein](#), a storage industry veteran and a guest columnist for us, points out that the company has done a clever job of building a nice, lightweight wrapper around a bunch of drives to produce a system that is fine-tuned for Internet-driven uploads and downloads. Orenstein, who was co-founder of storage startup Nishan Systems, says that the big challenge with a system like this is managing drive failures over time and developing on the platform.

Nevertheless, this move by Backblaze is interesting because it addresses the current logjam in the hardware business. At our [Structure 09 conference](#), Facebook's VP of engineering, [Jonathan Heiliger](#), lamented how the chip industry and hardware makers fail to address the needs of the big spenders: web companies. Facebook had to build its own mousetraps to meet its specific needs. If your startup has open-source hardware designs that meet the needs of today's web-based businesses, you can easily do an end run around the incumbents.

Glubman said that the company is giving away design without any licenses because it doesn't want anything to come between the design and cheap storage. As long as companies keep innovating, the company is happy with the karma points it notches up. "We hope that the people actually contribute back to the community created around this hardware design," he said.

Will Backblaze's big dream come true? It's hard to say. What is safe to say is that if more companies start contributing their hardware designs in an open source manner, we can expect to see more innovation in the hardware business. So far, hardware innovation has been hampered by the high costs that go hand in hand with innovative products. A few more open-source hardware designs and we'll soon see tinkering minds get to work.

I would recommend you read our research report, [The Future of Data Center Storage](#). (\$79-a-year subscription required.)



# TECHNOLOGY BLOG



[Previous](#)

[Blog home](#)

## Petabytes on a budget

Backblaze was shocked by the cost of data servers and cloud-based storage systems, so it decided to build its own drive pods. But along the way, it has also raised interesting questions about the long-term costs of cloud computing, and how firms like Amazon can turn a profit

[Comments \(1\)](#)

[Buzz up!](#)

[Digg it](#)



Posted by  
**Jack Schofield**  
Wednesday 2 September  
2009 17.36 BST  
[guardian.co.uk](#)



**A** [larger](#) | [smaller](#)

**Technology**  
[Cloud computing](#) · [Amazon.com](#) · [Computing](#)

**More from Technology blog on**

**Technology**  
[Cloud computing](#) · [Amazon.com](#) · [Computing](#)

Just after I'd bought my first terabyte hard drive, up comes a blog post about assembling cheap petabytes of storage (1 petabyte = 1,048,576 gigabytes). Backblaze says it provides unlimited storage for only \$5 per month, so, [the post says](#), "After looking at several overpriced commercial solutions, we decided to build our own custom Backblaze Storage Pods: 67 terabyte 4U servers for \$7,867".

Like similar systems, Storage Pods are just a way of packing bunches of cheap hard drives together. And since Backblaze's photo shows them stacking pods, the built-in fans presumably provide sufficient cooling -- at least in the context of, I'd assume, an air-conditioned data centre.

Plenty of questions could be raised about Backblaze's approach: not just cooling but how reliable the Pods are, and how you're supposed to deal with failing drives. But the more interesting bit was why they decided to build, and open source, their own design. Tim Nufire writes:

Before realizing that we had to solve this storage problem ourselves, we considered Amazon S3, Dell or Sun Servers, NetApp Filers, EMC SAN, etc. As we investigated these traditional off-the-shelf solutions, we became increasingly disillusioned by the expense. When you strip away the marketing terms and fancy logos from any storage solution, data ends up on a hard drive. But when we priced various off-the-shelf solutions, the cost was 10 times as much (or more) than the raw hard drives.

The accompanying chart puts the price of a petabyte of raw drives at \$81,000, and Backblaze pods at \$117,000. "This translates to just three-tenths of one penny per gigabyte per month over the course of three years." A commercial alternative could cost \$826,000 (Dell) to \$1m (Sun).

The scary bit is that using Amazon's S3 cloud computing service would cost \$2.8m over three years, and I don't think Amazon is the high-price supplier. (Microsoft's Azure, for example, costs more.)

<http://www.guardian.co.uk/technology/blog/2009/sep/02/backblaze-petabytes-cloud-computing>

Twice ever thus, of course. If you're old enough to have rented a TV set or a VCR, for example, or a car or a flat, you'll know that the initial costs are relatively low, but become infinitely high over an infinite time period. Renting avoids the up-front cap-ex because you don't have to spend \$1m with Sun, but you will inevitably pay much more for the hardware in the long run. If you didn't, the cloud company would go bust (possibly taking your data with it into oblivion -- or possibly to Nigeria).

Cloud computing companies can make some cost savings from their much greater utilisation of resources, and that's certainly true for processing power and memory. But they can't save much on hard drive space without moving your data offline, perhaps to tape, and you could do that yourself. Basically, your data fills just as much space in the cloud as it does on your own hard drives.

There is, of course, much more to cloud computing than there is to online data storage. The problem is that you really can't do much with cloud computing unless you also store your data online, so you don't have much flexibility in that respect.



And as we know, when you don't have a choice, that's exactly where you are going to get overcharged.

---

[Previous](#)

[Blog home](#)

---

 [Printable version](#)  [Send to a friend](#)  [Share](#)  [Clip](#)  [Contact us](#)  [Article history](#)

---



## Why Netflix's CDN should scare the storage industry

Derrick Harris Jun 6, 2012 - 12:22 PM CDT



Lest storage vendors thought they were immune to disruption that open source hardware is having on the server industry, Netflix's (s nflix) [new Open Connect content-delivery network](#) might make them think again. While Open Connect directly targets commercial CDNs, it's [based upon \(or at least inspired by\) open source storage designs](#) first released by Backblaze [al-](#)

[most three years ago](#). Backblaze's design evolving and expanding its range into the data centers of a Fortune 1000 company is significant in the same way the evolution of modern man was for neanderthals.

By way of background, Backblaze is a cloud storage provider focused solely on backing up lots of data for cheap (like [\\$5 a month for unlimited capacity cheap](#)). In order to do that, it had to build a storage system that could hold massive amounts of data without breaking the bank. As of last July, Backblaze's architecture had [evolved to a point where a 135TB pod cost less than \\$7,400](#) to build from scratch.

Understandably, the architecture generated a lot of interest from companies and organizations wanting to leverage it to soothe their own IT budgets, but none of them are Netflix. EMC's (e emc) Pat Gelsinger said recently that the storage component of Facebook's (s fb) Open Compute Project, [called Open Vault, isn't yet ready for primetime because nobody is running](#) — or would run — mission-critical workloads on it. That might be true of Open Vault today — the project just launched earlier this year — but it likely won't be for long. If you consider a CDN that serves Netflix streaming video mission-critical, the criticism is already invalid for Backblaze's designs as Netflix has adapted them.



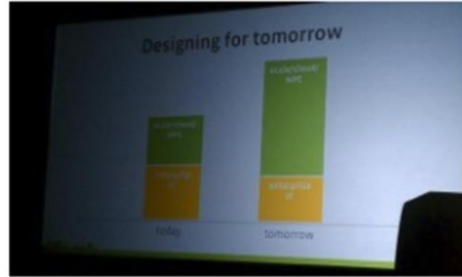
Netflix's 4U, 100TB server

It's worth noting, too, that open source hardware isn't the only piece of the stack threatening legacy storage vendors such as EMC. I've heard it suggested recently by someone experienced in building out large-scale cloud infrastructure that the Hadoop Distributed File System has the potential to become the default file system for large infrastructures once it works out some of the limitations around performance and availability. [One of the biggest of those limitations — the NameNode](#)— has been eliminated in the latest version of Apache Hadoop and is already integrated into Cloudera's new CDH4 release.

## Can storage deal with the open source disruption?

As with [Open Compute's effects on the server industry](#), though, open source storage doesn't need to spell doom for legacy vendors if they're willing to adapt. One reason is that, at least in the short term, there are still plenty of customers that don't operate at Facebook or Netflix scale and can afford to pay a premium on smaller deployments that offer the features (and vendor support) those customers demand.

If the server shipments tell us anything, though, it's that the rise of cloud computing and web giants will ultimately take a toll on the storage market, too. Fewer, but very large, customers will be responsible for a greater percentage of sales, and they won't necessarily want all the bells and whistles that make enterprise storage



products so expensive. And if VMware (s vmw) is correct, even mainstream enterprises will soon want to follow the examples of web giants like Google (s goog) and Facebook [by running relatively dumb hardware managed by really smart software](#).

If this scenario plays out, storage vendors will have to reassess how they deliver value and earn their money. That might mean adopting open source designs in their own gear while shifting their focus a lot more heavily toward software and services, or perhaps unlocking their storage-management software from the hardware and certifying it to run on open source gear.

Perhaps we'll get some ideas for what the future storage and markets look like at our [Structure conference June 20 and 21](#), where we'll dive into the topic with Facebook's Frank Frankovsky, Netflix's Adrian Cockcroft and VMware Steve Herrod. Whatever the case, it looks like something will have to give.

*Feature image courtesy of Shutterstock user [Zadorozhnyi Viktor](#).*