

9 STEPS TO CONTROLLING COST IN THE CLOUD

A practical look at the challenges
and critical steps to control costs and
optimize your cloud architecture



TABLE OF CONTENTS

9 STEPS TO CONTROLLING COST IN THE CLOUD	3
1. GET BETTER VISIBILITY INTO YOUR DATA	5
2. REVIEW UNDER-UTILIZED RESOURCES	6
3. CREATE A STRONG TAGGING STRATEGY	7
4. USE RESERVED OR SPOT INSTANCES (WHERE IT MAKES SENSE)	8
5. CREATE A SCALE-IN AND SCALE-OUT POLICY	9
6. CONTROL YOUR CLOUD PRACTICE WITH GOVERNANCE POLICIES	10
7. RE-EVALUATE YOUR DESIGN CHOICES	11
8. FIND THE RIGHT BALANCE	13
9. MAKE THE MOVE TO MANAGED SERVICES	14
CONCLUSION	15

NOT SEEING THE ROI YOU EXPECT ON YOUR CLOUD INVESTMENT?

HERE ARE 9 ACTIONABLE WAYS TO MAINTAIN CONTROL AND PREDICTABILITY OF COSTS, AND OPTIMIZE YOUR CLOUD INVESTMENT.

You've migrated to cloud to scale your services, increase efficiencies and unlock greater business profitability. But maybe you're not seeing the return on investment you expected. Or maybe your costs are spiralling out of control, and you're lacking visibility into where those costs are coming from.

It is a common belief that the cloud is more cost-efficient than a traditional infrastructure solution, since it's an operating cost rather than a capital expenditure. Sometimes, though, a cloud solution can end up being more expensive than anticipated for several reasons, from poor governance to misaligned system architectures and more.

While cloud transformation offers real financial benefits, those benefits are not guaranteed. In fact, the third report in Aptum's four-part Cloud Impact Study, *A Bright Forecast for Cloud*, found that 57% of organizations experienced higher than expected costs from cloud; and more than a third (35%) admit they're wasting IT dollars due to inefficient use of cloud platforms.

As organizations migrate workloads to various cloud platforms, visibility and control into cloud environments becomes difficult to achieve, often leading to unforeseen costs — a key challenge for 81% of respondents to the study. Without the insights necessary to evaluate performance and enforce remedial measures for issues that arise, the cloud environment becomes harder to manage and efficiency difficult to ensure.

“Without necessary insights to evaluate performance and enforce remedial measures for issues, the efficiency of cloud becomes hard to manage. This is the case for more than half of organizations (57%) who say cloud has resulted in unanticipated costs, and over a third (35%) who admit they are wasting IT spend due to the inefficient use of cloud platforms.”

– *A Bright Forecast for Cloud*, Aptum (2021)

“Working in the cloud is totally different from working in a traditional infrastructure environment. You’re moving away from the fixed costs of running an infrastructure, and into a dynamic model where usage and costs are changing on a daily basis,” says Kwong Lum, Lead Architect, Cloud, with Aptum. “Cloud delivers efficiencies and even cost savings, but it needs to be implemented and managed effectively to ensure you get those benefits.”

A recent **Gartner report** cited several challenges faced by organizations using cloud services. They include:

- Multifaceted pricing where the numerous options make it difficult to know which is the best pricing scenario for your use case
- “Extreme granularity” that makes it difficult to enable chargeback
- Resource sprawl that leads to unnecessary charges
- Constant change in cloud offerings which makes it difficult to predict how new services or features will impact costs
- Lack of standardization between cloud platforms, which means lack of standardization among billing formats

There are other challenges too. But whether you’re using Microsoft Azure, Amazon Web Services (AWS), Google Cloud Platform (GCP) or a multi-cloud environment, there are practical ways to control your costs and even save money, whether in business-as-usual processes or for more tactical purposes.

If you’re struggling to control and optimize cloud costs, or failing to attain the ROI you expect, this eBook is a must-read. We dispel many of the myths around costs in the cloud, while providing 9 practical steps to control and predict costs, and optimize your cloud architecture — from simple, low-cost actions to more complex, tactical solutions.



1

GET BETTER VISIBILITY INTO YOUR DATA

It sounds simple enough: Cloud creates cost efficiencies because you only pay for what you need. You choose the best deployment model for each workload to maximize cost savings. But in practice, this isn't as simple as it sounds, since cloud can limit visibility and control of resource usage.

You need cost optimization data and analytics — but you need them at different levels of your organization. The executive leadership team, for example, needs a different level of reporting than the engineering team. But how do you get the relevant data to those who need it? How can you build visibility of cloud costs by workload and by department, and identify any inefficiencies?

In a legacy environment, you typically buy a machine with all the performance you're ever going to need, and then some. The model for cloud, though, is to run on the minimum configuration that will give you the performance you need day-to-day so you're not overspending on resources.

Getting better visibility into your data requires an environmental assessment – What are you running that's permanent, and what is temporary? What is stable and what is in flux?

“With stable instances, you can often leverage additional cost-saving measures like buying reserved instances, or buying a software subscription to lower your costs. You could gain up to 60 per cent savings by doing that,” according to Kwong Lum. “But you're not gaining anything by purchasing a reserved instance if the amount of consumption is never going to pay for itself.

For example, if your virtual machine is turned on and used only 10 hours per month, it won't make sense to buy a reserved instance for it. Similarly, if you are evaluating an application where performance stability has not been established, it wouldn't make sense to commit to a 1- or 3-year term for a reservation that is under or over requirements.

It's critical to do a cost analysis assessment of your environment – what you're running, how you're running it and the length of time you're running it for.

KEYS TO GETTING BETTER VISIBILITY INTO YOUR DATA:



- Equip yourself to monitor and control your cost in real time
- Understand the nature of your workloads (stable, fluctuating, seasonal, etc.)

2

REVIEW UNDER-UTILIZED RESOURCES

Third-party cost analysis tools are readily available in the market, and cloud platforms come with built-in tools to help manage costs. The bigger issue is that many organizations take a 'set it and forget it' approach, where no one looks at the cost analysis tool until there's actually a problem — and by then it can be harder, and take longer, to fix.

A better strategy is to take a proactive approach to cost management, using your cost analysis and optimization tools to review under-utilized resources. This can help you see the overall performance of your environment and determine if you're overprovisioned — and if you are, make recommendations on the right provisioning so you don't pay for resources you don't need.

Cost optimization isn't a one-and-done activity. It should be done continuously, and automation can help to provide recommendations on usage. Native cloud tools like Azure Advisor can help monitor performance and flag underutilized or overutilized virtual machines. For example, it can help you resize virtual machines to lower the cost of managing multiple VMs. You can start/stop VMs during off hours or evaluate VMs based on CPU and network utilization, giving you the choice to decommission, resize or shut them down.

AWS also provides several reporting and cost-optimization tools, such as AWS Cost Explorer, which allows you to see spending patterns over time, observe your use of reserved instances and identify areas that need further inquiry. The AWS Cost and Usage Report details your hourly AWS usage across accounts for DIY analysis.

KEYS TO REVIEWING UNDER-UTILIZED RESOURCES:



- Be proactive at hunting unused resources
- Stop VMs when they're not in use

3

CREATE A STRONG TAGGING STRATEGY

Providing the right visibility into data — and ensuring each stakeholder group in the organization has the right data — can be accomplished with a strong tagging structure. Importantly, tagging helps you stay on track with your budget, identify overspending and even put caps on spending in certain departments.

A robust tagging strategy enables public cloud infrastructure components to be identified in various ways relevant to their use. For example, tagging can identify groups of resources that are owned by a department or project; as a temporary or permanent resource; or

by the ability to do chargebacks to departments. Tagging resources makes it easy to strategically deal with all aspects of the cloud in a clear and concise manner, and implement a level of governance that prevents over-deployment, controls spend, and assigns clear owners and responsibilities.

“The importance of tagging your resources and assigning them budget can’t be underestimated to ensure you have meaningful information in your cost analysis tools,” says Martin Poirier, Cloud Solutions Architect with Aptum. “Tagging your resources – including assigning multiple tags to the same resource – lets you set your budget, then track spending on a continuous basis. You can even set alerts so you’re notified if, for example, usage in a particular department hits a certain spend level.”

For example, when you create an Azure policy that dictates how your environment is configured, that policy can ensure all resources require a tag, so users can’t spin up a resource without it. And with AWS Budgets, you can set custom budgets that trigger alerts when cost or usage exceeds (or is about to exceed) a set dollar amount.

“In cloud computing, organizations are confronted with the difficulty of creating accurate cost estimates. They are often hit by bills that they apparently can’t explain and struggle to identify items that are responsible for spending. As a result, financial management is often overlooked until spend is out of control.”

– [How to Manage and Optimize Costs of Public Cloud IaaS and PaaS](#)

Gartner, (2020)

KEYS TO A STRONG TAGGING STRATEGY:



- Group resources with tags to track where the spending flows
- Set multiple budgets depending on target audiences

4

USE RESERVED OR SPOT INSTANCES (WHERE IT MAKES SENSE)

For certain use cases, reserved and spot instances can provide significant discounts from on-demand pricing. Cloud providers also offer payment plans or discount pricing, which are worthwhile for certain workloads. While these options can save money, they also come with caveats — and must be used with care since they won't make sense for all workloads.

RESERVED INSTANCES

Reserved instances enable you to prepay for capacity over a specific period of time, typically one or three years. The advantage? A significant discount compared to on-demand pricing. But that deep discount could end up costing you money if you don't really need the capacity. According to Kwong Lum, "Say, for example, you're getting a 40% discount through reserved instances. If there's 744 hours in a month but you only run that machine for 100 hours a month, it's not going to pay for itself."

Using reserved instances requires a clear strategy. Some organizations purchase reserve instances from the get-go and then, two weeks after their go-live date, realize their capacity is grossly oversized or undersized. But at that point, they're locked into a one- or three-year contract. It's best to purchase reserved instances based on historical knowledge of cloud usage rather than in pre-planning mode.

SPOT INSTANCES

Another area where you can save costs is spot instances. For example, when Microsoft or Amazon have extra capacity, they offer spot instances at a discount. Both vendors claim their customers can receive a discount up to 90% off on-demand prices through spot instances, with no term commitments. This could be a good approach for certain applications such as containerized workloads or test and development environments.

Like reserved instances, you can save money — but with a caveat. "As demand for that instance moves up, you could lose it without any notice, so it's not good for running critical workloads," says Kwong Lum. "But if you have a DevOps test project and you don't care if it disappears, spot instances are good for that. Or if you have large-scale data sets that are

load balanced, it's generally not a big deal if you lose one instance. But you do have to be careful."

DISCOUNT PRICING

Payment plans are available that can provide discounts compared to on-demand pricing — depending on your usage. It's a decision that requires an environmental assessment and clear understanding of what's best for your specific workloads.

Microsoft offers discount pricing through its [Microsoft Azure Enterprise Agreement, Enterprise Dev Test Subscription and Cloud Service Provider \(Partner\) Program](#). Amazon also has [flexible pricing models for AWS](#), including its Compute Savings Plans, EC2 Instance Savings Plans and Amazon SageMaker Savings Plans.

There are also licensing terms, such as the [Azure Hybrid Benefit](#), that let you use your on-premises Software Assurance-Enabled Windows Server and SQL Server licences on Azure. That means if you already have on-premises licenses with Microsoft, you can reuse those licenses in Azure. Martin Poirier suggests "You should look at that every time your contract expires with Microsoft. But it can become a challenge because you need to keep your cloud team in sync with your enterprise licensing agreement."

KEYS TO USING RESERVED OR SPOT INSTANCES:



- Consider the varying contractual mechanisms cloud providers offer to save
- Purchase Reserved instances based on a concrete KPI, not assumption

5

CREATE A SCALE-IN AND SCALE-OUT POLICY

Not all workloads are equal. Typically, costs increase as demand scales up. But you also want to scale down unused resources if demand drops. Some workloads are highly stable with a predictable performance requirement. Other workloads vary by time of day or by events.

When a workload is predictable, you can apply rules to scale your environment as required, so you don't have to pay for what you're not using. But if you have unpredictable loads — such as in e-commerce and retail with promotions and seasonal increases in demand — automation is required to monitor demand.

At its simplest, a scale-in and scale-out policy will help you apply automation to scale up when demand increases, and scale back down when demand drops. But building in automation to automatically scale up or down isn't necessarily easy, and it takes time to do it effectively.

KEYS TO A SCALE-IN AND SCALE-OUT POLICY:



- Automate scaling for unpredictable loads
- Build automation to scale based on demand



6

CONTROL YOUR CLOUD PRACTICE WITH GOVERNANCE POLICIES

Good governance can help you gain both cost savings and optimization in the cloud, but don't expect to see results overnight. "It takes time for the cost savings to appear. They require you to build a strong governance policy, and your internal processes must adapt to it — say, when your supply chain is better trained to understand the cloud. Then you start to see the return on investment," says Martin Poirier.

Cloud governance should be built into a project during technical engagement with relevant business departments or stakeholders – application owner, finance, procurement, suppliers and others. Tagging is the technical solution to this, but governance (via an overwrapping policy) sets parameters around what users can do in Azure or AWS.

For example, your governance policy could restrict users to a geographical region. "If you have data that can only stay in the U.K., then your governance policy should say that you can only spend resources in the Azure U.K. data centers, and you can't provision resources anywhere else in the world," says Kwong Lum. "For controlling costs, you can also specify virtual machine types to control what kind of virtual machine users can spin up."

KEYS TO CONTROLLING CLOUD WITH GOVERNANCE POLICIES:



- Governance policies can be technically enforced
- Governance must be a collaborative practice engaging many business departments and stakeholders



7

RE-EVALUATE YOUR DESIGN CHOICES

There's a lot of low-hanging fruit known to save costs on cloud. But if you're still not getting the desired ROI, re-evaluate some of your design choices. For example, have you mirrored a previous on-premises environment in the cloud? Do your cloud strategy decisions and design choices need to align with usage and decouple from legacy provisioning patterns? It might be time to modernize your applications, make them cloud-native or get away from IaaS deployments.

APPLICATION MODERNIZATION

Modernizing your applications could involve refactoring (moving applications to cloud infrastructure) or re-architecting your applications so they're better suited to a cloud environment. But these projects are complex, involving a complete rethink of your strategy for a specific workload — and it can be hard to estimate your cost savings.

"This is the part that's quite challenging to estimate because you're trying to scope the cost of rebuilding an application, and in platform-as-a-service you'll find that your cost analysis tools lack the features to give you good documentation," says Martin Poirier. "Optimizing infrastructure-as-code is not the same as optimizing in the cloud-native application."

With legacy applications, it's not possible to modernize them overnight, and if they're sitting on legacy hardware you'll need to take a staged approach. There's no one-size-fits-all approach, so each application must be considered on a case-by-case basis. In some cases, it may need to be retired or replaced by a software-as-a-service application.

"It's about continuous improvement, continuous assessment and re-evaluating against not only the current environment, but potential alternate environments. That's a benefit of going to a vendor-agnostic service provider, because maybe most of your workloads work well in the Azure environment, but a portion of an application or a portion of your infrastructure would be better suited to AWS. Taking a vendor-agnostic approach can help with that."

– Grant Duxbury, Director, Pre-Sales Engineering, Aptum

DATA STORAGE

The right mix of storage contributes to controlling costs. Make sure you have lifecycle management and the right tiering — hot, cool or archive. Otherwise, your cloud storage could cost you more than anticipated in the long-run. Whether it's Azure, AWS or another cloud platform, it's easy to increase spend on cloud storage when you need it, but once you're locked in, it's hard to reduce that spend. So getting it right from the get-go is important.

"It's so important to understand if you're choosing the right storage for now and for the long-term," says Martin Poirier. "If you have a lot of archival storage but decide you don't want to archive with a particular cloud provider anymore, it can become expensive because the cloud provider charges a premium to read from the archive. So the colder the storage, the more expensive to go back and fetch from it. Be sure to make the right decision in the first place."

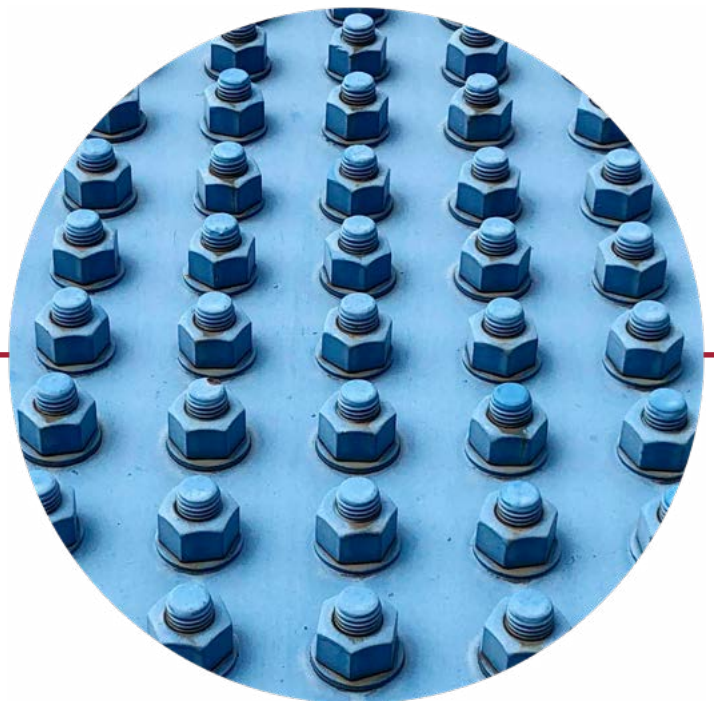
DATA TRANSFER

Since pricing is based on billing zones and regions, be sure your egress is limited in those billing zones and regions — unless your service levels require it for availability, or as part of your disaster recovery plan. In other words, you don't want to pull data from another region where you'll end up paying egress charges, which can quickly add up. That should be considered as part of your overall cloud adoption framework.

KEYS TO RE-EVALUATING YOUR DESIGN CHOICES:



- If there's no low-hanging fruit, review your architecture choice
- Re-architecting is a project in itself that needs to provide a ROI



8

FIND THE RIGHT BALANCE

When it comes to designing a workload, there will always be trade-offs. It's important to understand those trade-offs and find the right balance between cost and reliability, performance and security.

As Martin Poirier says, "Finding that balance is an art. When you design a solution, you're given a budget and you might have to make some trade-offs. You should never trade-off on backup or security — that's not a good idea. But you can do some smart trade-offs such as reducing your mean time to recover, or in terms of the SLA, which allows you to reduce the complexity of the design and therefore cut costs. It's always about finding a balance, but sometimes you need to spend if you want to achieve your objectives."

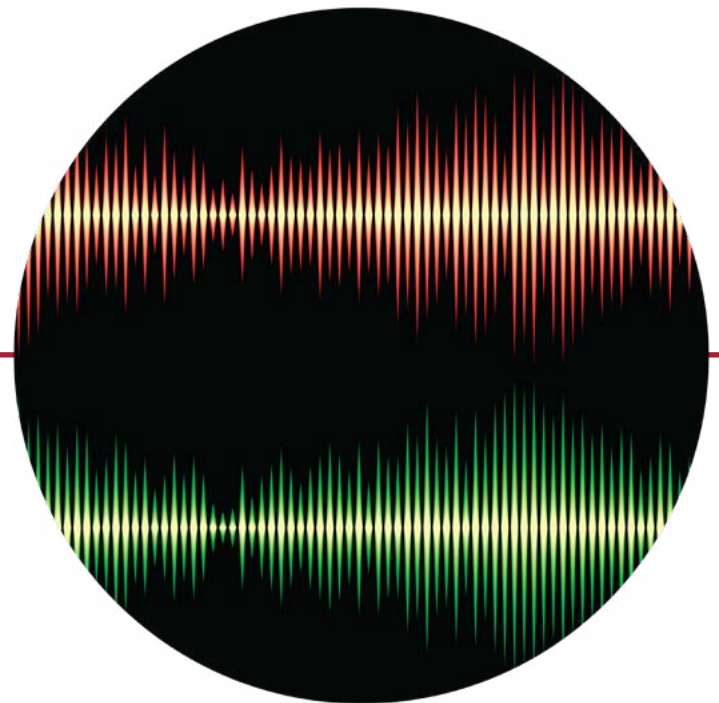
You should also consider which investments might increase your costs initially — such as spending on automation and systems monitoring — but will save money over time. Once again, having cost analysis tools in place, as well as a strong governance structure, will help you see which investments are likely to pay off down the road.

Similarly, spending money to boost reliability, performance and security may increase your overall costs, but could save in ways that are hard to measure. For example, it's hard to put a price on preventing the reputational damage caused by a major security breach.

KEYS TO FINDING THE RIGHT BALANCE:



- With a limited budget, trade-offs must be made
- Automation is costly up-front, but will pay-off down the road



9

MAKE THE MOVE TO MANAGED SERVICES

As infrastructure morphs into a hybrid setup, costs and complexity expand — often beyond initial expectations — and it becomes less likely an organization possesses the diverse range of skills necessary to manage a variety of infrastructures.

Aptum's Cloud Impact study revealed that 69% of organizations want to accelerate their cloud deployments, but need expert help. The data also suggests that cloud deployments without the benefit of expert guidance from a Managed Service Provider (MSP) can increase associated costs if cloud optimization is not a fundamental principle from the start.

"This has created healthy demand for MSPs and infrastructure service providers to help manage and optimize cloud usage on behalf of mid-tier and large enterprise organizations," says Philbert Shih, Managing Director of Structure Research. "Service providers have established economies of expertise through repeated trials and operational scale. And they are able to package all this know-how into a consumable service that end users can benefit from on day one."

"Organizations increasingly find that the journey to cloud is not always a straight path. Cloud adoption comes with a learning curve and it is often difficult to address these challenges with existing internal IT resources."

– Philbert Shih, Managing Director, Structure Research

KEYS TO MAKE THE MOVE TO MANAGED SERVICES:



- A MSP brings expertise and know-how to accelerate cloud benefits
- Nurturing cloud insight unlocks new approaches and services

CONCLUSION

There are several immediate, practical steps to address cost visibility and optimization in your cloud environment. If, however, they still don't drive the expected efficiencies or ROI improvement, a deeper review of the digital estate — such as architectural choices and even cloud platform choices — may be necessary.

It is also important to realize that some workloads will, indeed, cost more to operate in the cloud than on-premises. It's a myth that cloud is the perfect destination for all workloads. Rather, a hybrid IT model can help to maximize scalability, performance and security — and, of course, cost savings.

“Anything you do in the cloud will have repercussions, so it's important to validate your choices. That validation could be performance. It could be downtime. Or it could be cost obligations,” says Kwong Lum. “Don't tie yourself into a three-year contract, and then realize it's not something you truly want. Always do an analysis and assess the possible repercussions of certain actions you might take.”

To understand true total cost of ownership, and optimize the cost of your cloud investments, it's essential to consider all the components of cloud architectures. A great place to start is working with a cloud service provider that has hybrid expertise and a vendor-agnostic approach. That way, you'll ensure you harness the power of cloud to deliver greater cost efficiencies — without sacrificing reliability, security or performance.



ABOUT APTUM

Aptum is a global hybrid multi-cloud managed service provider delivering complex and high-performance cloud solutions with an integrated secure network. Using our Data As Infrastructure™ approach, Aptum solves complex technology challenges with total solutions and tailored options that drive tangible business outcomes and maximize the value of clients' technology investments. Aptum's cloud and global network solutions, underpinned with expert managed and professional services, offer genuine choice and adaptability with international reach spanning North America, Latin America, Europe and the United Kingdom.

Many of our services incorporate practical ways to control and optimize your cloud costs, and even save money, whether in business-as-usual processes or for more tactical purposes in the cloud. For example:



MANAGED MICROSOFT AZURE SOLUTIONS: We adopt a holistic approach, with experienced cloud architects and project managers who work directly with customers as trusted advisors. Aptum provides 24/7 support from Azure experts, and you gain control of your cloud spend with a single monthly invoice for your Azure services.



MANAGED AMAZON WEB SERVICES: Aptum's AWS experts ensure your cloud architecture is designed to meet your specific needs and optimize your investment. We provide a full range of services from building and testing your AWS environment, to migrating data and applications and expert ongoing management of your platform.



APTUM CLOUD CONNECT: Aptum provides a private, high-performance, totally secure connection between multiple clouds, your data centers and your users. A dedicated connection to your cloud applications eliminates uncertainties associated with security, availability, throughput and unpredictable latency.

For more information on controlling costs and optimizing your cloud architecture, visit [Aptum.com](https://www.aplum.com)