

EPISODE 1441**[INTRODUCTION]**

[00:00:00] JM: Vantage is a system for optimizing cloud costs. It provides tools and interfaces for developers to analyze how they're spending on AWS resources, and has recently expanded into GCP as well. Vantage users gain an easy interface into their costs that would otherwise be hard to analyze via the raw AWS console.

Ben Schaechter is the founder of Vantage. In a previous show, he gave an overview of Vantage, and in today's episode, he returns to discuss the engineering internals of Vantage and provide a macro perspective on where people are wasting money in the cloud.

If you're interested in reaching 250,000 developers per month, check out Software Engineering Daily for sponsorship, you can reach us at sponsor@softwareengineeringdaily.com, to reach our audience.

[INTERVIEW]

[00:00:46] JM: Ben, welcome back to the show.

[00:00:48] BS: Hey, thanks. Good to be here.

[00:00:50] JM: Last time, we talked about the basics of Vantage and some cost management elements. I'd like to start the conversation by talking about some macro trends in AWS usage, starting with specific cost management errors or areas of waste. Where are the biggest places where you see waste in AWS deployments?

[00:01:15] BS: Yeah, for sure. So, I would say this kind of ranges depending on the specific customer and what's going on. But typically, what we see is when people use some defaults provided by AWS and may not understand all the options, they have available to them. This is an area where they can run into an issue. So, let me give you a few different examples. By default, the majority of customers will use on demand EC2, and if you aren't familiar with on

demand or spot or reserved or savings plans, essentially, the thing that you may not know at the end of the day is that if you can commit to some specific usage, you can yield some fairly significant savings on EC2. And what this looks like is essentially, you just have to know the certain amount of instance hours or if you want to take a super basic example, let's say you have a single EC2 instance, you know it's going to be on for the full month, and let's say you think it's going to be on for the full year, you can do an upfront commitment, where there's actually no prepayment, and save between 20% to 30%, or sometimes more, at the minimum commitment level on EC2.

So, those tend to be like the largest areas where we can find savings or lack of optimization, but the same kind of thing applies to a variety of AWS products. So, AWS offers the same committed usage model for RDS in the form of reserved instances, with EBS volumes, you may be using the wrong volume type, and with a slew of other products. S3 is another common case where every single file that's uploaded typically by default is done with standard storage. However, there's a variety of options that you can change or tweak that can yield savings. I would say, just to distill this down, the first thing that we tend to see as unoptimized is compute. That's usually the largest area of the bill in terms of what comprises the bill from a customer perspective. But then, services two, three and four things like EBS, S3, RDS, those are other areas where there's kind of these micro optimizations that you can make as well.

[00:03:25] JM: And when the average customer sees a glut in spending on those raw EC2 instances, do they typically need to make some drastic change, like, totally rearchitect their application? Or is there some kind of just configuration change in EC2 that they can make?

[00:03:46] BS: Yeah, great question. So, actually, we see this as two different things that the user can do. The first thing is, there are financial engineering implications where they don't actually have to change anything related to their infrastructure to yield some of these savings. And the two examples of this are AWS savings plans, which is essentially AWS is modern answer to committed use discounts. There's another option that's reserved instances, which has greater mindshare out there. In concept, it's the same thing as AWS savings plans, but gives less flexibility. But these two things are things that you don't actually have to make any infrastructure changes on, you don't have to change anything about your architecture, and you can yield significant savings.

So, this is usually low hanging fruit that we recommend customers take a look at. In terms of a second swath of things that you can do from an optimization perspective, there are things that do require changes on the customer's infrastructure to be able to yield some of these savings. For example, with S3 storage types, if all of your files that are in an S3 bucket are defaulted to standard storage, but for example, you want to move that from one storage class to another, or AWS offers a service called S3 intelligent tiering, where you can have AWS monitor and shift these files from one storage class to another. They do require changes, typically configuration changes, or one-time migration changes. And then beyond that there's likely other larger architectural things customers can do to really move the needle on cost savings or cost optimization.

[00:05:25] JM: Can you go a little bit deeper into how a customer typically responds after seeing gluts in expenses when it comes to compute or storage?

[00:05:37] BS: Yeah, I can. The thing that we typically see in the value prop that Vantage provides is really first and foremost, visibility of costs. So, typically, the user flow from the Vantage perspective is a customer will sign up, they'll get integrated with Vantage. These days Vantage will automatically provision a number of, for lack of a better word, business intelligence dashboards that give you visibility into your infrastructure, both by AWS service, as well as resources and even understanding where my costs are coming from.

The optimization part of things is kind of twofold. Vantage has a devoted section on the console, that's called a cost recommendation. And this is our automated way of always profiling your infrastructure and surfacing and recommending things that you can do specific to your infrastructure. So, an example is, we'll take a look at all of the EC2 instance hours that are running, and very specifically recommend to you what AWS savings plan you should buy.

Another example is if you're using the wrong storage type in a specific area, will surface that to you. But then there's a number of things that people get value out of just in terms of diving deeper into, "Hey, we had an infrastructure change go out, let's see how that impacted things from a cost perspective." And they'll go into Vantage and see what the impact is, both in terms of cost going up and cost going down. And then there's kind of the story that we see pretty

regularly these days that we refer to as cost of bugging, where, at least at the time that we're recording this podcast, Vantage is the only provider that's able to give per resource specific drill downs into category and subcategory costs.

I'll give a really clear example, which is, let's say you have a giant RDS instance that you need to take a look day by day of the costs that are coming from provisioned IOPS versus storage snapshots, versus the requests that are coming in, or the underlying box that you're paying for. Typically, what we see the value that people get of Vantage is they'll see their cost go up and dive into one of these resources and see specifically what caused that cost to go up.

So, "Hey, did we accidentally misconfigure something and it's getting a lot more requests? Or did we provision too much in the way of provisioned IOPS or things along those lines?" So, in terms of the original question that you had, it's kind of twofold. There's the visibility that people can get and analyze things from where they're at from a cost perspective, and optimization is not always top of mind, which may seem counterintuitive. They just want to make sure things are trending generally in the right direction. And then when you do have these occasional check-ins on optimization, you have the automated recommendations from Vantage to at least have a head start on where to go from there.

[00:08:31] JM: Let's talk a little bit about engineering. We did an overview this in the last episode, but I'd like to go deeper today. So, when I plug into Vantage, what's the onboarding process? And what's going on beneath the hood? How is Vantage interacting with my AWS infrastructure?

[00:08:50] BS: Yeah, for sure. So, this is an area that we're just, I would say, we have a best in market experience, and we've written this all up on a blog, actually multiple blog posts, I encourage people to go and take a look. So, the way that vantage works, is it is entirely self-service, you can come on board, you can register an account, you can verify your email. Technically, the way that we integrate with accounts is through something called a cross account IAM role. That cross account IAM role essentially allows us to mimic being you for your AWS account with a certain set of permissions. These permissions are very basic at the end of the day. They're all read permissions, largely taking a look at the ability for us to essentially call cost explorer and listed describe API's on your behalf.

So, what Vantage is doing is we're pulling in information from cost explorer, we're creating what's referred to as a cost and usage report integration, which essentially allows us to get the raw billing data and then we're profiling all of these direct AWS service API's. So, listing EC2 instances, EBS volumes, S3 buckets and things along those lines, and this allows us to show you a few things. The first is where are my costs going high level, which we get from costing usage reports and cost exploring certain cases. But then the next question that we can answer is, okay, I know my costs are going up, why? And with our service integrations, which were just kind of always profiling in the background, we can say, here's where your costs are going, and here are the specific resources that are causing these costs to go up. By the way, you can drill down into these specific RDS instances, or you see two instances or whatever is in your account, to kind of join together all these AWS primitives, and give you a really high-fidelity way of interacting with your costs.

So, happy to dive deeper on more specifics around this cross account IAM role, but that's technically what's happening behind the scenes.

[00:10:51] JM: Yes. Does that role get full permissions into being able to see all the internal usage and spend?

[00:11:00] BS: Yeah, I want to make sure I answer this correctly. So, we're not getting full access to your accounts. It's actually the way that we create this role is we give you a CloudFormation template, which you can go in audit, and on our documentation, it has a full list of permissions. There's a very long list of permissions. But the reason for that is it's largely looking at a whittled down list of IAM permissions, that only gives us the ability to see cost data, and then metadata about various resources. What these maps to in the AWS world are, typically list and describe actions. So, for example, we can see what S3 buckets are in your account, but we can't read from those S3 buckets. Or we can see what RDS instances are in your accounts, but we cannot query the actual underlying database.

So, the permissions that we're getting essentially allow us the bare minimum scope down that Vantage needs to do to do its job, which for us, is cost and usage report, cost explorer, and then

what looks like a long list of permissions for AWS service API's, but essentially looking at list and describe actions for the majority of AWS services. Does that kind of makes sense?

[00:12:15] JM: Yeah, definitely. I'd like to talk a little bit here about differentiation. I think we talked about this a little bit in the last episode. But there are so many of these cost management platforms, is there a differentiator in your approach in terms of how you gather information versus other cost management platforms?

[00:12:39] BS: Yeah, good question. So, the thing that's been surprising about the space is, we, as far as I know, are the only self-service platform out there. I don't know why this is the case, and I'm sure me mentioning this, if there's competitors that are listening in on this interview or things along those lines. But every other provider seems to be top down, sales driven, and Vantage is taking the exact opposite approach, which is self-service, bottoms up. Anyone who's listening to this can go and create a Vantage account in five minutes. They can connect their AWS account and go from there. And then from a feature perspective, we actually, over the span of the last year have majorly closed the gap in terms of us providing features of what other incumbents have provided. We're actually excelling in a number of ways that we're now offering features that the majority or in fact, all of other competitors in our space, don't have.

So, one very clear example that actually just launched a few weeks ago, happens to be one of my favorite products that we've launched, if not the favorite, we call Savings Planner. Savings Planner is a tool that allows for forecasting, and what if scenario planning. You can model arbitrary future compute spend, you can invite your finance team or your finance counterparts, or whomever in your organization to model out what specific commitments could look like. We're the only provider in the market that has something like this and it's been a huge hit with all of our enterprise customers.

That's one of the clear examples. As far as I know, we're the only provider that also shows granularity down on a per resource basis for I think, somewhere between 40 and 50 AWS services at this point. So, in some of the things that we were mentioning, throughout this interview, on cost of bugging, a lot of our users get major value out of drilling down into specific resources, see why certain costs are actually increasing on a per resource basis. Some of these products that have been around in market for 10, 11, 12 years in some cases, specifically like a

cloud health or cloud ability. We just have a number of their customers transitioning over these days for either lack of sophistication, lack of modernization. In some cases, the pricing model doesn't make sense, and so on and so forth. But yes, it's a crowded space. I think Vantage is doing a great job on just product execution, and then just our approach on a go to market basis is also very different.

[00:15:12] JM: It also seems like there's differentiation, when it comes to what you can layer on top of the basic cost information aggregation. You can do forecasting and just take more advantage of the data that comes off of these different services that are getting monitored. Can you talk about some of the features that you've built on top of just gathering basic costs or expenses? Obviously, you can easily gather the various servers and storage systems and caches and load balancers. You can gather that cost data, but being able to project where that information is headed and the cost trends, if you could do that, you can obviously help the user prioritize where to cut costs. So, can you talk about some of the features that you can build on top of that basic cost discovery layer?

[00:16:15] BS: Yeah, for sure. I can even expand on some of the things that I mentioned before, which is, the way that we think about Vantage is that we want to be the full representation of your P&L as it relates to cloud costs. So today, we have support for AWS, in another couple of weeks, we'll have support for Google Cloud, and Azure makes sense as the likely next candidate. However, there's a long list of other services that makes sense here that are typically line items, two, three, and four. These are things like Datadog, and CloudFlare, and PlanetScale, and Sumo Logic, the list goes on.

Now, what's interesting about all these providers is they're largely wrapping whatever the base primitives or recommendations are on top of AWS. AWS is always going to tell you the best way to optimize within the confines of your AWS account. What Vantage aims to do is by integrating all of these providers, we can give recommendations cross provider, and we've actually done a couple of these already this year, and they've been super popular in terms of maybe the best thing for your organization is not to use CloudFront, it's for you to use Cloudflare. Or in specific examples, maybe it's not the right thing for you to use RDS, you should be using PlanetScale, where pricing can make sense there.

So, to the original intent of your question, there's a number of things that we provide, in terms of optimization on where things can go with AWS, and we offer this, I guess, to your original point, kind of like enrichment features, or business intelligence features and savings planner, the modeling, and what if scenario planning tools, one of those where it's finance schemes, we're doing all of this in conjunction with engineering, in spreadsheets, or over Slack or things along those lines. Now, they have an interactive kind of space on Vantage to come in with their actual data, simplify everything down, save these models and share them. There's a number of forecasting features that we provide.

The things that are going to be super interesting, as people listen to this interview now and kind of throughout this year in 2022, they're going to see a number of features that are going to be more cross provider recommendations, or more collaboration features on top of some of the optimization features that we have on top of AWS. But I think increasingly, people are looking at what are the best options beyond just the core AWS service has provided, because we're just seeing a maturation happen from all these products in market from other service providers, other infrastructure providers, and so on and so forth.

[00:18:53] JM: You mentioned like a difference in cost between, for example, using an AWS native database, and using something like PlanetScale, which is a newer database system. We've done several shows about that. How dramatic can the cost savings be between those AWS native services and specific providers?

[00:19:19] BS: Yeah, so the answer is, it depends, which I know is always kind of the answer when it comes to anything infrastructure cost related. It depends. So, we actually just published a blog post that I say if you're curious about this, we give a full breakdown of here's what RDS gives you, here's how they think about pricing, here's PlanetScale, here's what they give you, here's how they think about pricing. The biggest distinction between these two providers is RDS, you're paying for something all the time that's sitting there, which for a database totally makes sense, right? And PlanetScale has a more usage-based approach that kind of going back to what I was saying on the "it depends" answer.

If you have a largely unutilized RDS instance, they can make a ton of sense to use PlanetScale, you can probably reap significant savings and we give an entire breakdown of how you can

think about this in that blog post and are actually working on kind of formalizing or automating, looking at what your RDS instance usage is on AWS, and potentially providing a cross provider recommendation in the Vantage console of when it can make sense to use PlanetScale. The flipside is that if you're really hammering your database, RDS can actually make a lot more sense in certain cases than a PlanetScale.

So, the answer is, sadly, it's complicated, and you could have significant savings one way or another, depending on your usage patterns. I don't know if I want to give like a specific quote on the amount of savings one way or another. But just because I don't have the information on top my head, but I wouldn't be surprised if one way or another, you could get 30%, 40%, 50% savings versus RDS which is like, significant, right? Again, it just depends upon your organization, what the usage patterns are, and things along those lines.

[00:21:08] JM: Coming back to just the basics on AWS. Let's say I install Vantage, and I get this report back of my expenses across storage and compute, and all the other services I'm using. Is it a straightforward practice to figure out like which of these services I should be optimizing? Is it as simple as just looking where the most dramatic costs are and discovering, and then just by looking up how I can save in these individual services? Is there some procedure, obviously, you can get a report back, but what's the action item on getting that report back? Is it only reporting or is there also prescriptive line by line, ways of how you can save money on each of these different services?

[00:22:06] BS: Yeah, so when a customer signs up, and they connect their account with Vantage, the first thing that they get is an interactive dashboard of essentially their top costs. Usually, when someone comes on board, it's not a surprise of like, "Hey, I'm spending the most on EC2 or RDS or things along those lines." If there is a surprise, that's usually the first thing that they'll kind of diagnose. Sometimes data transfer can be something kind of sneaky, that people are unaware of what they're using.

To your question, okay, I understand my costs, I understand what your services are coming from, what do I do next? And there's kind of two things that we offer from the Vantage side. One is, we do have automated cost recommendations that give you specific, it'll say along the lines of, "Hey, for EC2, we think you can save this dollar value with this, which equates to this

percentage of your overall bill for the last month, and give an automated description of what we recommend for you to do.” And where there are specific resources that match that cost recommendation, we’ll also link you to those specific resources in the Vantage console.

So, let's take an example. An example is S3 might be one of your top cost centers, you can go to cost recommendations, after connecting your account on Vantage and we'll say, “Hey, for S3, we think we can save you 20% when you shift from standard storage to infrequent access. Here is the specific buckets or objects that we think that this is a candidate for and someone on your team can go in there and evaluate this information from Vantage and choose to make an architectural decision going forward.” And so really, what we provide is visibility in automation for understanding your costs and what you can take action on. The cost recommendations are meant to give you a head start on those things.

The other thing that's actually just, it's public, it's hosted at handbook.advantage.sh. We call it the cloud cost handbook. It's entirely open source to so contributions that come from the community. But what we do is for AWS, and we want to expand this to Google Cloud and other providers is give us super, super informative, easy to understand kind of editorial viewpoint on the things that you can do per AWS service to save money. So, it'll essentially go through each individual AWS service, and give you the top three things that you can think about from a cost perspective. And we actually get like a fair amount of feedback from the community that this kind of stuff is helpful, or they didn't know about certain concepts, or they didn't know that they had certain options to them available on certain AWS services. So, our stance is visibility gets you 80% of the way there. There's another 10% that's just in the form of these cost recommendations. And then there's probably another 10% just on like education, understanding, making people kind of aware of the options that they have available to them. That's a verbose answer to all the things that we kind of give customers from our end.

[00:25:09] JM: You mentioned you have implemented the GCP integration recently. Have you noticed any important differences between the cost structures of a typical GCP deployment versus an AWS deployment?

[00:25:26] BS: Yeah, so we are just to set expectations here. So, we are in early access with Google Cloud right now. And by the time this is published, hopefully will either be at the end of

that, or in general availability. There are options for people who are entirely in on Google Cloud to use Vantage. Or if you have costs across both AWS and Google Cloud, you can kind of see that in a single pane of glass in the Vantage console.

In terms of differences between Google Cloud and AWS, I don't think we're far enough along on our Google Cloud journey to give broader thematic things of what we're seeing in terms of like, this is what compute costs look like on AWS versus Google Cloud or storage versus Google Cloud. Throughout this year, hopefully, we'll get that information and be able to share some of it. The things that we're seeing are just the differences in terms of how these providers actually expose their billing data, and I would say that Google Cloud does a much more modern implementation of billing data than AWS, which is just really nice for us to be working with, on behalf of our customers, and essentially, everything's backed by BigQuery there. So, it's all a queryable, we can pull in this information a lot easier than actually processing the underlying raw CSV files from AWS. So hopefully, more to share on that in the future, and we'll definitely be something that we publicize, but we're still early enough with Google Cloud at this point that no major things to share on that front.

[00:26:55] JM: Who is the typical user of Vantage? Does it get exposure to the financial teams or the senior leadership? Or is it particularly in the purview of the engineering teams?

[00:27:10] BS: Yeah, so this is something that we've seen kind of grow and evolve over the span of the last year. It's really stage dependent on the organization. So, I'll speak through what we see at different stages. At a startup or individual level, it's typically the easy answer, which is it's usually the technical co-founder, or like lead DevOps engineer, lead engineer. So, this is typically like a co-founder and CTO or something along those lines, really small team, and they'll invite just the rest of their startup. So, whether it be their CEO wants to see costs on a regular basis and get an email about them, or just have their entire startup engineering team in the Vantage console, kind of clicking around and taking a look at what's going on. That's what we see kind of at the lower end.

In terms of, I would call it like, it's usually referred to as like mid-market, I guess, more objectively, but for us, we see it as like SMB growth stage startup. It's usually like head of infrastructure, or lead DevOps engineer, who will sign up for Vantage. But then depending upon

the sophistication or maturity levels of the organization, in terms of like what employees they have, they'll sometimes invite like, well, first of all, usually all of their engineers, or at least the engineers that have an impact on anything cost related. And then on the management side, or finance side, if they have like a VP of finance or finance lead, they typically will invite them specifically for commitment, collaboration, modeling.

So, “Hey, we want to buy this amount of savings plans in this month”, they'll show it in our savings Planner Tool, send that model to their finance team who approves the budget, and kind of has all the information there for them to get going. We've had what I would call a very large poll from enterprise users in the span of the last year. So, these are publicly traded companies, and what we typically see there is usually an embedded cloud costs team, or FinOps team, where there's usually a person with a job title of head of cloud costs or lead FinOps analyst or things along those lines, where they're kind of on the usually on the platform team that's taking a look at costs, kind of on behalf of the broader organization. And there might be one or two finance counterparts, but usually it's on the engineering side. I would say on the enterprise side, it varies dramatically in terms of sometimes it's just that person who's in the Vantage console, other times they want to invite 250 engineers to assign to specific cost reports and get alerts and things along those lines. So, it really varies, I would say like just dependent upon how the organization culturally is thinking about costs. That's typically what see across the board.

[00:30:03] JM: For complicated deployments with, let's say, some number of EC2 instances, Kubernetes clusters, maybe a database that's running on, I don't know, maybe they decide to manage their own database on top of another EC2 cluster. What's the practice of using cost management information to right size the instances that this various infrastructure is running in?

[00:30:30] BS: Yeah, interesting question that pops up a lot on the Kubernetes side of the house. So, there are general right sizing things that you can do just on pure EC2 or RDS and Vantage pulls in CloudWatch metrics alongside your cost information. So, you can take a look at, “Hey, are we actually highly leveraging this vCPU or memory or disk IO or whatever with the specific EC2 instance?” There are some specific recommendations that Vantage will make into, I guess they would be considered right sizing recommendations. So, like, “Hey, you're not leveraging the vCPU or memory on this machine, you can change it to this other instance, type and save X percent.” And that's kind of one swath of work that we hear about occasionally.

Although typically, these days, we feel like organizations usually doing benchmarking in an upfront manner where they have a pretty good handle on the EC2 instances that they are using from a performance perspective. And so, then it becomes, “Hey, we have a Kubernetes cluster, is our cluster too big?” It’s really the distillation of the customer question. Or really, is it being utilized in a way that we’re maximizing our spend? So, Vantage, I guess, since the last time we talked, we do have full-fledged Kubernetes support. And in this, we give a breakdown of your cost by Kubernetes service, namespace, label things along those lines. But the thing that people get a lot of value out of is also just seeing the unallocated amount of your cluster for which there’s new Kubernetes pod running. And as a result, you may have some waste, depending on the size of your cluster.

So, you do get this out of the box in Vantage if you make our Kubernetes integration, which is pretty light, just over time, you can monitor on allocated space within your Kubernetes cluster as a percentage of your cluster and make right sizing decisions from there from a cost perspective. I’m trying to think if there’s anything else beyond Kubernetes, that may make sense here from a cost perspective. Nothing’s coming to mind. But I would say those are the three big swaths of cost savings potential would be EC2, right sizing, RDS right sizing, and then Kubernetes cluster right sizing.

[00:32:50] JM: Are you able to gather any interesting macro statistics from all of the deployments that are using Vantage given that you’re seeing so much information?

[00:33:03] BS: Definitely. So, we haven’t really shared scale of the amount of data that we see. But I’m comfortable saying that we are tracking now in the hundreds of millions of dollars of annualized AWS costs. I expect that that will hit a billion this year, over a billion. I think is like, actually, I feel pretty comfortable with saying that just given the growth. I mean, when you have line item data for that scale of costs, you definitely see some interesting stuff. We’ve exposed a little bit and in general, I try and err on the side of not exposing data just because I think it’s like something that should be held internally, although we do aggregate and anonymize and expose certain things.

So, one thing that we've done is, we host and maintain `leaderboard.vantage.sh`, which gives a ranked list of all AWS services as the amount of costs that they incur for customers. And it updates every single day. It's pretty interesting because the top 10 things are probably not a surprise. It's like okay, EC2 or data transfer, EDS or things along those lines. But what is interesting is when AWS launches a new service, and you can see how quickly or not quickly it's growing from a cost perspective. So, that's one piece of information that externally people can go and take a look at today. It's just kind of generally interesting information. We do have a lot of stuff that we've looked at internally in terms of like, how much is Fargate growing versus ECS or EKS? Or how much are the specific EC2 instances being adopted? Or Graviton, what is Graviton adoption by customers versus other chipsets? And so, we do have a lot of extremely interesting information we would love to expose this data. It's really a team bandwidth thing, which I guess the pitch here is like we do have a number of open jobs. We're hiring. We'd love for someone in the data science kind of fashion to come in and help us with this stuff. But yeah, a ton of super interesting data. The thing that's interesting is like the bill tells you everything about what people are doing on AWS at the end of the day. So, a lot of just like, really, really interesting data.

[00:35:16] JM: Have you thought about what other features you're likely to build after expanding to the other cloud providers? What else could you build on top of this platform?

[00:35:28] BS: Yeah, good question. So, in terms of the near-term roadmap, the things that we're looking at are what I would call more like, table stakes type features. Essentially, replicate what we've done on AWS, and apply it to Google Cloud and Azure. And if you haven't used Vantage, or even if you have used Vantage, one thing that we take super seriously is just ease of use, simplicity, and like an intuitive design. So, there's a challenge on just making sure that across these providers, things just work in the way that you expect them to work in a Vantage kind of friendly way. There's a lot of work, both engineering and design wise, in terms of being a cusp, like we're talking about some enterprises that are like 80/20 split across AWS and Google Cloud, and just making sure the reporting functionality across those clouds is done in an accurate and seamless manner, is a good amount of work.

In terms of other things that we have, besides what I would call table stakes, enterprise features, which are things like role-based access control, SAML integration, things along those

lines. I think there's a lot of work that we can provide in terms of visibility for organizations to make decisions about their infrastructure from a cost perspective. So, let me give like a couple examples. One is today, people are not going to lift and shift from like AWS to GCP. They're not going to lift and shift from AWS to Azure. But I think the thing that is always kind of a question in customer's minds is what is the premium that I'm paying by being on this provider, that it may make sense to switch to another provider?

So, in terms of what I've referenced earlier in the conversation around, we want to be the full P&L representation of your cloud costs, I think a whole other component of that could be what if analysis for moving certain workloads or entire workloads from one cloud to another, or from one provider to another, and our entire stance from the Vantage side is let's just give you the information you need for your organization to make some of these decisions. And it's a decision also might be, by the way, I know that this is not something that I personally, like opine on a bunch, but at a certain stage, it may make sense for you to move certain workloads off of the cloud and move to an on-prem type of situation. Again, at the end of the day, like what we want to do from the vantage side is just give you the information to be able to make these decisions. But I think that's kind of the realm of where we're going to be going into a lot as we onboard these other providers into 2022 and beyond, and hopefully more specifics to share there as the year goes on.

[00:38:05] JM: How do you validate and test the accuracy of projections, and I guess, just data that you're gathering?

[00:38:16] BS: Yeah, good question. So, we produce forecasts that are typically done based off of the usage that's in your account per service. And then we take a look at all these various services and roll them up into a general forecast. The way forecasts are done are also done per – we have a primitive called a cost report. And a cost report is basically a set of conditions over a subset of your infrastructure. So, for example, you could create a cost report for a specific tag on your infrastructure and get a forecast from that. We do a couple different things in terms of what we've historically done for gut checking these forecasts and improving them going forward. And then what we're going to be doing just internal advantage.

In the early days, what we did, which was kind of cheating, but it was helpful is we would benchmark ourselves versus what AWS cost explorer actually was forecasting. And that got us actually pretty far along. But then the other thing, too, that we do is we record, essentially a forecast of where we think you're going to be at intra month, and then we do a look back and learning on top of that forecast, as time goes on.

So, for example, if we think you're going to be, let's say you have \$3,000 a month in costs, and we think you're going to be at \$10,000. We'll record that forecast, and then at the end of the month, see what happened for the model to essentially learn going forward, based off of everything that's happening. Our forecasting models are also trained in an anonymized fashion across thousands of different organizations at this point, so they're getting more accurate as the user base just grows on Vantage. So, that's historically what we've done, and there's going to be just a lot more of this going forward because I think one surprising thing to me is actually, there's a lot of value that customers are getting from forecasting these days in a way that I actually didn't expect, and I continue to do a large number of our enterprise kind of prospective customer calls, and I've always been surprised by how much utility people are getting from service and resource specific forecasts, which to be totally transparent was not something that we did super intentionally from the get go, was just kind of an afterthought like, "Oh, yeah, of course, you need to have forecasting." But we've really invested in it a lot, and I think it's going to be an increased area of investment going forward, just given how seemingly important is for these organizations.

[00:40:40] JM: You spent some time at AWS prior to starting Vantage. I'd love to know if you've gotten any conversations or feedback from people at AWS or any insights that have related to Vantage?

[00:40:55] BS: Not really, as much as you might expect. I think the thing that is always, I think the outside perspective on like advantage versus an AWS, it's like, are you at odds with one another? I actually, I think I may have said this in the last time we spoke, and I really do emphatically believe this to be true. I personally have a great relationship with AWS. I think the organization has like Vantage and AWS are friends, we are AWS partners, at the end of the day, as well. I am in what I would call on official contact with probably I don't even know at this point, like 15 to 20 different service teams, just us checking in with their product managers or

engineering to make sure that one we have pricing correct, or two, there's things that they want to have included in the products, because there's things that we can help AWS with as well, in terms of like, it's probably better for them to have people running on Graviton, and it makes sense from a price perspective. So, we're happy to promote those types of things.

I mean, I would say like, AWS gets a bad rap. But at the end of the day, they're trying to do the best thing by their customers as well. I think in some regards, it's good if people are optimized on AWS from a cost perspective, because typically, what we see is once you're optimize, you actually end up spending more on public cloud, because you have that faith that you're not wasting money. So, no, as far as I know, we haven't had a single bad interaction with AWS. I would say they're actually helpful in more ways than people would expect. We also have begun having conversations with Google. I would say our relationship with Google is as good as with AWS at this stage. And like we see ourselves as the friends, these providers at the end of the day. So, yeah, nothing bad to share, only good things for the most part. And I would say like the unofficial helpers on the AWS side, in terms of like, people I can DM on Twitter or talk to you over Slack are just really, really helpful, good people.

[00:42:49] JM: Have you seen any interesting cost savings from people who put an emphasis on replacing their core infrastructure with lambda functions? Like moving away from even having long lived containers and just trying to move all their functionality into lambdas?

[00:43:06] BS: Yeah. So, what's interesting is I'm familiar with this model. There are, I think, two or three organizations I've spoken to fairly recently, and out of consideration for them, I won't mention their names, but I'm sure a number of people know who they are, that they're entirely like serverless organizations built on lambda. I think the thing that I've found is, unless everyone in your organization is fully bought into this, it's just a really hard thing to do in practice. And I think like, objectively speaking, it can make sense for a lot of organizations, but it's just hard to get right from the beginning and instill this as your engineering practice throughout the entire organization, is like the truth of it.

I would say from my perspective, it's usually the exception, not the rule and to a large degree, where it's very rare, actually, for me to see an organization all in on like a lambda. But at the end of the day, in the event that an organization is all in on lambda, their cloud costs are like,

dramatically smaller than I would expect. So, that's one thing for people to maybe hear. But I personally don't have a viewpoint. We at Vantage run entirely up on top of Fargate. We pay a premium for that, but we're happy with it. There's likely cost related things we could do there. But it doesn't make sense to rip out our entire workflow for something like that as an example.

[00:44:26] JM: What's the hardest engineering problem you've had to solve to date building Vantage?

[00:44:31] BS: Yeah, so very clear, these days, the amount of data that we are storing, indexing, structuring, and serving up to the user is starting to be fairly dramatic. And the thing that I mentioned around when you have line item data for hundreds of millions of dollars annualized and soon to be billions of dollars, it's just a lot of data. So, our engineering team is really – I mean, we've been spending, I would say the majority of the engineering effort is around just making sure this entire process is correct and good and stable. Because our users who are signing up for our customers, who are signing up for Vantage at this point are, we're starting to talk to like some fairly large organizations out there. And with that comes orders of magnitude larger amounts of data, and so this problem is just continuing.

But I would say, this is the engineering challenge that I expect to just be the case going forward, and our ability to show costs on even a per resource basis for an organization that may have tens or hundreds of millions of dollars in cloud costs annually, it's a hard challenge to make sure that we have that all together for them and served up in a way that's quick and easy for them to understand and everything along those lines. So, that is the clearest engineering challenge for us right now. Actually, nothing else even comes remotely close to that, I would say.

[00:45:56] JM: Cool. We're nearing the end of the time. Anything else to add about Vantage or what you've been working on engineering wise?

[00:46:04] BS: I think you've covered a lot of it. The only thing I was going to say is like, we're hiring across the board, and the team is growing fairly substantially. So, I'm happy to throw a pitch in for that. But if that's like more unappealing, then happy to admit it as well.

[00:46:19] JM: No problem. Cool. Well, Ben, thank you so much for coming on the show. It's been a real pleasure, as always.

[00:46:25] BS: Yeah. Thank you so much. It's been awesome to help back on the show basically a year after we originally did it. And I hope, in a year from now we can do it again and we have as much, if not, more progress to share.

[END]