WHITE PAPER

Amazon Web Services Cost Surprises

Unexpected Costs Experienced by AWS Enterprise Customers

March 2019

Dao Research is a competitive intelligence firm working with Fortune-500 companies in the Information Technology sector. Dao Research is based in San Francisco, California.

Visit www.daoresearch.com to learn more about our competitive intelligence services.
Executive Summary

Customers have been attracted to Amazon Web Services (AWS) for many reasons, such as the ability to start small and grow as needed. And, embedded in this pay-for-what-you-use strategy, is a further claim that they will continuously lower costs. While this is a strong marketing message, the reality is that many hidden costs inflate the overall total cost of ownership of AWS. In fact, AWS can cost significantly more than customers expect and not just in one or two areas but for a host of technical and business-oriented reasons. As a result, some customers decide to reevaluate their options for other cloud providers who provide pricing transparency, simplicity, accountability, and governance.

There are three primary factors for these hidden costs. First, AWS pricing is complex. Unless you have a detailed understanding of every operational facet of every application or development environment, there is no way to properly budget your expenses. Second, the ease of provisioning encourages requests for more resources. In a typical enterprise, without strict management of policy enforcement and governance, hundreds and even thousands of resources are provisioned and often abandoned. Those unused resources continue to accrue charges. And lastly, although AWS enterprise agreements offer some discount relief, they actually exacerbate higher expenses than expected because they provide no usage oversight.

Based on Dao Research interviews with AWS customers, this is a reality for standard use cases, such as application development, analytics and data management, and even infrastructure services.

The key findings of the study were as follows:

- Study participants cited the ease of scaling up resources as a strength of AWS, but with a lack of proper planning this can result in runaway charges and stale resources that require manual effort to decommission. Several companies shared that they have dedicated resources, at a cost of $300K in annual salary, to monitor and manually decommission resources.
- Several AWS customers were surprised at the additional compute and storage required to achieve performance on par with their on-premises environments. They reported having to pay up to double the expected AWS fees to achieve performance relative to on-premises deployments.
- The costs of data transfer and storage request and retrieval were cited as known costs, but they were higher than expected due to scenarios such as retrieving data from long-term storage (Glacier to S3), application development environments frequently being torn down and rebuilt to save on storage and compute, and even charges for failed data transfers. In some cases, data transfer costs amounted to upwards of 30% of their total AWS cost.
- The cost of enterprise support came as a surprise to many customers. Some companies we spoke with found themselves discontinuing the support in large part because the support costs are based on the underlying cost of AWS services, which often increases unexpectedly for the aforementioned reasons.
- The billing for AWS in many cases does not provide the detail customers need to support charge-backs and multiorg/multiclient environments, and customers must overlay manual processes or other software systems to parse and reconstitute billing. One large enterprise shared they were spending $200K per year in maintenance costs for their custom billing solution on top of AWS.

"What started out as a few hundred dollars per month in AWS fees ballooned to over $20,000 per month within a year. Due to the pricing policies and high cost we experienced in a short amount of time, we had no choice but to shut it down and move to another cloud provider."

Implementation Partner Lead for Ad Tech Company
Introduction

The promise of the cloud is to transform IT economics, effectively outsourcing the datacenter responsibility to the cloud provider. This is particularly effective in development environments, where dev teams can rapidly spin up or down resources for dev and test. Historically, they have done this in “rogue” fashion by going outside of standard IT processes.

The expected benefits of consuming cloud services are to:

- Eliminate hardware purchases and hardware refresh cycles.
- Rapidly provision a managed resource.
- Start with smaller configurations and consume more resources as needed.
- Easily scale production workloads up or down depending on demand.
- Only “pay for what you use.”
- Enhance agility, in terms of rapidly implementing resources and adapting deployment size and scope, as well as taking advantage of new capabilities offered by cloud providers.

Some of the very same promises offered by cloud service providers can ironically lead to the erosion of the value proposition. In fact, many companies who signed up with AWS were surprised at the various ways in which costs were higher than their expectations going into a relationship with Amazon or as their deployment scaled.

To be fair, the issues were not all attributable to AWS itself but also to a lack of thoughtful planning and tooling that (a) would have made it easy to scale poor behavior, and (b) the lack of transparency to manage cost proactively.

Dao Research found several areas in which these issues occur:

- Runaway charges and difficulty monitoring costs based on services automatically launching or autoscaling.
- The need to increase infrastructure resources to compensate for lower performance compared to on-premises environments.
- Data transfer costs (e.g., zone to zone, cloud to external) and storage request and retrieval costs.
- Stale storage and compute resources from overprovisioning.
- The cost associated with enterprise support for business or mission-critical workloads.
- Operational costs and manageability resulting in costly labor and manual effort.
- Billing issues and determining costs based on organization and suborganization segmentation and the complexities of pricing based on location(s), volume, and service level.
Research Approach and Findings

The primary research phase consisted of an in-depth data collection and multiphase interview process. Dao Research identified and qualified eight customers and partners involved in AWS implementations inside medium and large organizations. These experts provided detailed primary research and data. The research focused on the procurement, operational, management, and usage experience and the associated cost surprises relative to expectations.

The research process and methods were as follows:

- Interviewed eight customers who shared their experiences trying, buying, using, and managing AWS from both an operational and a financial management perspective.
- Reviewed publicly available information and secondary research on a variety of AWS services.
- Researched industry blogs, support forums, and other online sources.
- Synthesized data and research findings to prepare this white paper.

Table 1 lists the companies analyzed and interviewed in this research study.

<table>
<thead>
<tr>
<th>Company</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>Healthcare Startup</td>
<td>CTO</td>
</tr>
<tr>
<td>Financial Services</td>
<td>Cloud Architect</td>
</tr>
<tr>
<td>Global Risk, Retirement, and Health Services</td>
<td>Executive Director, Technology</td>
</tr>
<tr>
<td>Regional Hospital</td>
<td>Director, IT</td>
</tr>
<tr>
<td>Banking and Credit Services Provider</td>
<td>Senior Business Leader, Strategy and IT</td>
</tr>
<tr>
<td>Home Goods Retailer</td>
<td>AWS Automation Engineer</td>
</tr>
<tr>
<td>Pharmaceutical and Medical Device Company</td>
<td>AWS Solution Architect</td>
</tr>
<tr>
<td>Ad Tech Company</td>
<td>Founder and CEO</td>
</tr>
<tr>
<td></td>
<td>Implementation Partner Lead</td>
</tr>
</tbody>
</table>
Customer Experiences with AWS

When deploying AWS, there are millions of potential considerations, and Amazon Elastic Compute Cloud (EC2) alone is only one of 90 services offered. Such AWS pricing complexity requires organizations to budget, buy, and consume their IT differently. Organizations must factor in costs beyond the core infrastructure as a service (IaaS), platform as a service (PaaS), and management services. Additional costs may also come from increased data transfers, greater network bandwidth required for the new service, ongoing AWS support costs, one-off implementation and migration activities, purchase of new tools, and changes to infrastructure management practices and internal team structures.

Participants in our study shared numerous examples of AWS cost surprises across several key use cases including application development, analytics and data management, and IaaS. They also shared some more general business-level insights and challenges in working with AWS relative to the billing and cost accounting processes and the effort required to track and monitor their AWS spend.

Runaway Charges

AWS, to their credit, makes it easy to scale dynamically without users actually knowing it, but this convenience can result in jarring costs, particularly if an organization lacked thorough planning or the dev teams are not in communications with the finance team. Based on customers we interviewed, it is not uncommon for a monthly bill to suddenly go from thousands of dollars to tens of thousands or even in excess of a hundred thousand dollars due to these considerations.

A cloud architect for a large financial services organization shared that it is easy in dev environments to incur a lot of cost associated with sprawl and scale. He told us, “If your automation is to tear things down—which is the one thing people tend to forget about—isn’t very good, maybe it just sits out there consuming money, right? You have all these people just ‘wild westing’ in your development environments. It’s been common in many large enterprises I’ve worked in with thousands of people. Nobody really knows each other’s name and they’re in there clicking buttons and the next thing you know you’ve got 100, 300, 500 instances and they’re just doing nothing.”

He went on to say, “I think that the true cost of development on Amazon always surprises people. I have never met a person who doesn’t eventually realize that it is not a free lunch. It scales with you nearly instantly. If you have bad planning, it’ll compensate for your bad planning. Right? That’s what it’s really doing for a lot of people.” From the magnitude of cost perspective for runaway charges, he shared that, due to the above, “you could burn through $100,000 or more within a single month.”
Another company we researched experienced runaway cost issues with respect to two elements of AWS pricing policies. First, an ad tech company experienced challenges with autoscaling by being limited in the percentage that AWS would scale up or down. For example, they wanted to scale up during periods of stress testing and scale down dramatically during periods of low usage in production. AWS policies would not allow them to do that. The other autoscaling “gotcha” is that some AWS servers, such as DynamoDB, are shared with multiple customers, who are charged for a prenegotiated partition, not actual usage. If a company uses up their contracted partition, AWS will automatically double the capacity and charge accordingly. All of these policies contributed to the ad tech company starting out paying a few hundred dollars per month for infrastructure services but within a year ballooning up to more than $20,000 per month. According to the company’s implementation partner, “We were just stress testing, not supporting millions of users or significant production capacity. Due to the pricing policies and high cost we experienced in a short amount of time, we had no choice but to shut it down and move to another cloud provider.”

Performance Costs Related to Compute and Storage

In terms of IaaS, several customers we talked to had to overprovision their AWS environments, relative to their existing on-premises datacenters, to achieve the same performance when moving workloads to AWS and not having compared the performance until after the fact. The performance impact was based not only on compute and storage but often on network bottlenecks. Many customers found that their AWS infrastructure cost can actually increase by 2x their expectation to compensate for the lower performance of AWS versus their existing on-premises environment.

An IT director for a global risk, retirement, and health services company found initially that performance was lower on the AWS cloud relative to their on-premises environment. This manifests itself in both the subscription provisioning and the cost of the infrastructure services but also in the disaster recovery (DR) plan. Specifically, the director said, “We couldn’t match apples-to-apples for the infrastructure that we owned, which is more dedicated compared to shared. So, we had to upgrade over time. So, let’s say for some more memory-intensive work, we went from 4 CPUs to 8 CPUs and from 16 GB to 32 or 64 GB of memory. So, while you could say that was an extra cost for us, it was actually double what we planned. And along with that it wasn’t just that because now you have to think about doing the same thing on the DR side.”

Based on this performance challenge example, the difference in cost for 100 instances of Amazon EC2 with the basic Linux configuration with 4 CPUs and 16 GB of memory (t2.xlarge) would be $16,427 per month, whereas with 8 CPUs and 32 GB of memory (t2.2xlarge) the cost nearly doubles to $32,855 per month.

An IT director for a regional hospital also cited the need to add more nodes/capacity to compare AWS with on-premises. He estimates that they would need 40% additional to meet the same on-premises demand. “I still have that gut feeling like if I upped my AWS capacity by 40% relative to my on-premises I still feel like it might be sluggish if I put on the same amount of load.”

IT Director for Regional Hospital

“I still have that gut feeling like if I upped my AWS capacity by 40% relative to my on-premises I still feel like it might be sluggish if I put on the same amount of load.”
Data Transfer and Storage Request Costs

AWS charges customers for storage request and retrieval and for transferring data from service to service within AWS services, as well as to external sources.

A research study by Centilytics1 found that data transfer costs can contribute upwards of 30% of the total AWS bill: “Data transfer into EC2 is usually free but, the outbound transfer has its fees. This is something that gets many new AWS customers to fall into the ‘unawareness’ trap. Generally, data transfer ‘OUT from Amazon EC2 to the Internet’ incurs substantial costs. These costs can rapidly mount up and keep on adding to your month to month cloud bills.”

The data transfer pricing varies widely by region as well, which can add to the cost if customers are unaware of this. For example, 1 TB of EC2 data transfer out to the internet costs $250 in Sao Paolo, compared to $90 in North Virginia (a 178% difference!). Several of our research participants cited challenges with data transfer costs. A cloud architect for a financial services company shared the unexpected costs associated with data transfer, particularly relating to their HA/DR environment. He said, “Then you have these other transitional costs that people don’t think about. So maybe putting data into S3 itself is kind of great. But suddenly, what you decide to do is you decide to build an HA/DR environment and you end up moving a lot of data in and around S3. Maybe you end up replicating it to other regions in S3 and all of those things are not great. So, they end up getting you coming and going literally in a lot of cases.”

He also shared that even when data transfers fail you still have to pay: “I’ve had situations where you start transfers, the transfers are huge and then suddenly there’s something wrong with the transfer and the next thing you know, you’ve actually spent a lot of money to move a bunch of data that’s now essentially bad. Right?”

With respect to storage request and retrieval, AWS customers were surprised by small but recurring fees beyond the basic storage fees. For example, there is a charge for S3 Glacier retrieval requests ranging from $0.025 to $10.00 per 1,000 requests, depending on the type of retrieval (bulk, standard, or expedited). And there is an additional fee for upload requests, at $0.05 per 1,000 requests. It all adds up to increase the total cost of AWS.

1 https://www.blog.centilytics.com/data-transfer-costs-everything-you-need-to-know/
The Cost of Stale Resources

A significant result of the ease of manual or automatic scale discussed earlier is the proliferation of resources that eventually become stale. Little to no automation exists for end-of-life resources, and determining how to decommission them is difficult depending on the type of environment, including dev, test, and production. Based on a conservative annual compensation average for AWS engineers, several companies in our study found this amounts to more than $200,000 per annum in additional cost for dedicated human resources to find and flag stale resources.

An IT director for a global risk, retirement, and health services company shared that development teams spinning up instances leads to proliferation of stale resources. He estimates about 20% to 25% of additional and/or wasted cost associated with this issue, saying, “Every time they do testing, they spin up a new VM in the new resource group. So that was a point where the bill continued and kept going higher and we were required to reroute it. Why? And this is what we found is that as much as it is efficient, it’s uncontrollable from a cost perspective.”

The cloud architect for a large financial services organization told us that most companies do not have adequate automation for end of life. He explained, “Many people never think about the far end of the life cycle like how to get rid of things so that they stop charging you. They build all this automation to put stuff in, they build all this automation to get stuff configured, but they just don’t consider what happens when it has to go. That actually is far more common in a development environment than people think about, right? That end-of-life cycle has to happen more in your development environment or you will have this out-of-control cost to run your development environment.”

An AWS automation engineer for a home goods retailer shared his experience monitoring his AWS environment with dedicated human resources to find and flag stale resources. He shared, “Obviously it has an effect on the cost and billing because when autoscaling is in place we need to worry about how many instances it is going to take and how it is going to run, and when we are not going to utilize it how it is going to be decommissioned. These are all things we need to keep in mind, and we need resources to manage these kinds of services. Right now, we have two engineers who are dedicated to the AWS services monitoring and manually decommissioning resources.” When factoring in AWS engineer salary cost plus an uplift for overhead and benefits, the cost for these resources is $300K per year.

“While the fact that you have to pay extra for enterprise support was not a surprise, we were quite surprised at how the cost of the enterprise support for AWS support increased based on the underlying AWS service fees. After the first year, we discontinued the service because we did not find the value above and beyond what our internal team could provide.”

Senior Business Leader, Strategy and IT, for Banking and Credit Services Provider
Costs of Enterprise Support

In addition to the subscription costs for the numerous AWS service offerings, those companies with more business or mission-critical operations were surprised to find that they needed to pay additional monthly fees to provide sufficient time zone or work hours coverage and to ensure service-level agreements for case severity and response times.

AWS enterprise support from the AWS website is “recommended if you have business and/or mission critical workloads in AWS.” The monthly costs for AWS enterprise support are as follows, with a starting monthly minimum fee of $15,000:

- 10% of monthly AWS usage for the first $0 to $150,000.
- 7% of monthly AWS usage from $150,000 to $500,000.
- 5% of monthly AWS usage from $500,000 to $1 million.
- 3% of monthly AWS usage for more than $1 million.

According to a senior business leader of strategy and IT for a banking and credit services provider, “While the fact that you have to pay extra for enterprise support was not a surprise, we were quite surprised at how the cost of the enterprise support for AWS support increased based on the underlying AWS service fees. After the first year, we discontinued the service because we did not find the value above and beyond what our internal team could provide.”

Operational Costs and Manageability

A key surprise area for many customers using data warehousing and analytics involved the amount of effort to implement, operate, and manage their Redshift deployments. Redshift customers experienced additional costs associated with initial implementation, data ingestion, tuning and optimization, and scaling Redshift up and down over time. Customers told us that roughly 25% of ongoing management effort is dedicated to performance tuning. These activities are performed by highly skilled database administrators who typically command salaries in the range of $90 to $130K per year.

Customers cited issues in both scaling up and scaling down Redshift. In terms of scaling up, the primary issue is that customers cannot scale Redshift in place but are forced to rebuild and configure the environment every time they scale up. Scaling down offers customers a similar cost surprise. According to customers we talked to, Redshift does not reclaim deleted space and, hence, does not reduce the cost of the service, which was the expectation going in. Moreover, customers cannot scale Redshift compute and storage independently, which leads to paying needlessly for additional compute or storage.

Customers also found Redshift lacking in some key data warehouse features. One example is the lack of enforced uniqueness in the data

---

2 https://aws.amazon.com/premiumsupport/pricing/
Cost Surprises for Customers Using AWS

structures. Instead, customers have to manually enforce rules for each application. Another aspect related to security and a limitation for customers who maintain multiple datasets for their unique clients is that Redshift requires configurations of multiple virtual private clouds (VPCs) to contain customer data.

Other Redshift customers shared the effort and workaround solutions required to bolster the performance of Redshift to meet their needs related to query performance. For data scientists running more complex queries, they were forced to offload to a separate Hadoop-based platform (EMR) due to performance issues, implement a caching solution on top of Redshift to do column encoding, build additional data structures, and add aggregation strategies to improve the user experience. According to our study participants, all these extra efforts and workarounds introduce complexity and add cost to Redshift deployments.

Billing Challenges

In addition to the use case–specific cost surprises discussed in previous sections, nearly all study participants noted issues centered around both understanding and effectively using the AWS monthly bill. Unlike traditional IT, which is centrally managed, the cloud is consumed in a distributed manner by many business units. This results in many accounts throughout the organization and is harder to manage. Several companies we talked to built their own solutions to provide the needed granularity for departmental use or cross-charging within their organizations. These homegrown solutions end up costing these companies additional time, effort, and costs in the five-figure range to build and maintain these tools. The alternative is to pay for a third-party AWS cost management tool, which is typically a percentage of the monthly AWS bill. CloudCheckr, for example, charges enterprises 2.5% of their monthly Amazon cloud bill.

Companies used Amazon’s VPC service frequently to segment their AWS environments when they were servicing multiple clients. However, the AWS bill did not give them an accounting at the VPC level but rather a conglomerated view. The use of ECS and containers also was found to magnify the billing and accounting challenges. The chief technology officer of a healthcare company said that the AWS bill does not provide adequate granularity especially with multiple VPCs: “If you're a very diverse environment like this healthcare company, we have 14 VPCs but when you look at your bill in Amazon, they don't give you 14 VPCs in the bill. They give you a list of your aggregated resource use.” Billing issues are further compounded when using ECS: “It actually gets a little weirder when you get up into things like ECS because now you have containers consumed inside of a VPC that are possibly sharing resources across microservices across companies. Instead of your application being monolithic and contained only in a single VPC, perhaps a shared set of resources is spread through microservices, woven from VPC to VPC.”

He further went on to share that the AWS bill does not provide adequate granularity especially in complex and dynamic environments: “I have worked with some financial institutions who will spin up in the course of a day hundreds of containers and spin them down at day’s end. Over the course of 30 days you will have potentially tens of thousands of containers that have come and gone. You are now well outside of that threshold where you as a human being can go through and manually audit what has now resulted in your bill. Right? Because the bill is aggregated together, if you had 14 tenants spinning up tens of thousands of instances and they're now gone, right, you would have to essentially have someone go through all of the logs, all of that process and try to, from those logs, glean what was done. Right? Really seriously, one mistake in an environment like that could be enormously expensive to a company. You could burn through $100,000 or more in a single month.”

An IT director for a global risk, retirement, and health services company shared that “given the billing complexity of AWS, we needed to leverage manual Excel spreadsheets and the capability of PeopleSoft Financials for cross-charging, regional currency, and cost differentials.”

The cloud architect for a financial services company told us that many customers develop their own capability to understand and utilize their AWS bill. He explained, “I have worked for companies who have written their own
billing product doing cloud lifts because of this, to wrap AWS. So now you have to have this auxiliary tool to basically deconstruct your bill to control these costs per VPC because you’re in a tenant model, right? Each tenant has a VPC.” He went on to say that building the initial billing solution required a team of three to four developers for roughly eight months, and he estimated that the annual cost of supporting it, complete with some backend processing in the cloud, was $200K per year.

Ironically, one customer shared the extreme case regarding the level of effort required to understand the AWS bill as nearly impossible in a monthly billing cycle: “You could find yourself in a situation where if you got a log that in 30 days you got surprised by a huge bill, it may take you more than the next 30 days to just analyze the log you got in the first place.”

These problems are exacerbated by the fact that the AWS native Billing and Cost Management console does not provide plug-and-play integration with existing enterprise accounting systems. To mitigate this, enterprises must either manually link their existing accounting systems to a third-party cloud budget management system’s API or use a third-party integration tool.

“I have worked for companies who have written their own billing product doing cloud lifts because of this, to wrap AWS. So now you have to have this auxiliary tool to basically deconstruct your bill to control these costs per VPC because you’re in a tenant model, right? Each tenant has a VPC.”

Cloud Architect for Financial Services
Conclusion

The perception of many customers is that AWS is the uncontested leader in the low-cost, pay-as-you-go cloud model. This may have been true in the early days of simpler cloud services, fueled by small developer teams without corporate oversight; however, as enterprises adopt cloud services pervasively, the hidden direct and indirect costs can surprise and sometimes overwhelm the expected costs.

These unanticipated costs are in large part due to the focus of AWS to easily enable and allow organizations to quickly spin up services but with a lack of transparency, accountability, billing granularity, and mature enterprise capabilities to decommission or effectively scale down deployments. The rapid pace at which AWS rolls out new services and gets their customers to consume them further exacerbates the underlying cost considerations. Some specific examples of unexpected costs cited by customers in this study include the following:

- Runaway costs of infrastructure scaling, with customers sharing examples of monthly fees jumping from thousands or tens of thousands to more than $100K per month.
- The need to increase provisioning of AWS cloud infrastructure, as compared to on-premises, significantly increases costs. Companies shared they needed to pay for twice the expected AWS service fees based on the need to increase the provisioning.
- Costs of data transfer and storage request and retrieval can be upwards of 30% of the monthly AWS bill, and these costs can vary significantly based on region.
- Several companies have dedicated resources, representing more than $200K in annual salary, to monitor and manually decommission stale resources.
- The cost of enterprise support increased based on the underlying AWS service fees.
- Additional operational costs are associated with performance tuning, with as much as 25% of management effort of highly skilled experts going toward ongoing performance activities.
- Billing-related costs associated with paying for third-party services such as CloudCheckr add nearly 3% on top of the AWS bill or significant manual efforts using spreadsheets or other accounting applications to better track and make sense of the AWS costs and effectively perform chargebacks. A large enterprise customer spent eight months to develop their own internal billing solution on top of AWS, which now requires $200K in annual cost to maintain in going forward.

It is quite a revelation that there are a multitude of Amazon and third-party tools designed to help customers manage the cost of AWS, with a half dozen AWS native tools alone for cost assessment and management. A recent Tech Target article\(^3\) highlighted the extreme case of the wayward cost of AWS services, saying, “Dropbox, a company built on AWS and an enormous user of cloud resources, saved a net of almost $75 million over two years when it moved storage off of Amazon S3 and onto internal, custom-designed infrastructure. While most enterprises don’t have million-dollar monthly AWS bills, nor Dropbox’s need for capacity, this example shows how blind usage of AWS can result in massive overspending.”

\(^3\) [https://searchaws.techtarget.com/feature/AWS-cost-management-tools-wrangle-unwieldy-expenses](https://searchaws.techtarget.com/feature/AWS-cost-management-tools-wrangle-unwieldy-expenses)
Organizations must factor in costs beyond the core IaaS, PaaS, and management services. Further costs may also come from increased data transfers, greater network bandwidth required for the new service, ongoing AWS support costs, one-off implementation and migration activities, purchase of new tools, and changes to infrastructure management practices and internal team structures. In conclusion, they should evaluate other cloud providers who provide pricing transparency, simplicity, accountability, and governance.