

Take Control of Cloud Costs

A guide to setting up an effective
cloud cost optimization program



Contents

Introduction	04
Core Principles	10
Key Components	16
The Path Forward	24





Introduction



Gartner

Worldwide end-user spending on public cloud services is forecast to grow over 40% in the next two years to \$362B.

There is little doubt that the cloud has cemented its place as the present and future of technology infrastructure. Well worn is the topic of its many advantages - speed, flexibility, scale, cost-effectiveness, and a never-ending stream of new capabilities - that are helping organizations leverage technology more effectively than ever before. However, this tremendous new capability has also brought about a new challenge - reigning in cloud costs that can quickly get out of control, eroding one of the core value propositions of the cloud.

[According to Gartner](#), worldwide end-user spending on public cloud services is forecast to grow over 40% in the next two years to \$362B annually. At that pace, public cloud spend will double roughly every four years. Looking at broader technology trends, the amount of money that organizations spend on technology is staggering. [According to IDC](#), spending on information and communication technology (ITC) globally is forecast to be over \$5T in 2021, growing to almost \$6T by 2023.

If you believe as we do that the public cloud is the future of technology infrastructure, then that is a massive potential amount of spending in the public cloud. Unfortunately, the same characteristics that make the public cloud a powerful tool also make it a powerful potential financial obstacle. The ease at which teams can spin up new environments to innovate and respond to market changes also means that they can rapidly increase cost just as easily. The traditional barriers to bringing on more infrastructure - cost, buying servers, wiring up infrastructure, purchasing and installing hardware, etc. - can't be relied on to control or slow down spend. [Gartner's](#) research has found that organizations that are not focused on cloud cost optimization may be overspending by 70% or more.

The fact that teams are given the flexibility to make their own infrastructure decisions quickly based on their specific needs, rather than be subject to centralized infrastructure planning, means that they can get exactly what they need, exactly when they need it. This flexibility also means that they can easily over-provision or lose track of spend.

Centrally-negotiated discounting, with its long cycles and limiting standardizations, has given way to direct purchasing of cloud infrastructure by teams and business units, lessening the centralized leverage of

organization-wide purchasing and negotiations.

Finally, the sheer complexity of cloud billing, with constantly changing services, prices, and capabilities, combined with hundreds of thousands of SKUs, means that it is harder than ever to make sense of charges. Billing for cloud services is finely-grained and highly complex — more akin to old school phone bills than the static and course-grained (yearly, quarterly, monthly) billing that we've historically experienced with traditional infrastructure.

So what should you do about this? Of course, this doesn't mean that organizations should stick with traditional, expensive and rigid infrastructure, but it does introduce the need for new ways of monitoring, measuring, and controlling costs. Organizations can use the public cloud as a powerful tool to innovate and evolve rapidly, but they can easily find themselves paying dearly for that capability. This guide will provide you with a framework for managing and controlling your cloud spend while empowering the teams driving innovation and productivity within your organization to make the most of all that the public cloud has to offer.

A New Way of Thinking

In our experience, one of the most common mistakes organizations make when embarking on a cloud cost optimization or cloud financial management (CFM) program is to treat it just like previous programs to manage physical infrastructure. In those efforts, the premium is on predicting costs and infrastructure needs with precision, often months or years out into the future. When you need to buy physical hardware, set it up, purchase software, and procure hosting (either in your own data center or with a 3rd party), this makes perfect sense. The cost for over-provisioning is incredibly high, and the risk of under-provisioning is possibly even higher depending on the use case.

The beauty of the cloud is that the cost of either scenario can be very low if it is mitigated properly. You can underestimate or overestimate with little risk - after all, one of the key benefits of the cloud is that you can adjust capacity on demand. The real risk with the cloud is long after the initial planning, in the form of ongoing waste and inefficiency. Optimizing a program for efficiency rather than planning accuracy requires a different thought pattern.



Core Principles



Much attention has been paid to how to effectively build and manage technology at scale in the cloud, and there is an emerging discipline, known as Cloud FinOps, that looks to provide organizations with ways to control and manage cloud spend, typically from an operational or financial perspective.

The book [Cloud FinOps, by JR Stormont and Mike Fuller](#), is an excellent source of information on the intricacies of cloud billing and offers a wealth of insights on how organizations are working to manage and control their cloud spend.

The framework described here offers a way for organizations to look at the value delivered by the public cloud holistically. To be truly effective, a cloud cost management program must consider the needs of a wide range of stakeholders, including engineering, product, operations, finance and lines of business. The program must be centered on maximizing all of the things that are great about the public cloud while giving everyone the tools to ensure that spend doesn't get out of control.

To understand the thinking behind this framework, it is helpful to start with some core principles:

01

Know what is important, measure it, and democratize access to information.

Every organization is different, and therefore what defines value in the cloud is different. If you are a cash-rich incumbent trying to keep pace with more nimble competitors, then speed is most important. If you are operating at scale and provide critical services, then perhaps reliability is most important. If you are in a highly competitive, cost-sensitive industry, perhaps cost control is most important. Regardless of your situation, value is generally a combination of factors and is defined in your context. The first step in creating a cloud optimization program is to understand and clearly define your goals and ensure widespread visibility to those goals and your organization's progress.

02

Maximize Value Rather Than Reduce Cost.

This sounds obvious, but most organizations don't measure value, whether it is when managing cloud spend, software spend, or services spend. Most organizations look at costs, and some look at defined outcomes, but few look at value, defined as outcome/cost. Value is focused on maximizing the units of outcome for every dollar spent. You can reduce your cloud spend by 10%, but if productivity declines by 20%, you are losing money. Likewise, increasing cloud spend by 50% isn't bad if you increase productivity by 100% or generate new revenue streams.

03

Enablement, not control.

The cloud provides a wealth of benefits, including the ability to more precisely allocate technology infrastructure to what is essential, thereby lowering operating costs. But the real point of moving to the cloud is unlocking the potential of organizations and their engineering teams to optimize their use of technology, and cost is just one element of that. This framework is built on the premise that it is more effective to enable teams to maximize value in the cloud than control their spending through restrictive policies, standards, and procedures. The more teams know about how they can maximize value, and the easier you make it to allow them to maximize value, the more progress you will make.

04

Optimization should be engineering and product-centric.

An optimization program should be interwoven into, and aligned with, an organization's goals and objectives from an engineering and product perspective. Priorities should be determined by and supportive of engineering and product objectives. Teams should embed optimization into core processes like product design, DevOps, and support rather than addressing it as an afterthought or separate consideration.

05

Leadership support isn't helpful. It is essential.

There are lots of great products and services out there to help organizations optimize their cloud, but in our experience, there is one thing that is an absolute determining factor in the success of optimization programs, and that is leadership support. No matter how much visibility you provide or how easy you make it, teams always have competing priorities, and unless leadership is clear that optimization is important, it won't happen. It is that simple.

Key Components



GOALS, KPIs & STANDARDS

Establish the goals, set the guardrails, and define the program measures.



VISIBILITY

Go beyond understanding changes in spend to understand what exactly is changing and where there is actual waste.



EXECUTION

Money isn't saved by identifying inefficiencies or waste, it is saved by taking action. Cloud environments at scale are complex so this can't be underestimated.



MEASUREMENT

Understanding who needs to take action, and quantifying the value of that action, closes the loop and drives improvement.

Goals, KPIs, and Standards

As with any initiative, it is important to start with what you are trying to accomplish. This includes the goals of your optimization program (we are trying to save X, we want to get efficiency to Y), the KPIs that measure its success (not all financial), and the standards that will drive the program.

Goals

Most readers would suggest that this is obvious - isn't the goal of a cloud cost optimization program to reduce cost? Well, not necessarily. Most organizations are doing more in the cloud over time, whether it is because they are growing, moving more to the cloud, or both. While top-line cost is important, and you should certainly measure that over time, it offers an incomplete picture. It is critical to measure how much value you are getting from your cloud. That could be measured in terms of revenue or customer growth, reduction in support costs, or the efficiency of your use of resources. Regardless, you should settle on a measure of value, and anchor on that.

KPIs

What if you meet your cost reduction goal, but at tremendous cost to other aspects of your business? Perhaps your support costs go up, perhaps application stability degrades, or perhaps you have less productive development teams. Think of KPIs as the guardrails of a cloud optimization program.

Examples of KPIs are service level metrics (uptime, performance, etc.), cost to support, and build quality. Often engineering teams will raise valid concerns about the impact to the stability or performance of applications, and it is important to both define what can't be impacted by an optimization program, as well as have data to support a program's impact on performance.

Standards

Standards drive your cost optimization program. Key items to consider are:



Tagging

Cloud environments are complex, particularly at scale. Tagging helps organize our understanding of a cloud environment and its usage by providing the metadata to categorize and organize data and insights.



Data Definitions

As with any program, communication is critical. A foundation of communication is aligning on a set of terms, technologies and data standards that are mutually understood, consistently used, and widely agreed upon.



Definitions for Measurement of Performance

To build a successful program, you must have a clearly defined, objective, and quantifiable set of key measures, including a clear definition of success for the program.



Technical Requirements and KPIs

While optimizing cloud usage is incredibly important, we can never forget that it is not always the primary value driver for cloud computing. For example, if a system is mission-critical, erring on the side of reliability over cost savings will always make more sense. Conversely, reducing cost by over 50% in exchange for a small risk of correctable performance degradation in a dev environment will likely make sense.

Visibility

To solve a problem, you first need to understand it. This starts with general cloud spend patterns but goes more deeply into an understanding of:



Spend

What are we spending money on? Who is spending the money? Are we getting the appropriate value for that spend? What changes in spend are we seeing? Why?



Opportunities

What are the opportunities to save money or improve efficiency? Are we acting on those opportunities? If not, why?



Results

What is the return on investment of our program? Is our value per dollar spent in the cloud increasing or decreasing over time?

The important thing here is to make sure that you have both depth (what specifically am I spending money on) and breadth (visibility across an organization). Changes in spend patterns aren't necessarily indicative of a problem or progress - you can see 10% lower cloud spend on 20% lower value (bad), or you can see 30% more cloud spend on 100% more value (good).

This is an area where tagging and standardization are critically important. To drive visibility, you must know what each instance is being used for, and by whom. Organizations follow many different processes and structures, and there is no right or wrong answer, as long as the patterns are consistently applied, followed, and enforced.

The level of detail matters for making informed decisions. It isn't enough to know that we are spending too much on X. We need to know we are spending too much on this specific instance, which is a non-production resource that is owned by Z group and managed by person A. Only then can you drive an efficient, effective process.

Execution

Perhaps counterintuitively, this is typically the hardest part of the process, particularly in large organizations. Determining what part of cloud spend is waste and what is necessary to meet the needs of the business is complicated, difficult to ascertain, and open to interpretation. Clear goals, KPIs, and standards are important because they set guidelines and give everyone a common set of terms and frameworks from which to operate. In other words, everyone is working from the same set of facts.

That is just part of the picture, however. At a smaller scale, making these changes can be viewed as bothersome, risky, or counterproductive, particularly if the mechanism to communicate, review, approve, and execute these changes is manual or cumbersome. In our experience, the more integrated the process is into the standard software development lifecycle, the more likely it is to succeed. When the ask and expectation for execution are clear, it is far easier to execute within the procedures and boundaries of the normal software development process.

Managing cloud spend gets more complicated with scale. A particular challenge is what to do with all of the “smaller” opportunities. \$1,000 savings opportunities may sound small, but 1,000 of them is \$1,000,000. 10,000 of them is \$10,000,000! Many organizations have thousands, tens of thousands, or even hundreds of thousands of resources. It isn't hard to see how that could add up quickly — after all, dollars are made of pennies. But how do you do this? Automated processes that are embedded into your software development lifecycle make this work as scale.

Measurement

As mentioned above, understanding the exact details of each opportunity is critical, especially if your goal is to enable the specific teams and individuals to take specific action. It also allows organizations to hold each other accountable for maximizing the efficiency of their cloud spend.

Providing a way to categorize, organize and assign optimization opportunities is critical to the success of any optimization program. There are many approaches to driving widespread adoption and buy-in from teams to help manage and reduce cloud spend - it really depends on the culture of a given organization. We've found that if you give teams and individuals the right information, make it easy for them to act on that information, and leadership makes it clear that the initiative is important, you will meet or exceed your goals. Gamification concepts, friendly competition amongst teams, or highlighting and rewarding the teams that are the most proactive go a long way towards building a successful program that the entire organization can get behind.

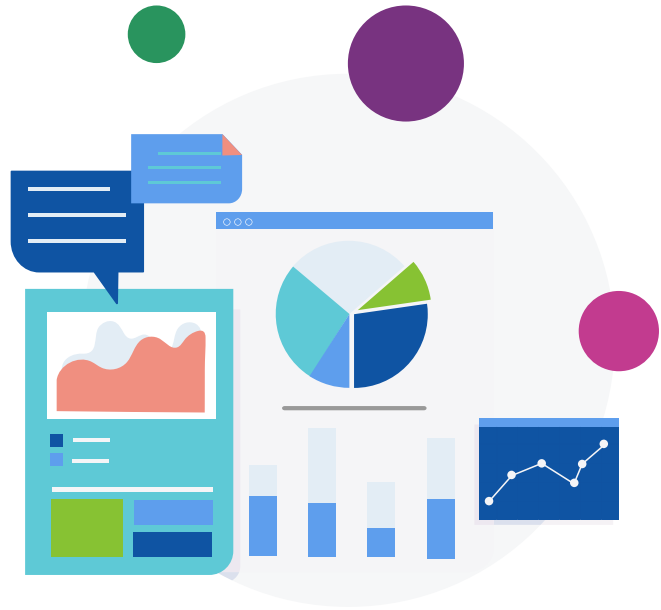
But how do you really know your program has been successful, and that the changes your program has driven are having an impact? You have to trace the full lifecycle, from increased spend to recommendations to execution, all the way to your bill.



This is important for two reasons:

Measuring the path from action to end result is really the only way to measure ROI.

Did the changes actually change your bill and the overall efficiency of your cloud in the way you expected? Are the teams driving the change seeing the value, and is it having a positive impact on their overall objectives?



Iterative change is the most effective, safest way to optimize your cloud.

For iteration to work successfully, you must iterate rapidly and constantly, with each iteration leveraging the feedback loop of data from previous iterations. Only then can you optimize with speed, certainty, and quality.

The Path Forward



We are still in the early stages of widespread adoption of the public cloud. There is no doubt that it already has been, and will continue to be, transformational for organizations of every size.

While the advantages of moving to the cloud and driving transformation are beyond doubt, they are not always guaranteed. For organizations to maximize those advantages, they need to think differently about managing cloud costs.

An iterative, data-driven approach to cloud optimization, driven by clear goals, measurable KPIs to ensure business value, and widely accepted standards, is key to a successful optimization program.

A modern cloud optimization program must be iterative, data-intensive, integrated, and automated in order to maximize value at scale while minimizing disruption. This doesn't need to be a massive effort, and you don't need a massive team to do it. You simply need a solid model, leadership support, and buy-in from teams across the organization.

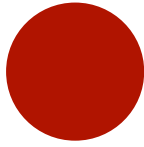
Our recommendation? Start small with a friendly team. Get some wins. Iterate. Learn. Measure. Once you have the formula and teams are seeing results, the good news will travel.



Thrive in the cloud.



Virtasant is a cloud technology company that specializes in helping organizations leverage the power of public cloud platforms. Our mission is to help organizations around the world thrive in the cloud by leveraging our proprietary methods, automation and technology, supported by a global team of 4,000 cloud experts in over 130 countries.



401 Congress Ave., #2560
Austin, TX 78701 USA



www.virtasant.com

