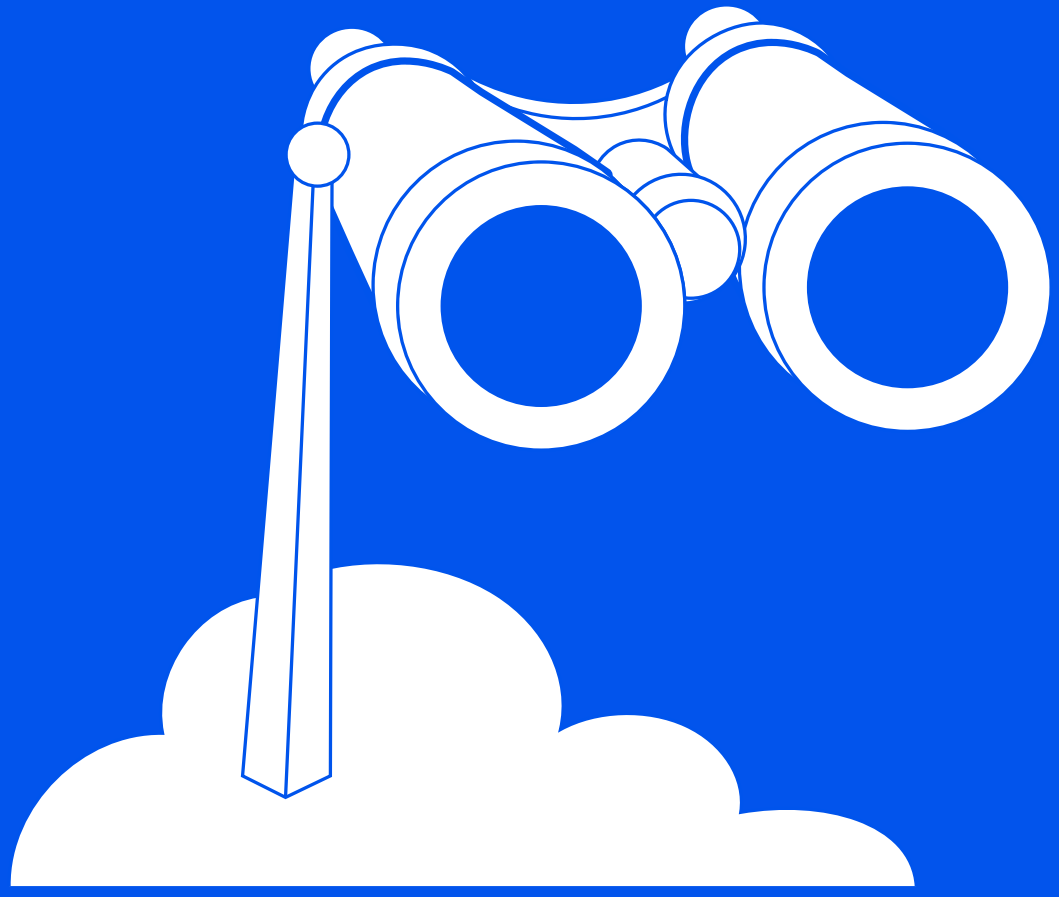
2023 State of the Cloud



Whitepaper

Executive summary

With cloud technology constantly evolving and growing increasingly critical to business operations, the responsibility of security professionals to stay abreast of the state of the cloud has never been greater in order to proactively address potential threats and ensure the safe and secure deployment of cloud solutions.

Over the past year, we have observed how cloud adoption has continued to grow with more organizations increasing their footprint in the cloud. Many new capabilities were introduced, with the number of possible API calls increasing by 15% in AWS, 20% in Azure, and 45% in GCP.

Although new services and their corresponding APIs expand the possibilities of how the cloud can be utilized, they can also broaden attack surfaces and create more challenges for cloud defenders. According to our data, 57% of companies use more than one cloud platform and therefore require greater knowledge and expertise from their security teams who need visibility into multiple platforms as well as the interfaces between them.

Besides novel cloud risks, well-known prevalent risks such as data exposure are also of concern. For instance, our data shows that 47% of companies have at least one database or storage bucket publicly exposed to the internet, and an attacker can discover and access an exposed bucket with a guessable name (e.g. “wiz-backup”) in less than 13 hours.

In this data-driven report, based on our scanning of over 200,000 cloud accounts, including more than 30% of the Fortune 100 environments, we analyze the latest industry trends and developments, presenting a factual and data-based assessment of the current state and progression of cloud technology. We examine how the cloud has evolved over the past year and attempt to shed light on some of the complexity of cloud environments, including aspects such as organizational usage of multi-cloud and both managed and non-managed services. We hope this can help cloud builders and defenders ensure they have the visibility and tools that they need to continue their cloud growth and protect their company’s assets. In addition, we will review notable cloud threats from last year and provide insight into the speed of compromise of misconfigured environments.

The year in review

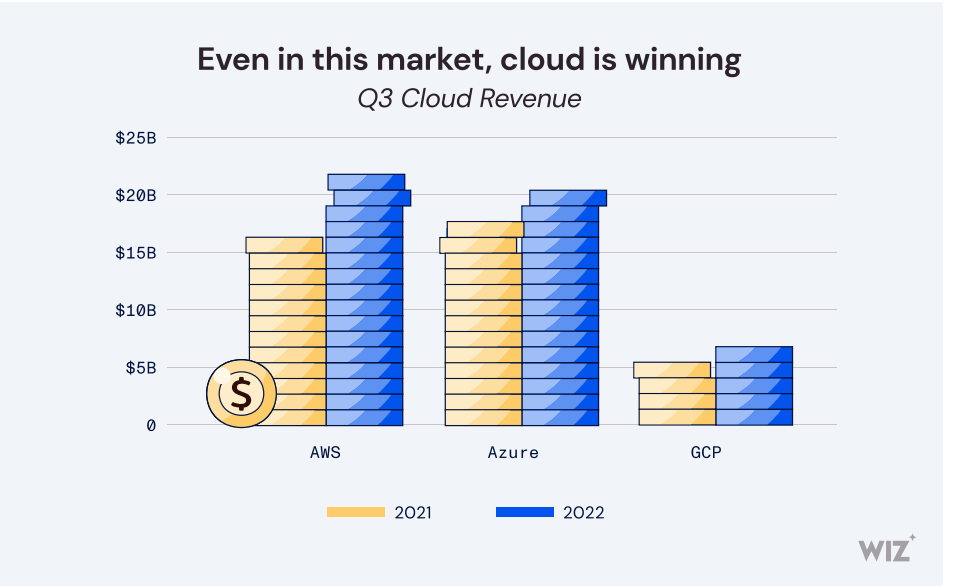
The sudden upsurge in 2021 of researchers attacking cloud vendors continued into 2022. A project sponsored by Wiz called cloudvulndb has been collecting these incidents, in part to look for patterns. Wiz's researchers found several critical cross-tenant vulnerabilities in multiple cloud providers (see AttachMe, Hell's Keychain, ExtraReplica, and more). By using the insights from the past incidents and our researchers' expertise, we developed the PEACH cloud isolation framework to offer guidance on better securing not only the cloud providers, but any PaaS or SaaS solution running multi-tenant environments.

Threat actors are becoming more proficient at attacking cloud environments. LAPSUS$–one of the most brazen–used its access at companies to move laterally through their cloud environments. Our guidance on mitigating the risks from this group is based on existing industry best practices but tailored to this threat actor's particular activity against cloud environments.

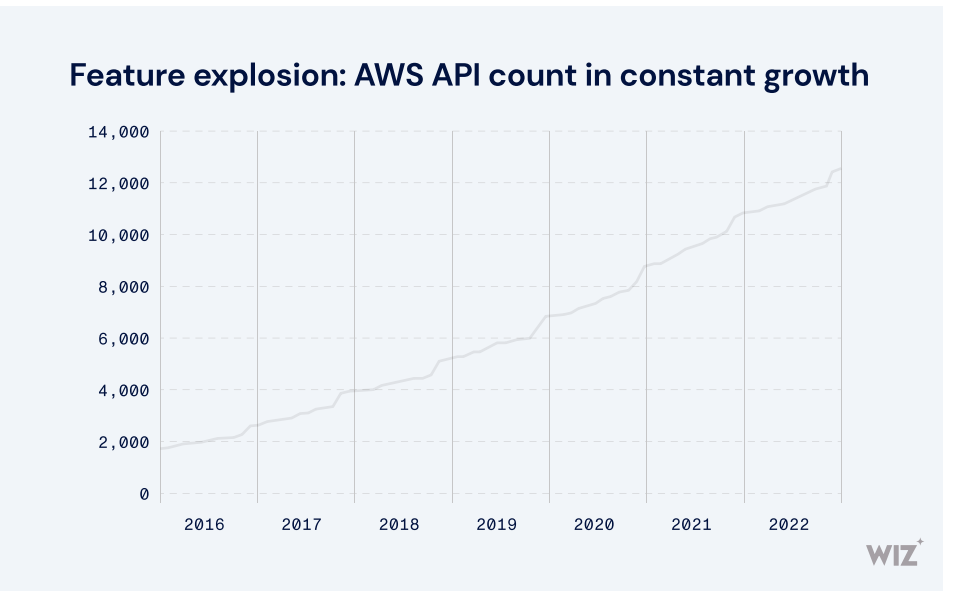
AWS's instance metadata service version 2 (IMDSv2) gained traction in a few ways this past year. A default deployment of an EC2 uses an older version of this service (IMDSv1) which does not mitigate SSRF and related attacks when the host contains other vulnerabilities. Even though a more secure version was released in 2019, the default IMDSv1 is still common (most likely for legacy support). AWS's GuardDuty received a new detection for when IAM role credentials are suspected to be stolen from an EC2 via this service. This risk is still relevant, as demonstrated by the threat actor UNC2903 targeting IMDSv1. A number of vendors also made changes to allow customers to enforce IMDSv2. Finally, malware built to run specifically inside AWS Lambda functions was discovered in the wild. Although this concept has been known for years, threat actors only just started using it.

Current landscape

Cloud usage continues to grow. Companies are shifting more of their workloads from on-prem to the cloud and both adding and expanding new and existing workloads in the cloud. According to the 2022 Q3 earnings of the top three cloud providers (AWS, Azure, and GCP), revenue increased across each cloud provider by a minimum of 20% from Q3 in 2021.

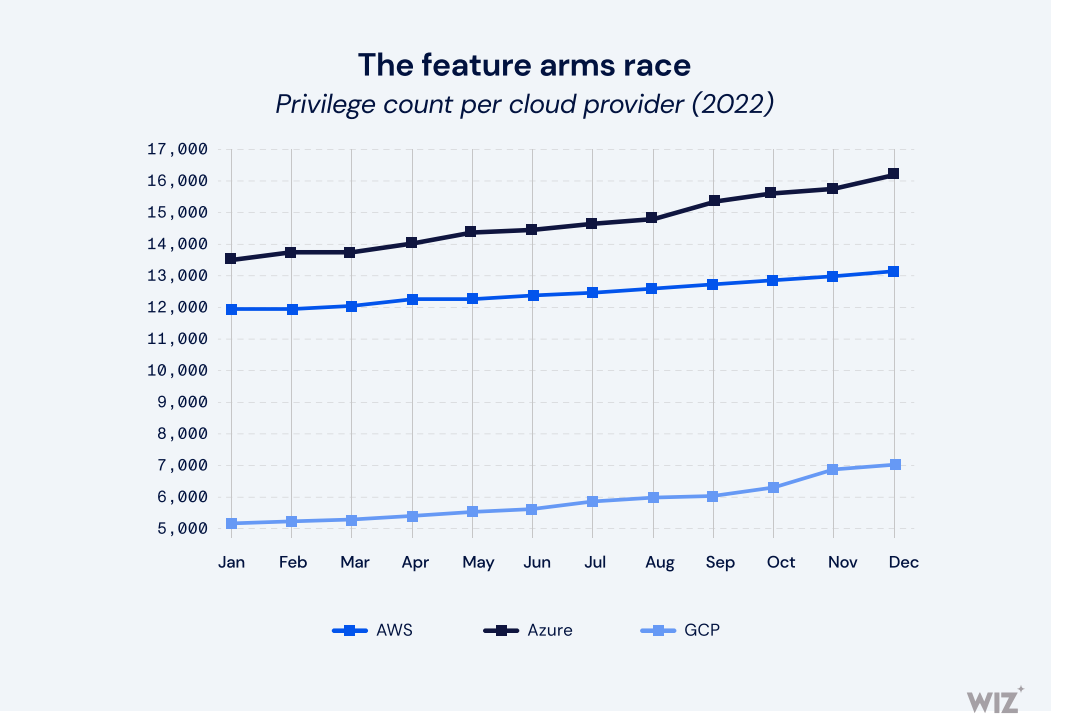


Cloud providers keep increasing their offerings and their complexity. In addition to growing in size, cloud providers are also becoming more complex. AWS has added APIs at a steady pace, with about 40 new services and 1600 new actions per year for the past 6 years[1]. The yearly spikes are due to the annual AWS re:Invent conference where they release large numbers of new features.



[1] The API counts data was obtained by walking the commits of botocore.

Additionally, the privileges available to control API access have increased in the past year across the top three cloud providers by 15% for AWS, 20% for Azure, and 45% for GCP[2].



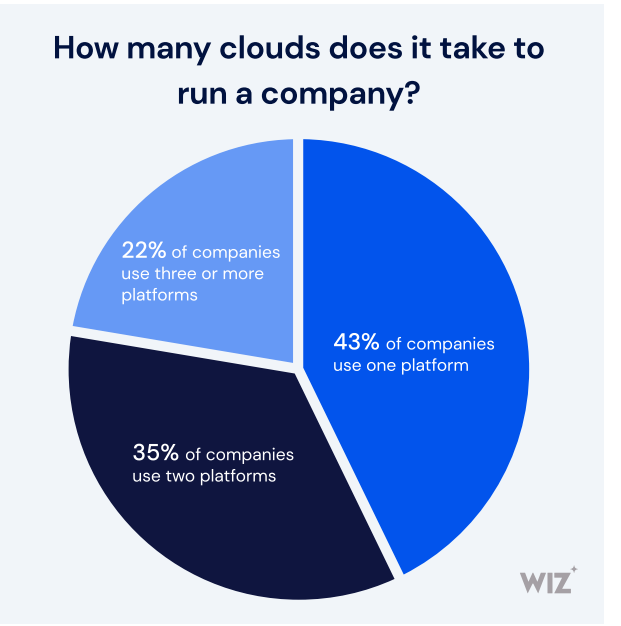
Our data set

This report is based on our scanning of over 200,000 cloud accounts (AWS, OCI, Alibaba cloud accounts, GCP projects, Azure subscriptions) including more than 30% of the Fortune 100 as customers.

Cloud usage

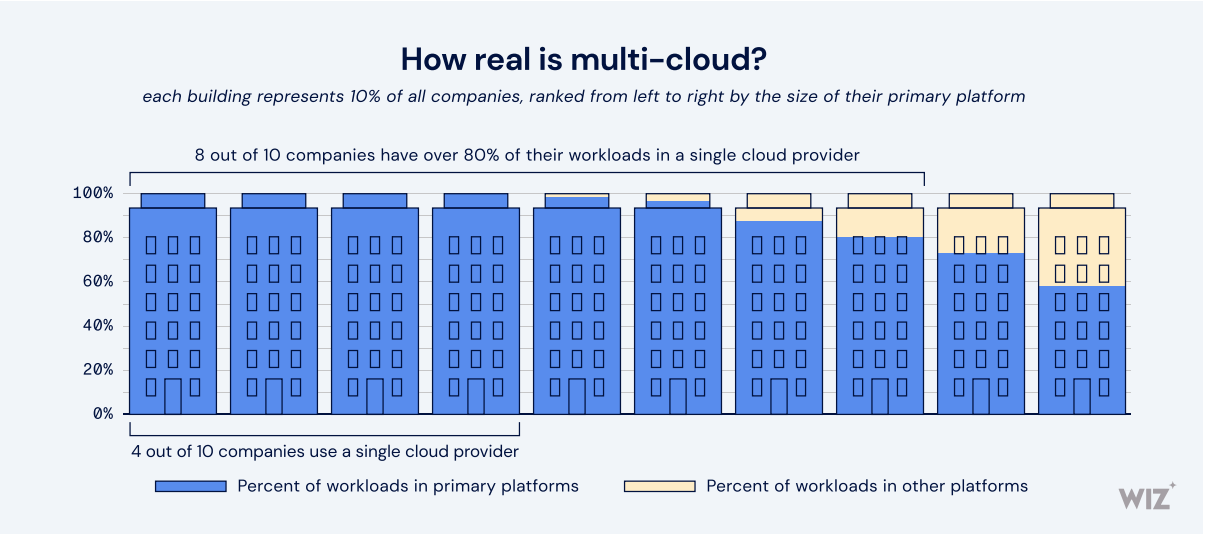
Is this the year of ~~Linux on the desktop~~ multi-cloud? According to our data, the idea of multi-cloud as a single architecture that spans multiple cloud providers is uncommon.

Most companies are only on one cloud and in cases where they are using multiple clouds, the majority of their workloads are on one cloud provider. In fact, about 43% of customers operate entirely on one cloud.



[2] Privilege counts were acquired from https://github.com/iann0036/iam-dataset.

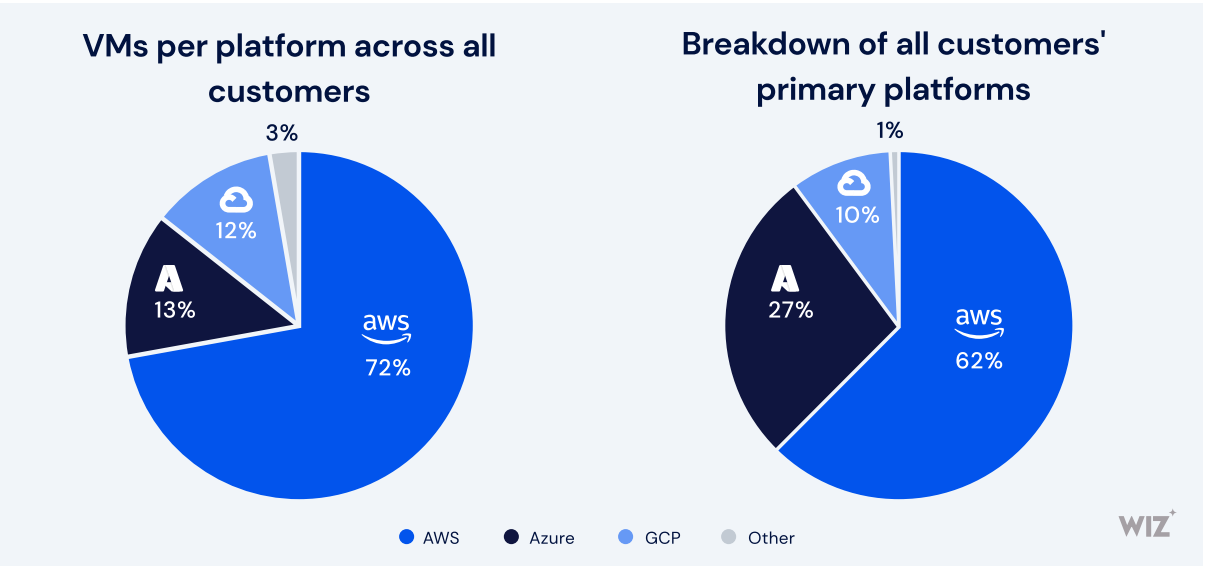
When examining how many workloads each customer is placing in each cloud provider, our data shows that about 78% of customers have over 80% of their workloads in a single cloud provider. In the following diagram we have sorted our tenants by the percent of their workloads in their largest cloud provider, and then bucketed these into deciles.



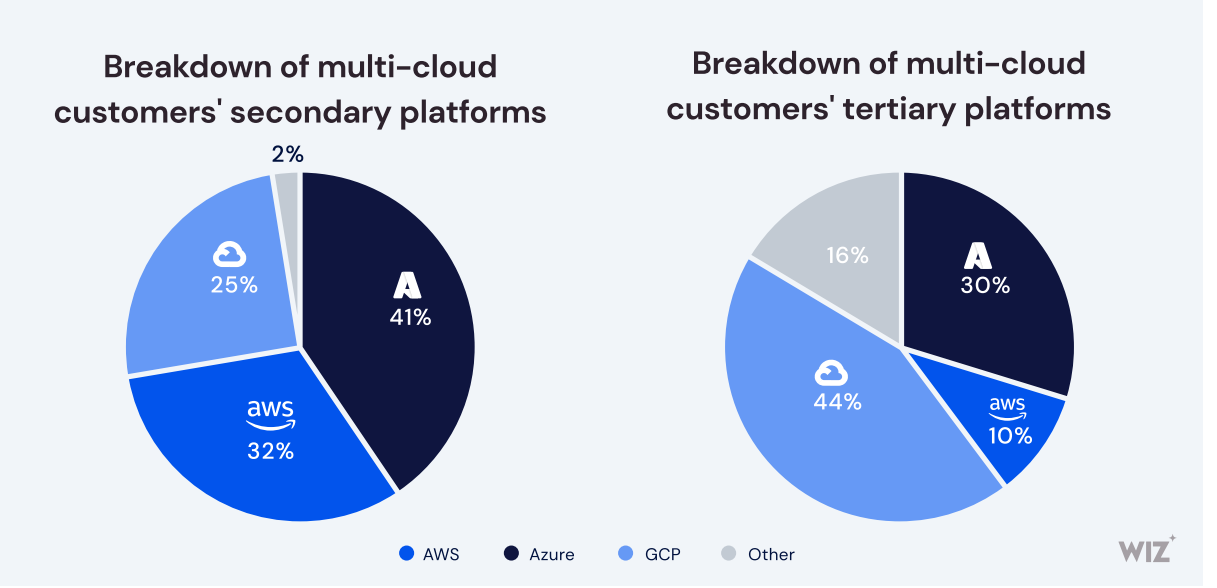
Most cloud customers concentrate their workloads in a few large accounts. Nearly all companies running on AWS have multiple AWS accounts, but the vast majority of these companies have a few disproportionately large accounts alongside many smaller ones. For over 97% of customers using AWS, the largest 5% of their accounts contain over 50% of their workloads.

In other words, although most AWS customers do not maintain a single monolithic account in the strictest sense, they do use a handful of what might be considered monolithic accounts.

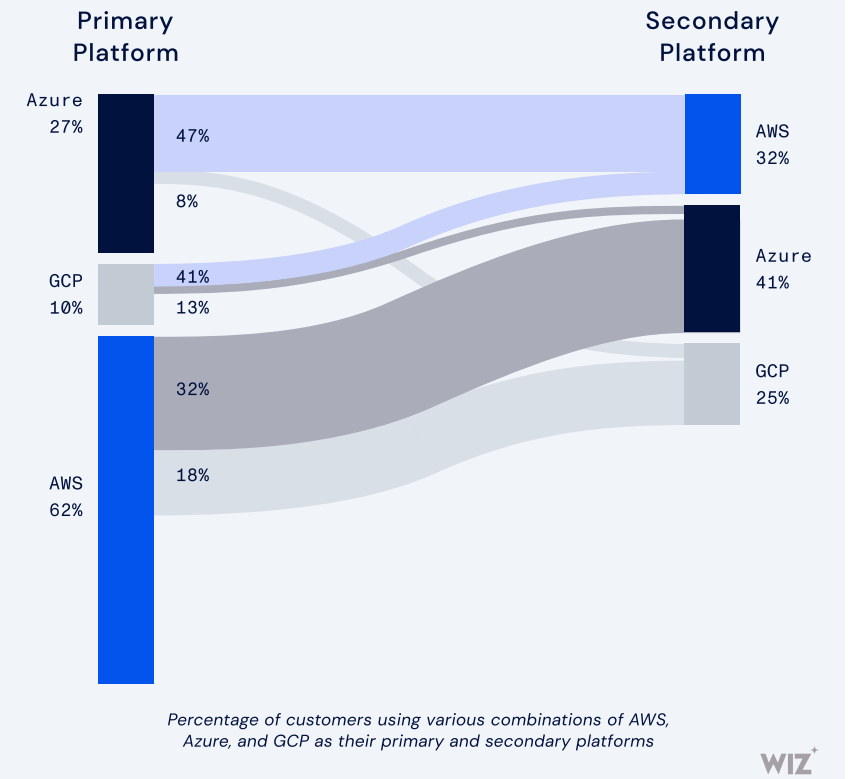
AWS is the most common platform, Azure the 2nd, and GCP the 3rd. Most workloads (in this case, virtual machines) across all companies are running on AWS (72%), and the majority of companies (62%) choose to place more of their workloads on AWS than on other cloud providers. In other words, no matter how we look at it, AWS is the most common primary platform among cloud customers.



We can also break this down further and examine which cloud providers our customers are using as their secondary and tertiary platforms (i.e. their 2nd largest and 3rd largest platforms, respectively). We can then observe that among companies using more than one platform, Azure is the most common secondary platform (41%), whereas among customers using more than two platforms, GCP is the most common tertiary platform (44%).



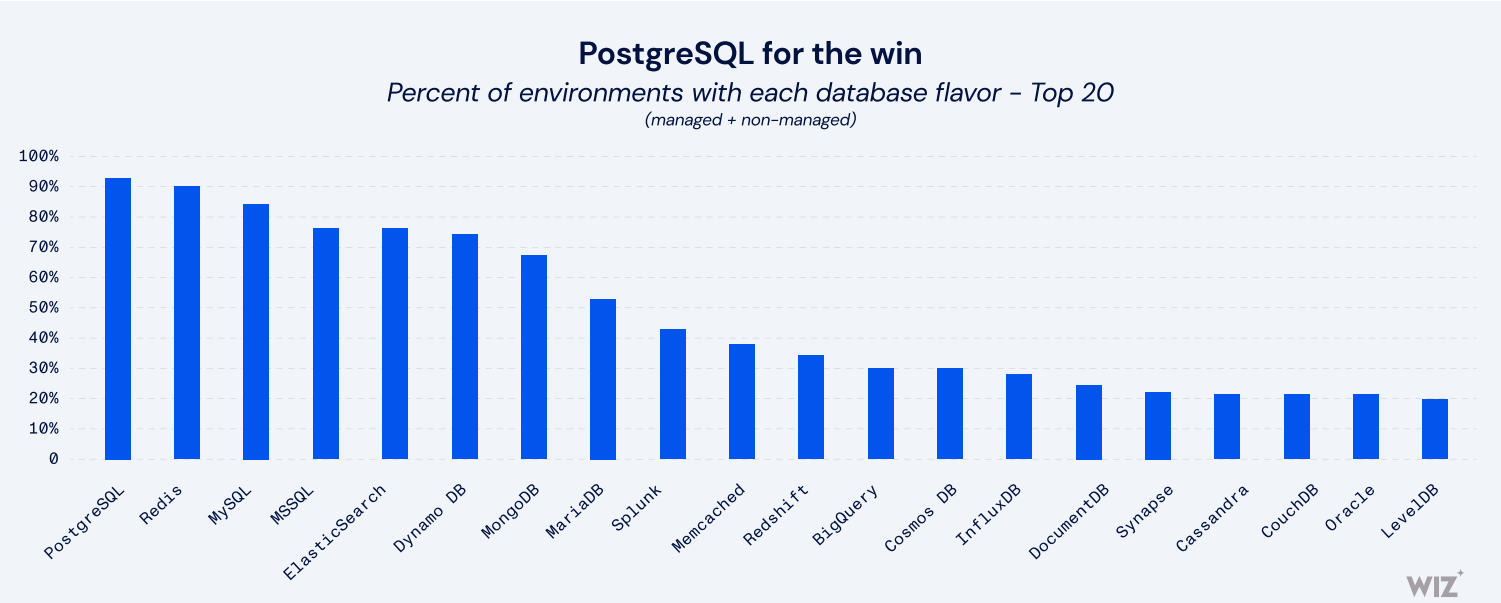
Another way we can analyze how customers are using multi-cloud is by checking combinations or pairings of different platforms. The following Sankey diagram breaks down primary platform usage on the left side and secondary platform usage on the right side, while the middle section shows the percentage of companies using each combination of two platforms. For example, while 27% of customers have most of their workloads in Azure (as mentioned above), 47% of this group are using AWS as their secondary platform, and 8% are using GCP.



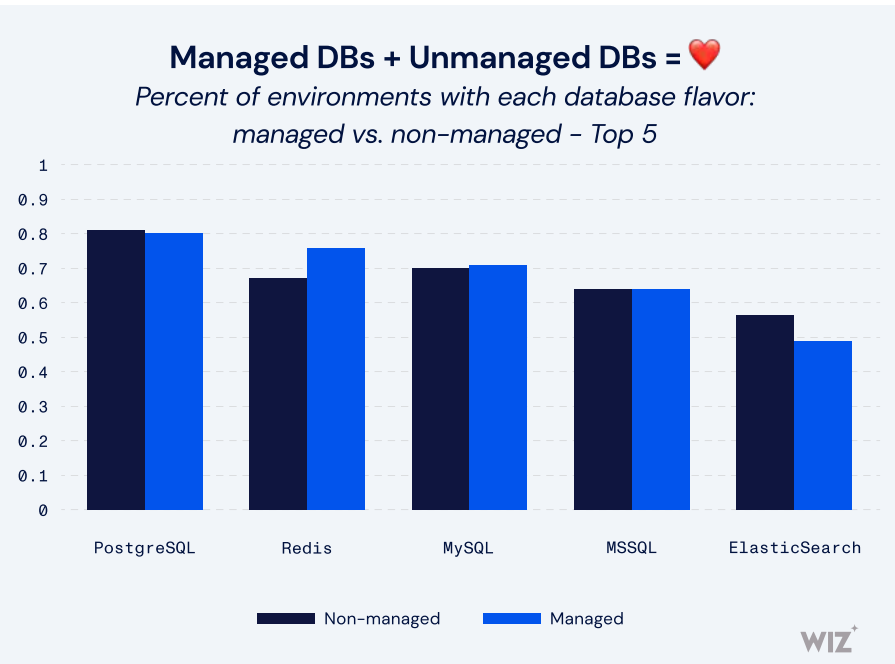
Percentage of customers using various combinations of AWS, Azure, and GCP as their primary and secondary platforms

Companies are using a healthy mix of managed and non-managed databases. Over 90% of AWS customers use managed database servers (for example PostgreSQL running on RDS), while that number is 87% in GCP and 82% in Azure. However, over 91% of companies have non-managed database servers running on IaaS (for example MySQL manually installed on an EC2), with only 6% of companies using managed database servers exclusively, and 90% using a mix of managed and non-managed.

As to which database flavors are being used, PostgreSQL, Redis, and MySQL are the most prevalent, with over 90% of companies having at least one instance of a PostgreSQL server (whether managed or non-managed).



Looking at the top 5 most prevalent database flavors, we can see that each of their managed and non-managed offerings are more or less evenly represented. For example, 76% of companies have at least one managed instance of a Redis database, while 80% have at least one non-managed instance. There doesn't appear to be any clear preference for managed or non-managed databases in terms of usage.



Data exposure

Data leaks are reported in the news every week. Attackers are aware of the value of sensitive data and the increasing difficulties in securing it. They continuously scan the internet for exposed databases and buckets. With the average cost of a data breach now over $5 million according to IBM's Cost of a Data Breach 2022 Report, eliminating this risk should be a top priority.

The risk of data exposure is shockingly common. 47% of companies have at least one database or storage bucket exposed to the internet (either managed or non-managed), and over 20% of those cloud environments with publicly accessible buckets have buckets that contain sensitive data.

Moreover, 13% of cloud environments have at least one publicly exposed non-managed database server, whereas for managed databases that number is 32%.

Exposed resources are compromised within hours. In experiments we ran where we created S3 buckets with names we assumed attackers might be targeting – taking well-known company names and adding "-backup", "_logs", etc. – we spotted attempts to list the contents of the S3 buckets in as little as 13 hours. In a similar test, we created an S3 bucket with an unguessable name, but referenced it in a commit to a public GitHub repo. Attempts at listing it occurred within 7 hours. This indicates to us the speed with which attackers could potentially find and exfil a publicly exposed S3 bucket.

From exposure to discovery: how fast do hackers find open buckets?

Attackers are searching for exposed resources. Exposed buckets with common names are discovered within 13 hours, while exposed buckets referenced in GitHub repos are found within 7 hours.