



For Cloud Professionals, part of the On Cloud Podcast

David Linthicum, Managing Director, Chief Cloud Strategy Officer, Deloitte Consulting LLP

Title: Matching cloud costs to need: Using FinOps for more effective cloud governance

Description: In the rush to support business continuity these days, and to support employees while they work remotely, many organizations have ramped up their cloud usage. However, some may be overprovisioning and using more resources than they actually need—which can result in huge cost overruns. In this episode of the podcast, David Linthicum and guest, David Bernstein, discuss how organizations can use the emerging discipline of FinOps to help them govern cloud usage and more effectively match usage to need to cost to value. Bernstein's advice is to relentlessly monitor usage, work with vendors to optimize services, and take advantage of both cloud's automation capabilities and new financial observability tools that have recently come to market.

Duration: 00:21:46

Operator:

This podcast is produced by Deloitte. The views and opinions expressed by podcast speakers and guests are solely their own and do not reflect the opinions of Deloitte. This podcast provides general information only and is not intended to constitute advice or services of any kind. For additional information about Deloitte, go to [Deloitte.com/about](https://www.deloitte.com/about). Welcome to On Cloud, the podcast for cloud professionals, where we break down the state of cloud computing today and how you can unleash the power of cloud for your enterprise. Now here is your host David Linthicum.

David Linthicum:

Welcome back to the On Cloud podcast, your one place to find out how to make cloud computing work for your enterprise. This is an objective discussion with industry thought leaders who provide their own unique perspective around the pragmatic use of cloud-based technology. On the show today, we have David Bernstein. David is the Managing Director at Cloud Strategy Partners, where he has helped dozens of software-as-a-service companies with their cloud platforms. David was an early part of the cloud computing community via Cisco, AT&T, NIST and IEEE, where he co-authored many of the first research papers, specifications and standards for cloud before cloud was cool. So David contributed extensively to the early development of UNIX at University of California. Pretty impressive, David. You and I have a very similar background. So what was it like being into the early days of cloud computing?

David Bernstein:

Well it was fun. I mean, cloud computing is kind of a new platform and it's really taken over here over the last ten years, obviously, but when things were just getting started, it was quite a revolution then. People had a lot of doubts about it, of course, but now I think it's powering the world. Pretty much if you're using a mobile app or a web site, you're probably using cloud.

David Linthicum:

So what's your day like? And I know it's changed recently now that we're going through this challenge, but what do you typically do? How do you talk to clients? How much time do you spend on research? How much do you spend time on writing?

David Bernstein:

Well, especially lately, a number of cloud ops people have been kind of remote for a long time. I mean, nobody goes and sits with the computers in their cloud data center much anymore, especially in a public cloud. You don't even know where the building is. But nowadays I spend a lot of time kind of continuing to help clients with their compliance, with moving to the cloud, and especially in FinOps now is a big topic as more and more people go to cloud, and they're seeing how to keep their costs in line with the footprint they're delivering. So that's been pretty recently that now the cloud has exploded, the whole business of cloud has become very important.

David Linthicum:

Yeah, and I think some of the pragmatic things that we're trying to relay to people out there who are going through these specific challenges right now, and how it's affecting IT currently and things that we can do to, in essence, mitigate some of the challenges and ability able to get through the other side, and I think in many instances, there may be a silver lining here that I think enterprises will figure out where their weaknesses are and take steps to shore things up. So ultimately, how are we getting to tooling, how are we getting to a cloud footprint, how are we getting to cost and efficiency? Ultimately, if they're not ready, what kind of things should they be doing right now, David?

David Bernstein:

Well, first of all, we're all very thankful for our IT people now with all the remote access and remote resources we've been using, so that's a shout out to all the IT people who are keeping the wheels on the bus here for us. And likewise, all of the ops folks who are keeping all the cloud infrastructure running now that we have lots more remote demand. That places lots of demand on the cloud infrastructure, and so this is something I think that companies have dealt with, are dealing with. They see that their planned budgets for cloud spending are probably insufficient because of the amount of remote work. So the first thing I think companies are getting together and figuring out, gee how should we cope with this. And so the culture aspect is sort of number one. Getting together with finance and saying, okay, we're serving this demand, let's make sure that we're looking after what finance is concerned with, which is how much all this costs.

David Linthicum:

So what should companies be doing right now in terms of tactical, in terms of fixing things that are currently in front of them in terms of IT concerns, and also how should they think strategically, at least short-term strategically?

David Bernstein:

Well cloud is all about agility, and so what you've got to do is you've got to respond quickly to what's going on and you've got to get not too far over your skis. And it's no different with FinOps here. So what we've got is a little bit of a different process with how rapidly things are changing. We want to keep the lights on, but we don't want to overdo it, so the first thing is get really friendly with your tools and start looking. Be information driven. Look at how your usage is going up and look at how your costs are going up, and clearly, if the costs are going up faster than the usage is going up, that's one point I would point people to is take a look at how your different contracts are set up, take a look at how your different prepays are set up, make sure that you understand that – to have your costs go up kind of linearly with your usage. That's the first thing right there is to know how cost is generated and how that corresponds with the actual usage.

David Linthicum:

So I like that as a key metric. And so, in other words, if we're monitoring costs, we're monitoring usage, if the costs are going up and accelerating beyond usage, then we have a potential problem, and we can get ahead of it by spotting that that problem is there. I mean, I know a lot of people kind of wait for the public cloud bills to come at the end of the month, and that's when they kind of figure out that they made huge mistakes in terms of their efficiency to consume this technology and other technologies. And so at the end of the day, is it going to be cost governance that really kind of comes into play, the ability to kind of monitor usage of systems, who's using it, what they're doing with it, how much we're paying for it, and the ability to have kind of a single pane of glass into where the trends are occurring, the ability to put policies in play to make sure we're not getting out of whack between the cost versus the value we're getting back from it? What are some of the advantages there?

David Bernstein:

Yeah. I mean, absolutely. This is still a little new. FinOps, which is kind of a new term, has really just sort of emerged within cloud ops teams. I mean, there's not a cloud op team that doesn't pay attention every day to what's happening, but there's been a lot of observability – technical observability tools that have come out. There are fewer—or it's a little more avant-garde—to talk about your sort of financial observability tools, and I think that that's something that people are really starting to pay close attention to, and there's some great tools out there. So teams should definitely start there and definitely be paying attention every day as a start.

David Linthicum:

So I get calls from clients and they're, like, we have the issue in front of us where we're trying to get ahead of what we're spending now that we're reacting to a particular situation which is kind of changing how we look at the market. So what are some of the maybe tactical things they can do that don't involve installing tools in the short term to kind of get ahead of this? I know manual processing's out there, things like that, so what would you tell somebody that has not gone the cost governance, the FinOps, route in the past, is looking to do that now, and probably is going to implement that over the next six months to a year but is looking for some solution that they can do today?

David Bernstein:

Well, I would say first of all, we always want to try to keep the agility and the responsiveness of engineering teams in place. So we want to do as much as we can with a carrot and not a stick. And so that means enabling visibility to the teams, working closely with them, and seeing what they're doing to deal with the capacity fluctuations. The first thing is, when you get a lot of demand, ops teams will go ahead and take out the big hammer and just expand footprints to keep the wheels on the bus. And I think that's a great tactic to make sure that we're delivering to customers who are using, end customers ultimately. But on the other hand, the center of excellence, the FinOps folks, you need to keep an eye on that and take a look at what the ops teams have done in a crunch first, because sometimes they expand footprints and then they just leave them there. So that's the first thing I would do is make sure you understand what's – how much gas was poured on the fire to start with there and get a handle on that. That's kind of the low-hanging fruit right there.

David Linthicum:

What do you think the three top mistakes are that are going to be made by enterprises that are reacting to the crisis that we're going through right now that probably they didn't need to make? I'd just be interested in your perspective.

David Bernstein:

Well, I think the first thing is overreacting. I mean engineers typically will go ahead and just increase the sizes of everything to make sure the system keeps working, so that's an okay move to stabilize, but then at this point, you have to constantly keep an eye that you're not just turning a knob all the way to the right is the first thing. The second thing I think is to just, again, understand what generates cost and where you might get into really tricky situations. Sometimes when you get outside of your contracts, you get into overages, you get into pure on-demand kind of payments, suddenly you're paying two, three, four times as much as you should be.

So I think contacting your vendors and understanding how to work together to make sure you understand the costs and footprint. And of course, the last thing is don't set it and forget it. Keep an eye on this and keep chipping away and do this in a way that you're in partnership with your stakeholders. When you're in kind of territory you're not familiar with, well then everybody should get together and work it out together. Finance, engineering, the services business units, make sure you're talking to each other so that you know the territory you're in and how to solve these anomalies as you find them. That's what I would say.

David Linthicum:

Yeah, I think that's great advice. And I think that what's going on right now, if you look at the stats with the public cloud providers many times that their demand has gone up 1,000 percent. The enterprises saw that we have a situation coming up where people are going to have to work remotely, we're going to have to move our applications and move some of our data in some sort of a centralized system that's managed by other people, so we don't have to maintain as much dependency on a physical data center because people may not be able to get into it all the issues are with a physical stop. And so they just start – you alluded to this in your last response – they just start provisioning storage and databases and putting things online without pragmatic uses for those things. So what are some of the best practices, what are some of the processes that they can put in place to ensure that they don't panic provision?

David Bernstein:

Right. Well, first of all, let's not forget one of the things we all love about the cloud is the elasticity. So let's start with where we can, on the autoscaling. Take another look at autoscaling around both your compute instances as well as databases autoscale now too, and so this is something that you can do that will help. The work from home scenario is very dependent on what time of day it is, and has some different demand curves around the time of day than previously. And so really take a look at autoscaling and make the cloud elasticity work for you is a place I would definitely get started.

Also, you don't have to take the same hammer to every different region. Some regions are differently affected or have different time of day or have a lot of demand and then maybe their situation will change in a different time frame than another region as countries respond differently and have different remote work policies. So take a look at each region separately, and try and do something appropriate to that region. So really kind of start slicing and dicing the reports and things that you get from your tools more granularly based on per hour, per region. Really look hard at the places where you could get some elasticity, and that's kind of a cloud 1A place to start.

David Linthicum:

What about emerging technologies. It's been around for a while at this point, but things like serverless, and now we have serverless container development, we have the ability to do container clustering with Kubernetes and other technologies out there, really kind of advanced ways where we can build the scalability into the application infrastructure, even into the applications themselves. What are some of the cost advantages there and what are some of the tradeoffs?

David Bernstein:

That's a great point, because those are inherently elastic services, but not every application is set up to use serverless right out of the chute, but maybe it's a nice time to start thinking about that. On the other hand, many of the infrastructure capabilities that we have, people have probably gone into containers, but maybe they really haven't looked closely at the packing policies and what kind of instances are underneath there and the autoscaling. Maybe they don't realize that there's even serverless databases that can come on and, in a transient way, address some capacity needs and then go away. So these are things that folks may not have realized that their cloud provider has come up with some pretty cool stuff in terms of elastic demand in the last year or two. Take another look at some of that, I would say. And also, you don't have to get too crazy. Sometimes people just maybe haven't really surrounded some of their databases with your kind of standard caches. These are things that you can do to help your hardworking databases without radically changing things.

You can use some of your tooling around some of the key engines that you're using, take a good look at them, understand if they are overprovisioned or if they can use some smaller or different types of instances. I guess really the devil's in the details, but there's plenty of tools that you could spend a little time looking into each one of these and get some help from your vendors or your experts that you use in each of these different engines. Each one's a little tricky but there's a lot of resources to get the most bang for the buck out of each one.

David Linthicum:

Yeah, I posted a blog last week talking about some of the trends in overengineering some of the applications that we're building out there, certainly the net new stuff. The thing with cloud is – and the great thing about cloud is we have so many services that we want to basically bind into our applications because they're cool and they may provide key functionality. Artificial intelligence based systems, data analytics, predictive analytics, the ability to leverage

serverless technology, the ability to containerize everything and put that into some sort of a cluster, or a container orchestration system such as Kubernetes, and all these things which people have a tendency to what I call manage by magazine, but now there are no magazines around, so it's manage by blog post, where they're moving in ways in which they kind of view them as being part of the cool kids, and ultimately they're overspending on what the applications should be doing pragmatically. What advice do you have there?

David Bernstein:

Well, this is the desire of some engineering teams to use the latest shiny object to put into their infrastructure because they want to learn about that, or it's always a risk, but I would say by and large people have a good reason for wanting to use a particular type of tool or particular type of database or service, but the real key is always – don't get attached to anything. Always ask the question. Look back at what your predecessor might have installed, and maybe it's time to move off of some things onto the newer technologies. Oftentimes we just add new things about deprovisioning the old, and so you end up with eight different NoSQL databases instead of just the one or two you really need.

So I think that it's a great time for folks to take a look at what is – what they're really using, what really is generating cost. Because as the demand goes up, these things that are kind of around, they start generating more and more costs, and some of them are third-party services and have different licensing schemes, so you expand that cluster, you have more network traffic, suddenly you're into overages. And that's one thing I'd like to point out is that it's not always – sometimes the part that gets you is not the part you're looking at. So with cloud, when your core increases, there's all these follow-on parts to it as well, all of the third-party services you have wired in if you use someone else's logging system or someone else's performance management system or someone else's – and you put all this stuff together.

As your cloud grows bigger, suddenly you're putting more demand on all these, shoveling more logs, generating more CDN traffic, generating – and so pay attention to those other services too because they may actually have a bigger cost increase than your core cloud footprint if you're not careful, because of the way those are typically purchased.

David Linthicum:

Yeah, the analogy I use in the blog – I don't know if you have a motorcycle or not. I own a few, but the thing is you can't leave a motorcycle stopped. There needs to be some law of nature where people are going to go out there and get aftermarket exhaust and handlebars and seats and engine control modules, things like that. End of the day, most people spend about one-fourth the cost of the motorcycle and they don't do anything to the motorcycle to make it run better. And so, in other words, they're not getting a return on investment for the spending of all these gizmos that they can do. In other words, they're not asking themselves should we do it. They're asking themselves can we do it.

And I think the same thing with applications. I think that ultimately I understand that the cloud is revolutionizing our ability to leverage next generation technology at a very inexpensive price and also very easy to find these things in particular applications. There really should be a projected need for these things, and at the end of the day, it's not necessarily the operational cost of these things taking up a lot of money and generating large cloud bills. It's a fact of the matter is that applications and deployments and solutions are much more complex than they need to be. And the complexity is going to cost money, because we have to put it in an operational state, we have to have people who maintain it, we have to have people to fix it. And so we're not necessarily focused on the long tale of this. What are your thoughts on this?

David Bernstein:

Well, I know in the old days we used to have kind of a rally cry around cloud—this will bring you back—of simplicity scales, and I'm sure you remember that. And I think to some degree, that's been a little bit forgotten, and that is that some of the – especially now when we have some increased demand on the systems. This complexity is tough. We're getting into some cases where people are doing their instance management and so on and they find out, boy there's a lot of moving parts to this. It is a time to take a look and say gee, can I do this in a more elegant way, can I eliminate some stuff that is just complexity? Or I put these cool features together, and now that my system is suddenly scaled up, these cool features aren't working exactly the way I thought or they're generating a lot of side effects, or they're generating a lot of cost. I think the original rallying cry around cloud, of simplicity scales, and will keep us honest. And so I think that's a good time. People are probably realizing, wow, when this scaled up, it got really complicated and entropy is setting in. What can I do to survive a bit better? Well, get simple. That's a great rallying cry to resurrect, I think.

David Linthicum:

Yeah, and have a culture of frugality. I mean, I think both of us have played the role of the designated buzzkill. Obviously when everybody wanted to move in a more exciting, cool direction, and ultimately have to kind of look at what business problems you're looking to solve, which is really what we need to be asking ourselves. At the end of the day, we're building IT systems, whether in the cloud or not, to serve the needs of the business, and that's our role and that's our responsibility. So where do you go to get your information on what's new in the cloud, David? So what are some of the resources you hit everywhere.

David Bernstein:

Well, clearly there are some folks that create great podcasts and blogs, so I'm glad to contribute to that. There's a number of other great companies that have a lot of exposure to customer projects and share their learnings there. Most of the vendors put out great information. I would say that, nowadays, we're going with fewer conferences and reading a lot more. That's fantastic. I do think there's also a good community of people that you can go to and share.

For example, in this area, there's kind of a new community, finops.org, rallying around kind of the ideas of how to do cloud financial management as a kind of first-class discipline, and that's been a great place to just swap notes with some other experts. I would for sure tell people who are part of the cloud ops team and looking after this whole scaling and cost ideas, shine a light on things, realize that this is a really complex problem. It's a little scary if some of the answers come down to prepaids, contracts, and readjustments, but the watchword in this is, do something. And then kind of adjust tomorrow and then adjust the next day and then adjust the next day. That's the whole philosophy around cloud, and that is the same philosophy around this when we've got this extraordinary situation with a lot of cloud growth going on. Keep an eye on it and keep fiddling with your tooling and with your team, and you will improve the situation.

David Linthicum:

Yeah, I think the key advice there is to be a continuous learner. So ultimately there's a lot of information out there and things you can maintain. So where can the audience learn more about you?

David Bernstein:

Well, I have a blog, and I also am more than happy to chat with folks, if they can drop me a note at david@cloudstrategypartners.com, and I'd be happy to talk to people about this, how cloud is the new computerization platform of the planet, and we've come a long way, David, in the 10 or 15 years you and I have been working on this, and I really do think that as this trajectory continues with 5G, it's only going to accelerate how inherent cloud becomes in our society. I'm proud to be part of the plumbers making it work.

David Bernstein:

Yeah, I'm proud to be part of it too. I think we're going to learn a lot from this crisis that we're going through right now. So if you enjoyed this podcast, make sure to like and subscribe on iTunes or wherever you get your podcasts. Also check out our past episodes, including On Cloud podcast hosted by Mike Kavis and his show, Architecting the Cloud. If you'd like to learn more about Deloitte's cloud capabilities, check out deloittecloudpodcast.com, all one word. And if you'd like to contact me directly, you can reach me at dlinthicum, that's L-I-N-T-H-I-C-U-M, @deloitte.com. So until next time, best of luck with your cloud computing projects. We'll talk again very soon. You guys take good care. Stay healthy, stay safe.

Operator:

Thank you for listening to On Cloud for Cloud Professionals with David Linthicum. Connect with David on Twitter and LinkedIn and visit the Deloitte On Cloud blog at www.deloitte.com/us/deloitte-on-cloud-blog. Be sure to rate and review the show on your favorite podcast app.

Visit the On Cloud library

www.deloitte.com/us/cloud-podcast

About Deloitte

The views, thoughts, and opinions expressed by speakers or guests on this podcast belong solely to them and do not necessarily reflect those of the hosts, the moderators, or Deloitte.

As used in this podcast, "Deloitte" means Deloitte Consulting LLP, a subsidiary of Deloitte LLP. Please see www.deloitte.com/us/about for a detailed description of our legal structure. Certain services may not be available to attest clients under the rules and regulations of public accounting.

Please see www.deloitte.com/about to learn more about our global network of member firms. Copyright © 2020 Deloitte Development LLC. All rights reserved.