Censys

Cloud Misconfiguration Mayhem

An Analysis of Service Exposure Across Cloud Providers

About Censys

Censys was started as a research project at the University of Michigan by the creators of ZMap and inventors of fast Internet-wide scanning. Today, Censys, Inc. is led by a team of industry and academic security and networking leaders, and provides both best-in-class Internet data and Attack Surface Management.

Learn more about Censys at: www.censys.io.

Who is Censys Labs?

Censys Labs is the research arm at Censys. We are a multidisciplinary group of individuals whose goal is to produce valuable and actionable insights into Internet-wide security trends to inform different stakeholders in the industry. We focus our research efforts on:

- Current risks associated with cloud environments
- Attack surface security trends and assessments across industries
- Impact of vulnerabilities and zero-days across the Internet

Table of Contents

Author: Megan DeBlois

Other Contributors: Michael Lopez, Jadon Montero, Derek Abdine, Zakir Durumeric, Justine Desmond

Design by: Allina Liu

Date: March 2021

Executive Summary

The increase and ease of cloud computing has created significant security challenges that every organization is trying to effectively manage. The cloud has amplified the age old problem around ensuring you have confident inventory of all your assets and that they are secure. This research aims to provide data-driven insights into service exposures on the cloud that can have severe business consequences in a few impact areas: data breaches, malware, ransomware, and services that may leave organizations vulnerable to attacks like credential stuffing.

This research leverages point in time data from our Universal Internet DataSet (UIDS) on March 8, 2021. We analyze service exposures across a dozen popular cloud providers in the industry, including the 2020 Gartner Magic Quadrant Cloud Providers.

Key Findings:

- **Our research identified nearly 2 million database exposures across cloud providers.** We found 1.93 million Internet-facing database services during our research and include the following database services: MySQL, Postgres, Redis, MSSQL, MongoDB, Elasticsearch, Memcached, and Oracle Databases.

- **We found more than 1.9 million RDP exposures across the dozen cloud providers we investigated.** Internet-facing RDP services have posed significant risk across the industry due to the rise in ransomware attacks. The total RDP exposures, however, only made up a small percentage of total services observed in all dozen cloud providers we investigated at only 2.2%.

- **Users of OVH were more likely to expose MySQL database services relative to other providers and users of Tencent were significantly more likely to expose RDP services.** We calculate the rate of prevalence of a service per 100,000 hosts in a provider to control for the sheer size differences in cloud infrastructure. Controlling for size, this metric gives us interesting insights into user behavior across cloud providers when it comes to exposures like database and remote administration tools like RDP and SMB.

More details and methods around how the research was conducted can be found in the remainder of this report. The purpose of this research was to provide insights into the prevalence of service exposure by practitioners using a variety of cloud environments which can be indicative of good, or bad, security hygiene on the cloud.

Some of the findings in this report may indicate differences in defaults by provider, as well as controls or the maturity of those controls that the cloud provider offers. We also consider usage bias by customers of different providers, meaning some users may be more likely to use a certain service over another, which may account for differences in the prevalence rate. Although our research team conducted some due diligence in terms of reviewing publicly available documentation and some default configuration testing among outliers revealed in the analysis, we did not thoroughly test all services across every provider. This is a potential area of future research.

1.15 million MySQL services were observed among the dozen cloud providers and made up the majority of database exposures.

Introduction

At Censys, we continually monitor the Internet for changes in behavior among exposed devices. The complexity, diversity, and ephemeral nature of the Internet today is much different than 20 years ago. More organizations have embraced the cloud, increasing their attack surface complexity and changing the shape of the Internet today.

Internet-facing assets can be particularly vulnerable to attack if in an unknown or unmanaged state. Findings from the 2020 Verizon Data Breach Investigations Report suggest that **"it might not just be an asset management problem, but also a vulnerability management problem on the assets you did not realize were there".** This illuminates a problem that is becoming increasingly challenging which is ensuring you have visibility of an expanding cloud environment and that there are not vulnerable services exposed to the Internet that could be consequential for your organization.
We analyze security hygiene in the cloud through the lens of service exposures to better understand user practices when it comes to network configuration across the cloud. We conduct a comparative analysis of the prevalence of service exposures in a dozen cloud providers.

Our main audience for this report is CISOs and practitioners alike to help them make data-driven decisions around which service exposures are most common in the cloud and could potentially impact their organization.

The initial concept for this work was inspired through our team's preliminary analysis of attack surfaces of the Fortune 500, assessing their cloud environments from an outside perspective. Based on publicly **available data, we found organizations are using on average 25 different cloud providers in their ecosystem.**

Impact of Exposed Services

Ensuring cloud assets are properly protected is critical for any modern IT risk management program. According to a recent survey conducted in 2020 by Fugue Inc., **misconfiguration in the cloud is the number one cause of cloud-based data breaches, with nearly half of respondents reporting that practitioners spend more than 50 hours per week managing cloud misconfiguration issues** [1].
Cloud misconfiguration is a broad topic, but for the purposes of this research we narrow our discussion to service misconfigurations such as database exposures or remote administration services like RDP and SMB.

Our definition of exposed services for our research are:

- Services that should normally be within the network perimeter; and
- Services on the network perimeter which, if misconfigured, allow potential unauthorized access to internal infrastructure.

[1] Fugue State of the Cloud Survey 2020

We selected these services by associating them with specific areas of impact to the business. The two areas of impact explored are:

1. **Database Exposures** as they are directly related to data breaches and data loss; and
2. **Remote Administration Services** which are often associated with malware, ransomware or leave businesses susceptible to credential stuffing attacks.

These two impact areas are significant due to the growing number of data breaches and ransomware attacks that have emerged over the last year, particularly with a growing remote workforce and the public cloud expansion we see today. Table 1 provides an overview of these impact areas and related service exposures.

Impact	Description	Service Exposure
Area 1: Database Exposures Vulnerable to data exposure or loss	Exposure or loss of internal, system-wide, or customer data.	MySQL, MSSQL, Postgres, Redis, MongoDB, Elasticsearch, Oracle, Memcached
Area 2: **Remote Administration Services** Vulnerable to things such as: - Malware / Ransomware - Credential Stuffing	Malware / Ransomware: Any software designed to cause damage to an organization's service(s). Ransomware is used to seize data or services and demand payment in exchange for their return. Credential Stuffing: Attempting stolen credential combinations against services requiring authentication to gain access.	SSH, RDP, SMB, Telnet, VNC, Team Viewer, PC Anywhere

Table 1, Impact Areas and Related Service Exposures

Methodology

This research leverages point in time data from our Universal Internet DataSet (UIDS). Data in these experiments is from March 8, 2021. We aim to answer the following questions as it relates to service exposures on assets in the cloud.

How much service exposure is present on different cloud providers?
What is the prevalence rate of exposure across providers, controlling for size?
What are potential explanations around cloud customer practices as it relates to varying service exposure rates of prevalence?

Data Fidelity: Censys data is collected via our Internet-wide scanning engine from three providers across the globe in the U.S., Europe and Asia. This multi-perspective scanning provides us visibility of **over 99% of the Internet**. Censys also can see 97% of services we scan, independent of which port they are running on. This is highly valuable for our research, so that we find services running in the cloud, even if running on nonstandard ports. The Censys Platform refreshes services in less than 72 hours on average, which is also important to ensure our data is relevant and fresh. We identified cloud user host IP addresses from our internal attribution system. We also filtered out **honeypot-like hosts** and hosts that have services that respond on every port.

Cloud Provider Selection: For our research we define Cloud Service Providers (CSPs) as providers who offer Infrastructure as a Service (IaaS). This is to capture the cases where customers are responsible for cloud network configuration and the configuration of services that they deploy. We categorize cloud providers into two groups to compare differences:

1. The 2020 Gartner Magic Quadrant 7: These include all very commonly used providers in the industry: Amazon AWS, Microsoft Azure, Google Cloud, Alibaba, Oracle, IBM, and Tencent.
2. **Other Cloud Providers:** We determined these providers based on market share and known cloud adoption and usage. These providers include: OVH, Rackspace, Choopa, Digital Ocean, and Aptum.

A Comparative Analysis of Service Exposure by Cloud Provider

The services we discuss in this section do not inherently imply misconfiguration or vulnerability. It is also important to note that all these configurations are the responsibility of the customer or practitioner, not the cloud service provider. As noted in the methodology, we also tried to increase data fidelity by removing honeypots, or hosts running more than 100 services.

Our focus areas for analysis are:

- **Total Count of Services by Provider:** Understanding the total amount of exposure of different services on the cloud
- **Prevalence Rate of the Service Compared Across Providers:** Normalizing the data and controlling for the difference in size, we compare the prevalence rate of services per 100,000 instances across each of the cloud providers.

Through our analysis, we discovered that the prevalence rates across providers vary. We posit some potential reasons for certain outliers and include documentation directly from the cloud service provider in English where available or explicitly state when we stood up test instances to check for default settings and configurations

Aggregating Cloud Instances and Services

Censys identified over 23 million hosts and over 89 million services, including services on non-standard ports across a dozen popular cloud providers.

Top providers with greatest numbers of hosts include:

Provider	Total Hosts
Amazon AWS	11,939,341
Google Cloud	2,873,424
Alibaba Cloud	2,091,473
Microsoft Azure	1,691,757
Digital Ocean	1,539,909

Table 2, Top 5 Cloud Providers by Total Hosts

A further breakdown of the total hosts can be found in Figure 1. Amazon AWS makes up the largest portion of cloud infrastructure across the providers we investigated with the next three being Google Cloud, Alibaba, and Microsoft Azure.

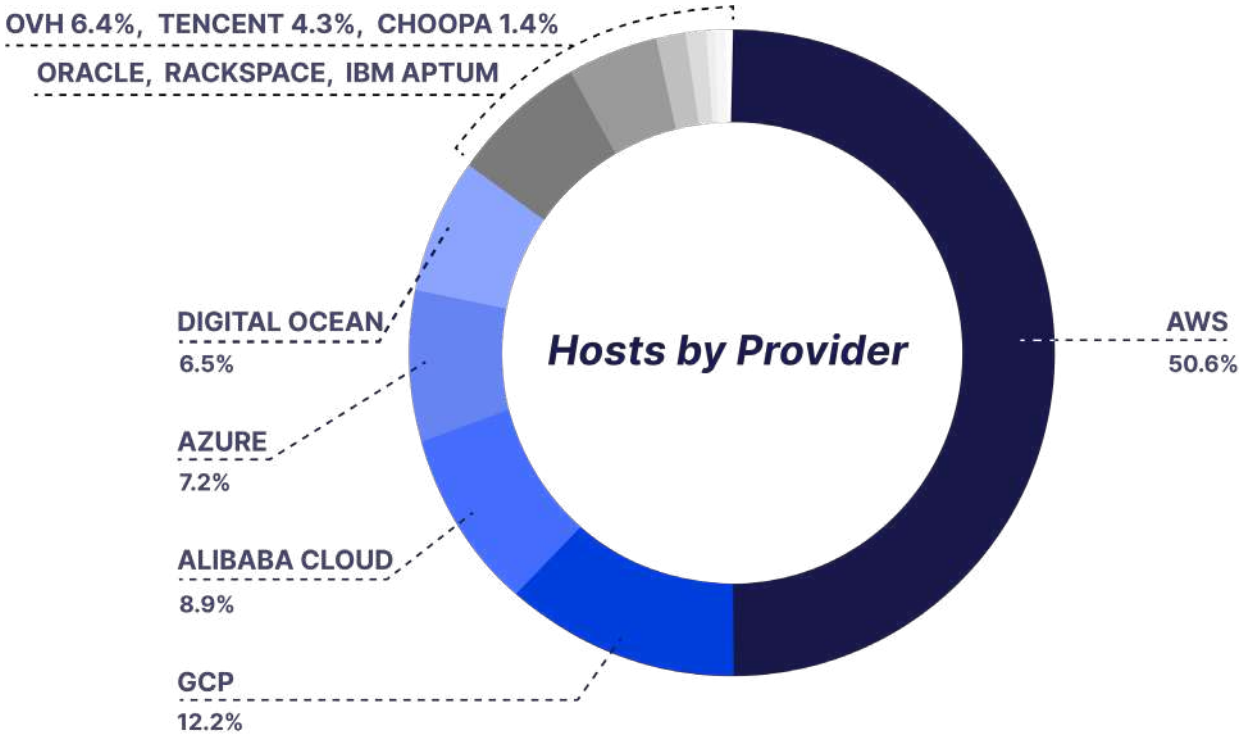

***Figure 1**, Total Hosts by Provider of the 12 Cloud Providers*

We've identified the number of services we see across the dozen providers below. The 5 providers with the greatest number of services include:

Provider	Total Number of Services
Amazon AWS	49,043,804
OVH	10,149,812
Alibaba Cloud	6,800,239
Microsoft Azure	6,103,424
Google Cloud	5,443,596

Table 3, Top 5 Cloud Providers by Total Number of Services Observed

A breakdown of the total service counts can be found in Figure 2 with Amazon AWS again making up more than half of the services we observed across the 12 providers.

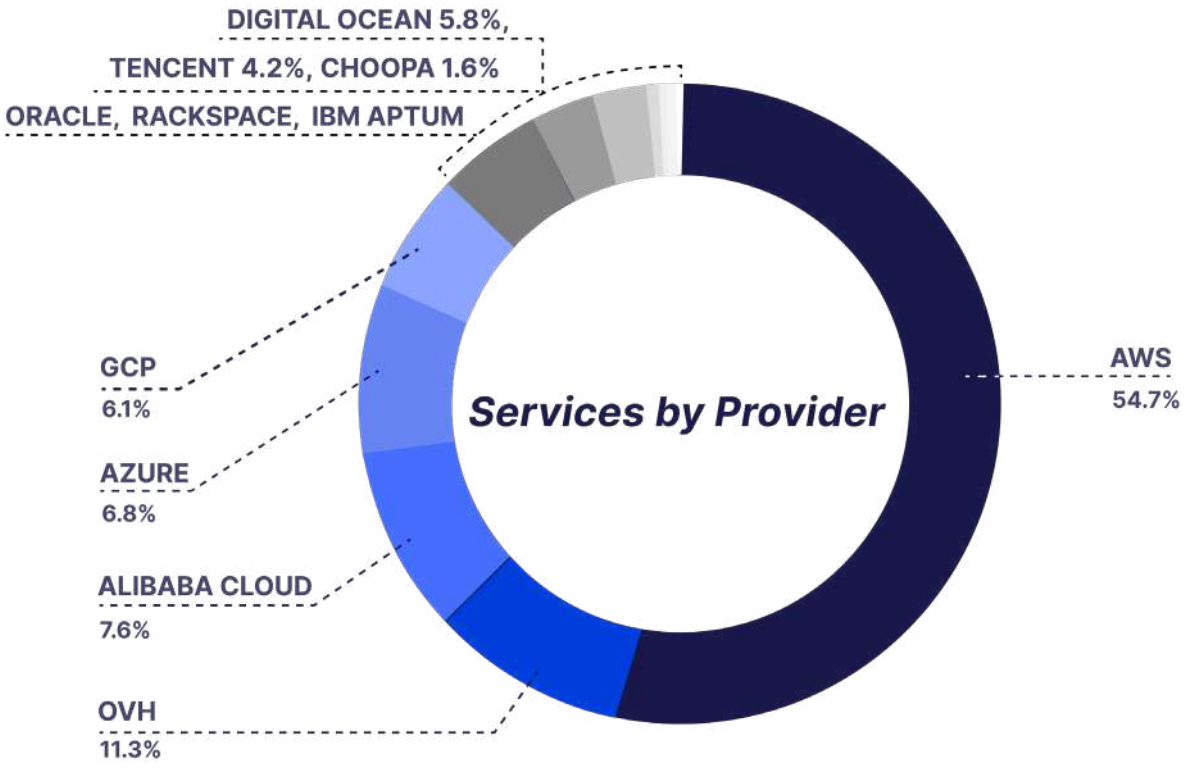

Figure 2, Total Services by Provider of the 12 Cloud Providers

As illustrated in Figure 1 and 2, it is important to account for the sheer size of the cloud provider since each company has a different scale of operations on the Internet. In the next sections, we will detail findings across the two impact areas discussed, looking at top related service exposures and their prevalence rates across the cloud.

Exposed Services Related to Data Breaches

The first impact are services related to database exposure. We selected this impact area due to the large amount of data breaches that blemish so many enterprises around the globe and at an alarming rate. Gartner claims that "by 2022, 75% of all databases will be deployed or migrated to a cloud platform, with only 5% ever considered for repatriation to on-premises"[2] . Given cloud trends Gartner mentioned above, as well as what we know about the concern many organizations have today around data breaches, we determined this was a very relevant area to explore. MySQL, Postgres, and Redis had the highest total service counts across the dozen providers. To narrow down our analysis further, we decided to more closely look at these top 3 across all providers.

We analyze the following database services across the dozen providers:

- MySQL
- MSSQL
- Postgres
- MongoDB
- Elasticsearch
- Redis
- Memcached
- Oracle

MySQL, Postgres, and Redis had the highest total service counts across the dozen providers. To narrow down our analysis further, we decided to more closely look at these top 3 across all providers.

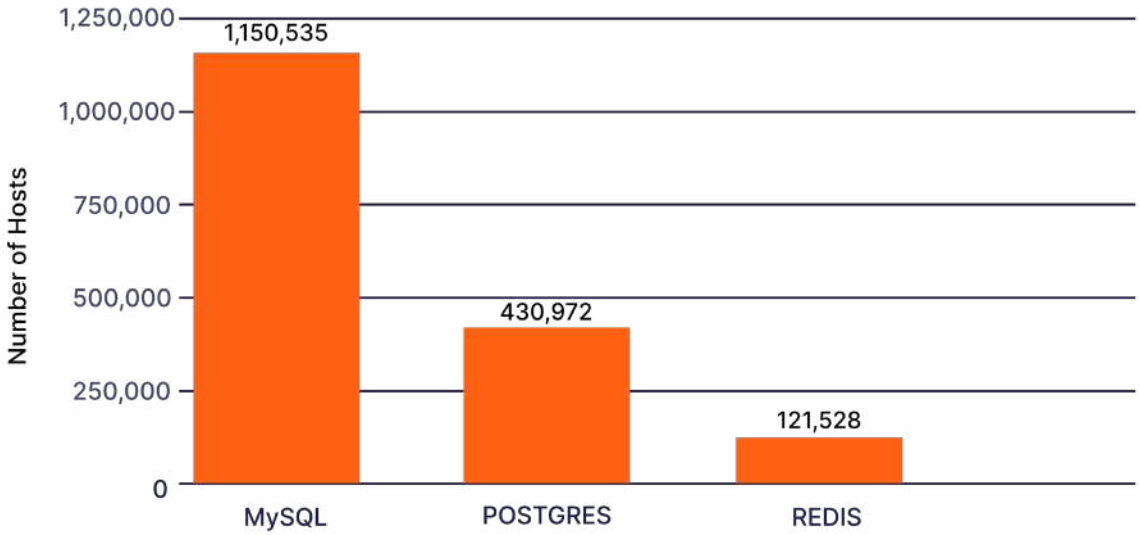

Figure 3, *Count of Top 3 Exposed Services by Impact Area: Data Breaches*

Figure 3 shows total service counts of the top three database exposures across the dozen providers. MySQL services made up the majority of database exposures with over 1.15 million MySQL services exposed. Postgres followed with more than 400,000 exposures and Redis with over 120,000.

[2] Gartner Says the Future of the Database Market Is the Cloud

MySQL

MySQL is a free and open source relational database commonly used across systems and cloud providers to store data. This database was the most commonly exposed database across cloud providers we investigated for this research. Figure 4 highlights MySQL service counts by provider from highest (left) to lowest (right).

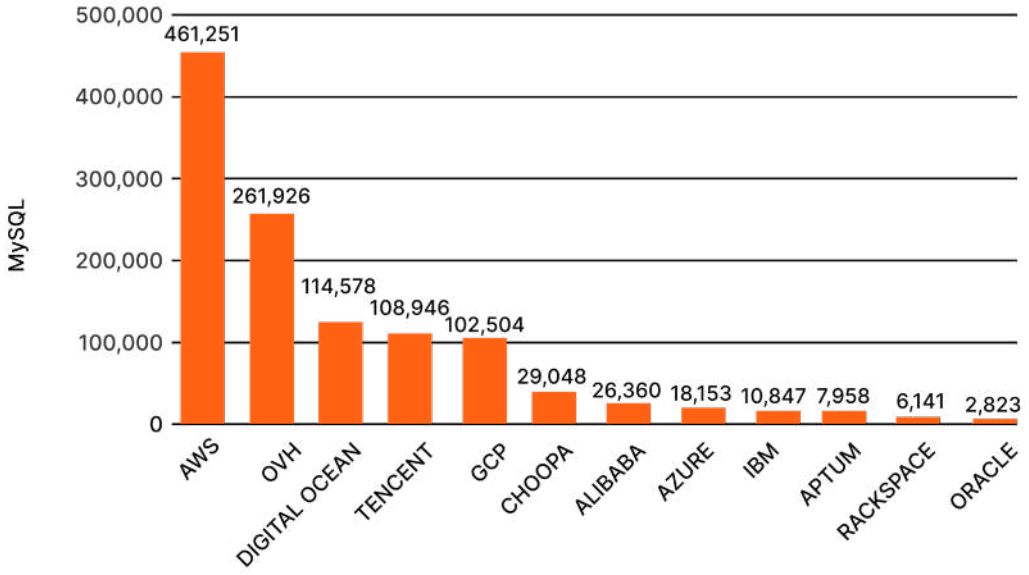

Figure 4, *Service Count of MySQL Exposures by Cloud Provider*

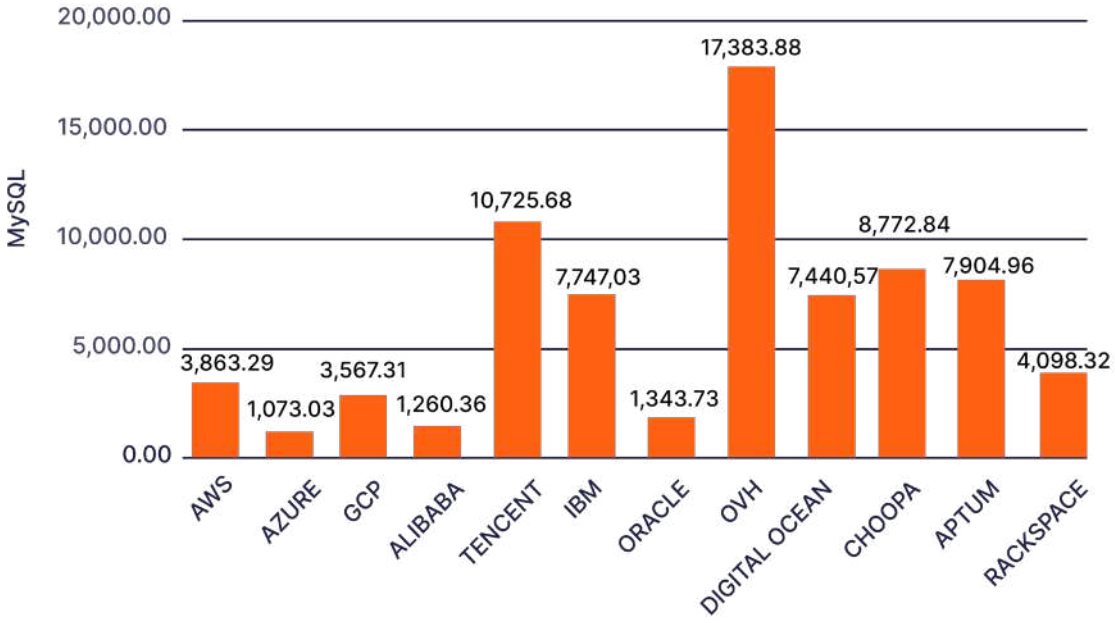

Figure 5, *Prevalence Rate of MySQL Exposures per 100,000 hosts by Each Cloud Provider*

Our team determined the prevalence rate of services per 100,000 hosts in order to conduct a comparative analysis across providers. This allows us to see which providers have higher rates of service exposure, controlling for the difference in total infrastructure differences. In Figure 5, the prevalence rate of MySQL service exposed per 100,000 hosts can be found below for all 12 providers, Gartner Magic Quadrant CSPs of 2020 on the left.

OVH is an outlier when looking at the prevalence rate. This means that users of OVH are more likely to expose a MySQL database relative to users of other cloud providers. We could find no obvious reason in their documentation as to why OVH users are more likely to expose a MySQL service than other cloud providers. One possible explanation could be usage bias in that customers of OVH may be more likely to utilize services that leverage MySQL than others.

Postgres

Postgres, also called PostgreSQL, is a free and open source relational database that can be used across systems and cloud providers. This particular database exposure was the second most common observed across all dozen cloud providers.

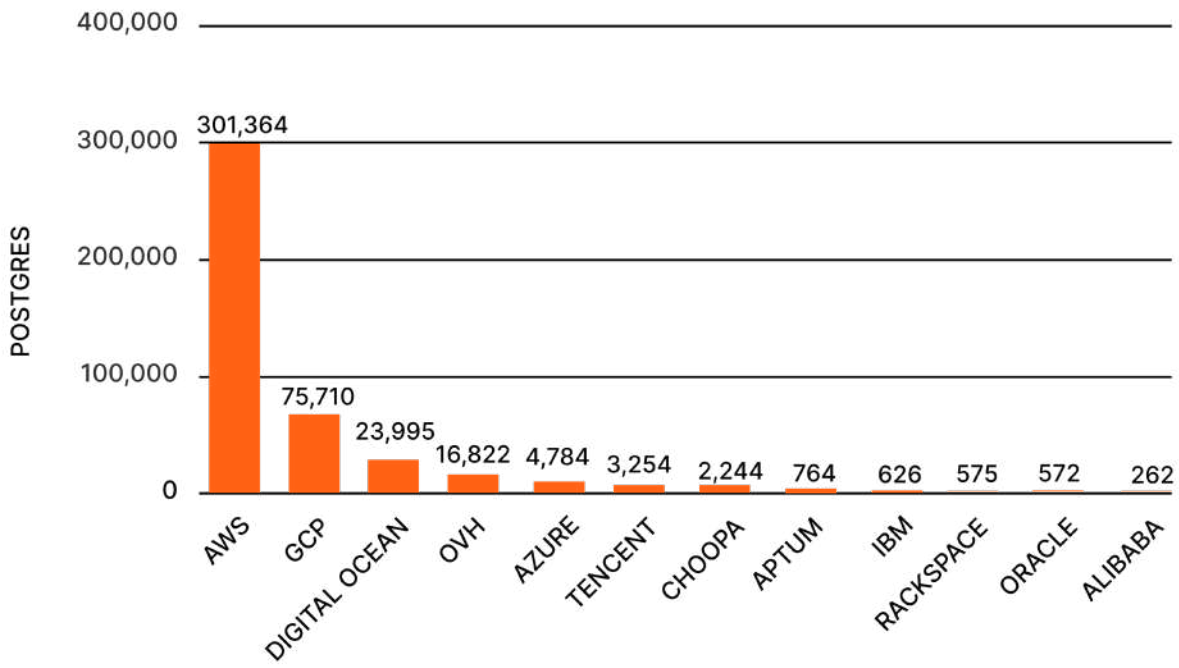

Figure 6, *Service Count of Postgres Exposures by Cloud Provider*

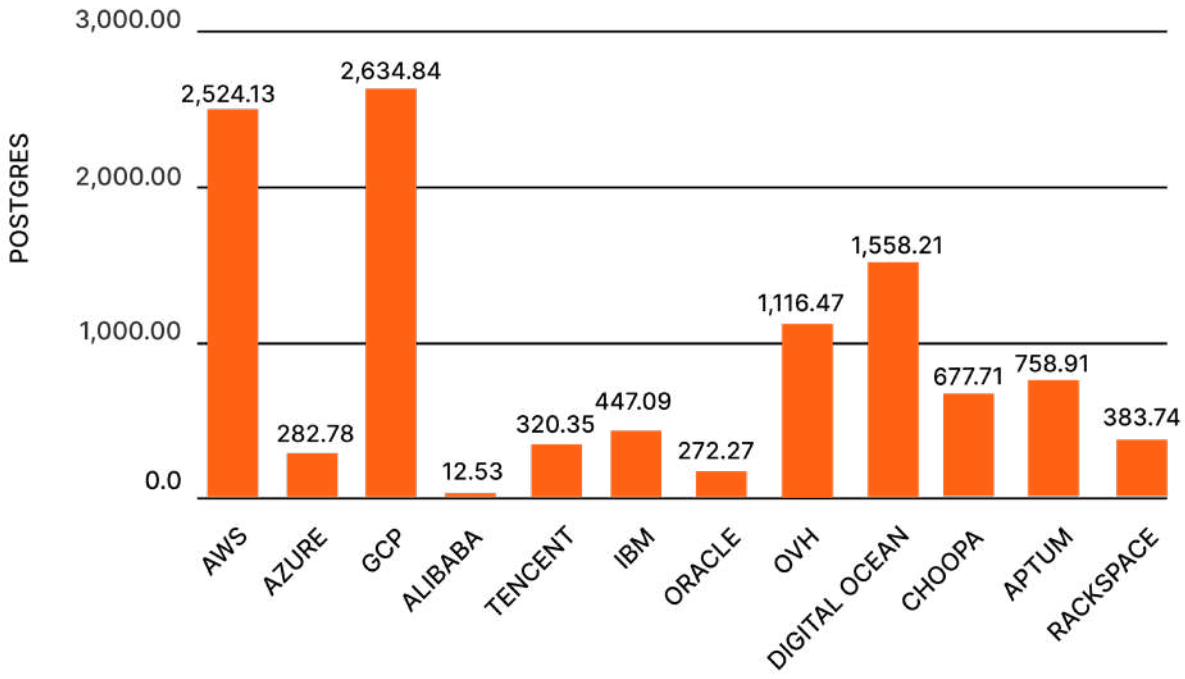

Figure 7, *Prevalence Rate of Postgres Exposures per 100,000 hosts by Each Cloud Provider*

The aggregate service counts by provider in Figure 6 help us understand the total amount of Postgres exposure in the cloud per provider. AWS is consistently showing higher counts of all services due to being the largest cloud provider on the Internet, and it's well known managed Postgres and MySQL offering, RDS. Google Cloud, Digital Ocean, and OVH all have higher counts of exposed Postgres services than Microsoft Azure.

The prevalence rate of Postgres exposure per 100,000 hosts is quite different from MySQL findings. Compared to all other other providers, AWS and Google Cloud both have significantly higher rates of Postgres database exposures.

To try and understand why AWS and GCP were outliers, we created PostgreSQL database instances on AWS and Google Cloud to review default setup configuration and pre-selected options. At the time of writing, we found that both default setting selections were toggled to assign a public IP address to the database, but with no inbound network connection. More configuration is required in order to allow inbound network connections to the database. In addition, both AWS and Google Cloud have their default authentication setting set to a password. This may be an easier and quicker user experience to setup the database, but databases are often used to store sensitive information be it company-related or PII. Good security practice dictates database services should not be Internet-facing generally speaking.

Another factor to consider is that there may be customer usage bias, meaning more users on AWS and GCP use Postgres databases compared to other providers. See Figure 7 for details and further comparison of providers.

Redis

The third most common database exposure in our research was Redis. Redis is also open source, but is used as an in-memory key-value store. Deployment is recommended on Linux systems according to Redis documentation.

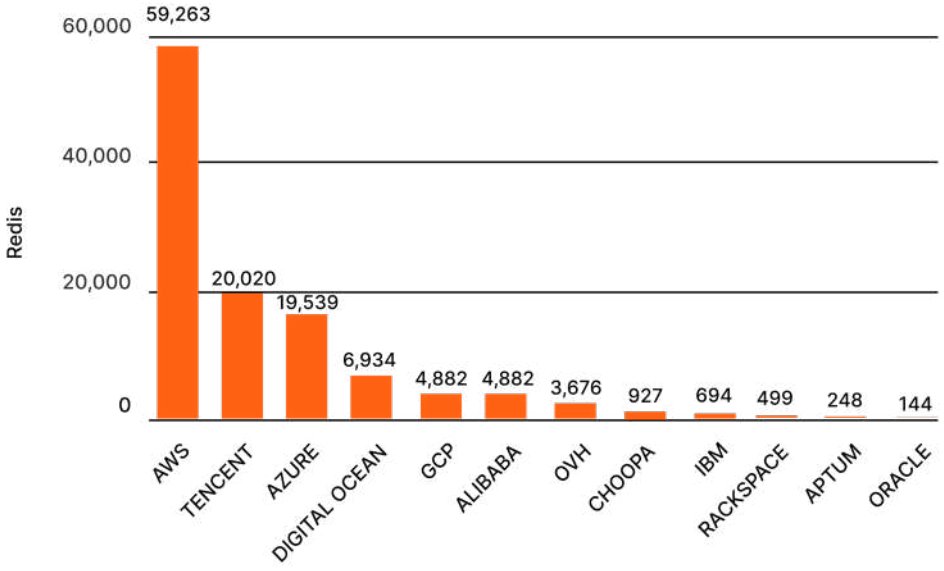

Figure 8, *Service Count of Redis Exposures by Cloud Provider*

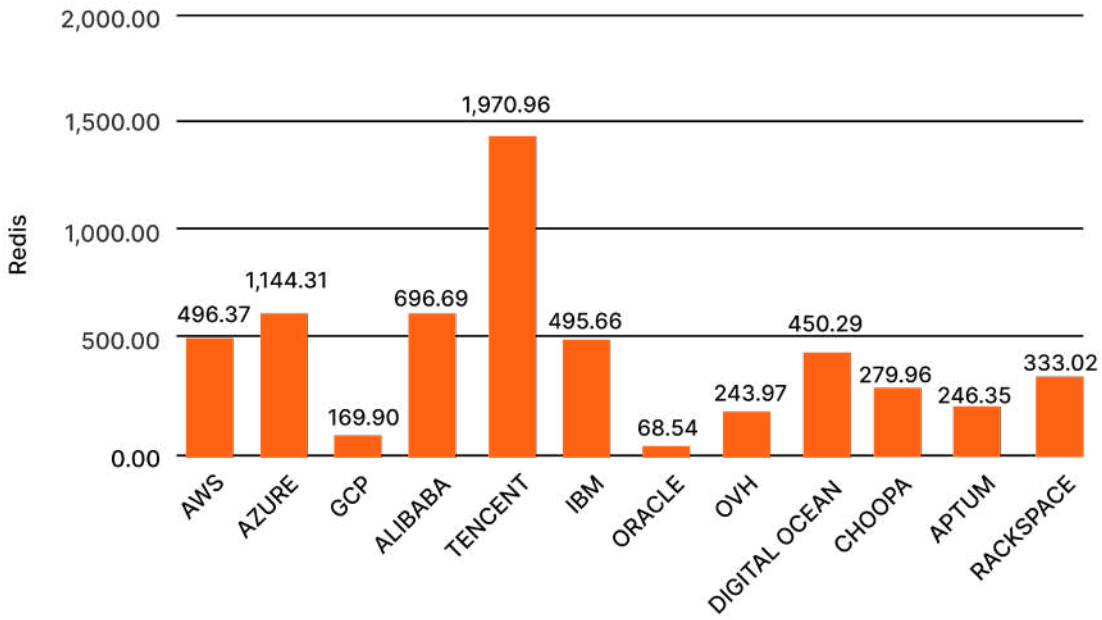

Figure 9, *Prevalence Rate of Redis Exposures per 100,000 hosts by Each Cloud Provider*

The aggregate service count by provider in Figure 8 is useful to understand the total amount of exposure on the cloud. The prevalence rate of Redis exposure per 100,000 hosts to compare across providers can be found in Figure 9. From the comparative analysvis of exposure rates, we see there is clearly an outlier, Tencent.
We reviewed Tencent's **documentation** around Redis configuration and it is unclear if by default a public IP address setting is selected during setup and configuration. Other explanations could be related to usage bias, meaning users of Tencent may be biased to using servers and services that utilize RDP.

Exposed Services Related to Remote Administration

The second area of impact we investigated were exposures related to remote administration. This impact area seemed an obvious one to explore given the increase in remote work across the globe over the last year and increases in cybercrime with the FBI reporting losses of more than $4.2 billion.

Some services listed below such as SSH are not inherently risky if exposed to the Internet. It is the way in which the service is configured, or version running, that could then be exploited and cause harm to the organization. The services we analyzed can also be used in attacks like: credential stuffing and vulnerable versions have led to successful malware and ransomware attacks over the years. More details around the top 3 services identified in this impact area and consequences of exposures can be found further in this report.

- SSH
- RDP
- VNC
- SMB
- TELNET
- Team Viewer
- PC Anywhere

From the list of services outlined, we identified the top three service counts across the dozen providers investigated.

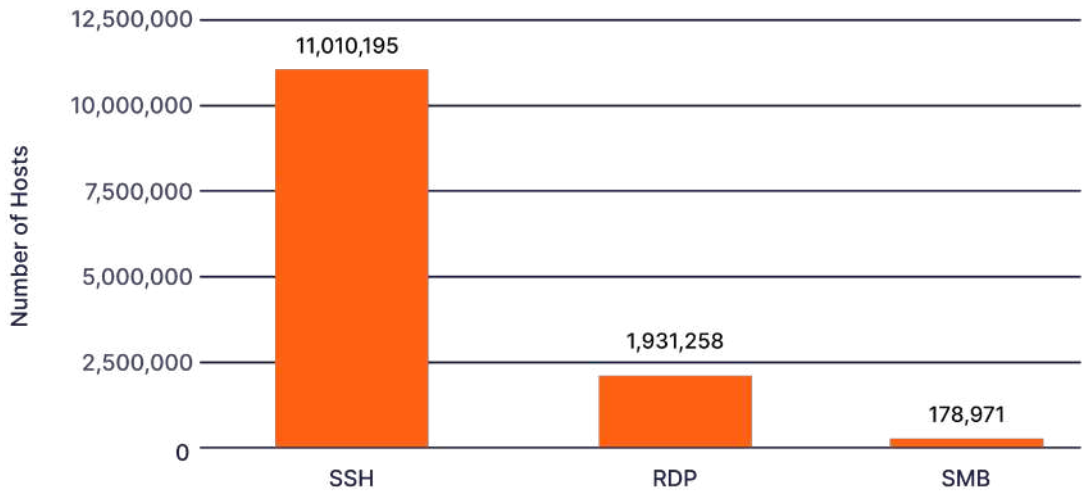

Figure 10, *Raw Count of Top 3 Exposed Services by Impact Area: Remote Administration*

These services include: SSH, RDP and SMB. Over 1.9 million RDP services were observed across the dozen cloud providers in our research. We conduct further analysis below for each of these remote administration services.

SSH

SSH, or Secure Shell, is a remote way to securely connect to a server to modify or access the system. SSH can be used via the terminal on Unix-based systems and by a client on Windows systems. SSH is not inherently a risk to organizations who have it exposed to the Internet. However, we saw so much of the service that we felt it necessary to include in in our analysis. We also know strong authentication practices are not always followed, such as disabling password authentication which is vulnerable to credential stuffing, a brute force attack documented by the ATT&CK framework. SSH is also known to be a way malware communicates with command and control (C2) servers, a protocol tunneling technique also documented by the ATT&CK framework.

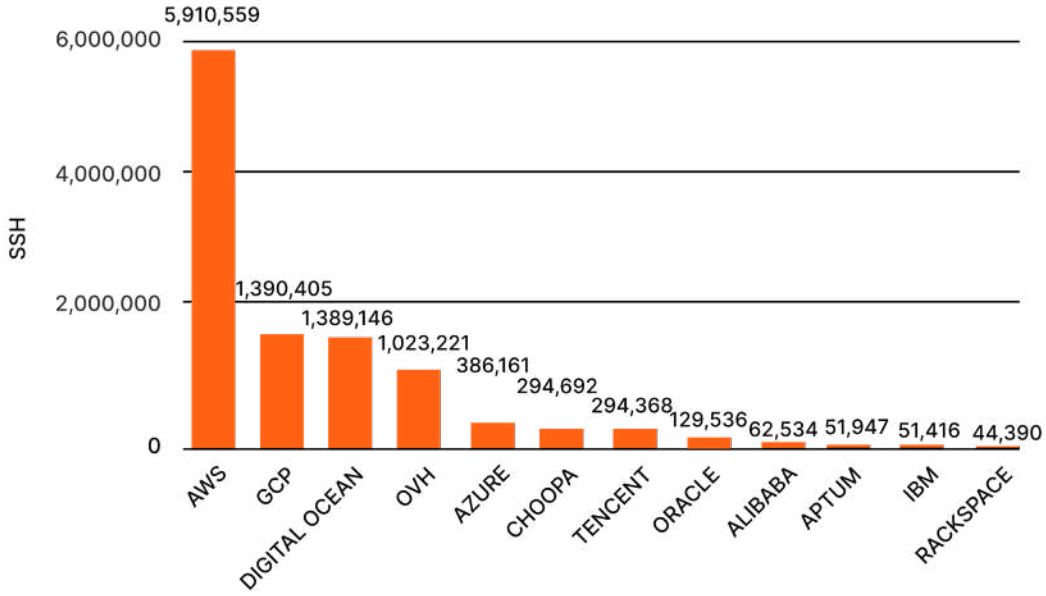

***Figure 11**, Service Count of SSH Exposures by Cloud Provider*

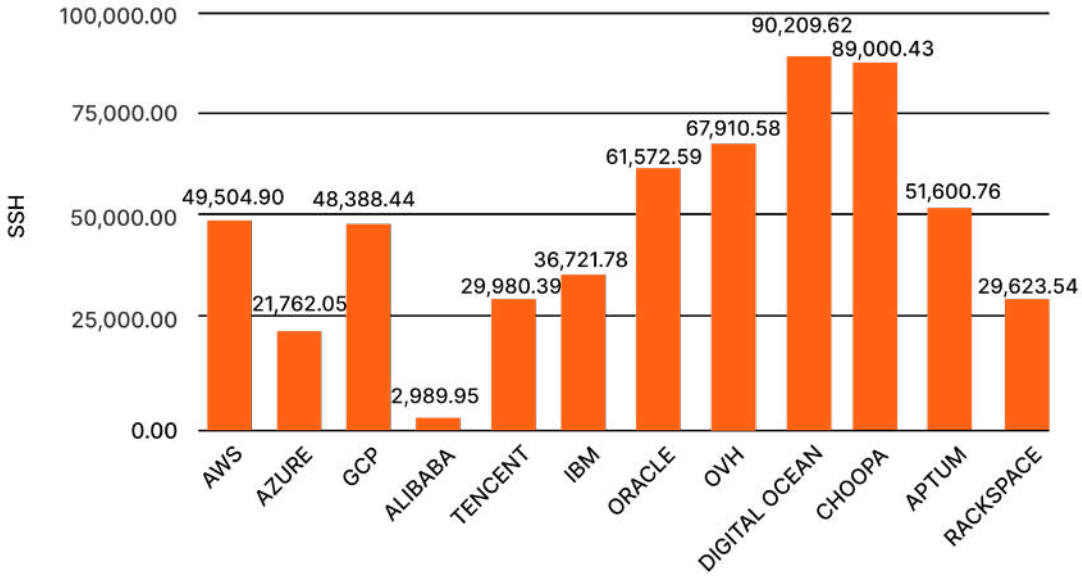

***Figure 12**, Prevalence Rate of SSH Exposures per 100,000 hosts by Each Cloud Provider*

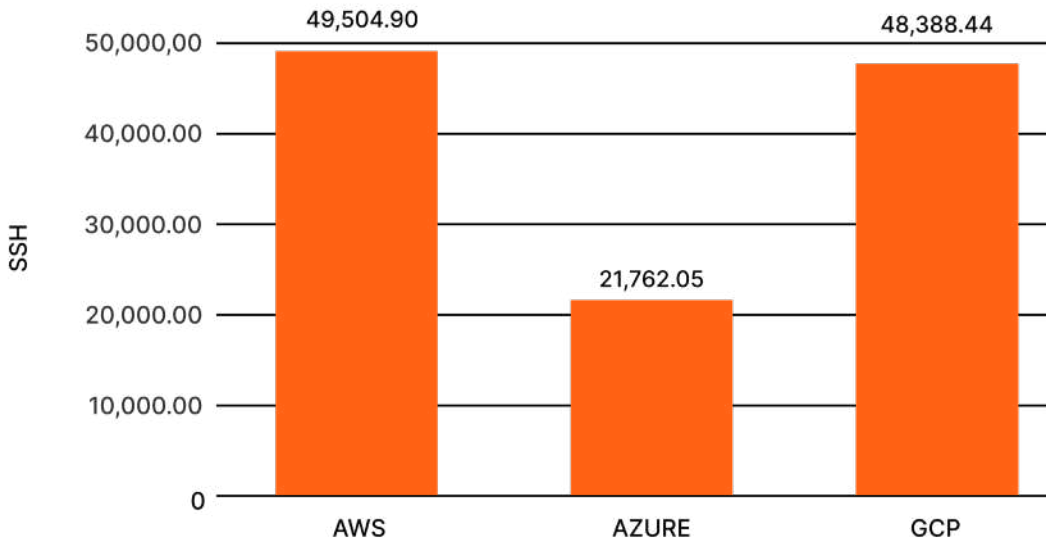

Figure 13, *Prevalence Rate of SSH Exposures per 100,000 hosts for AWS, Azure, and Google Cloud*

Figure 11 and Figure 12 highlight SSH service counts and exposure rates by provider, respectively. Looking only at Amazon AWS, Microsoft Azure and Google reveal a significant difference in exposure rates. AWS and Google have more than double the rate of SSH services than Azure. One reason this could be the case is the way in which customers on these platforms are using the infrastructure. For example, if more Azure customers use Windows systems, perhaps they are less likely to use SSH since a client is required.

We also investigated why Digital Ocean had such a high rate of prevalence relative to other providers. Looking at their **documentation**, SSH was mentioned as the primary way users interact with "**droplets**" or virtual machine instances on their infrastructure, which we know is a common use case for Digital Ocean. They also recommend customers use key-based authentication over password authentication for better security.

RDP

RDP, or Remote Desktop Protocol, is used to access desktops from remote locations. RDP has grown more noteworthy in recent years due to ransomware exploiting vulnerable versions to gain initial access into the network. Often this results in further compromised machines or compromised data. The RDP vulnerability that Microsoft issued **warnings** about in 2019 is called **BlueKeep**. See FireEye's report on **Ransomware Protection and Containment Strategies** to harden RDP to mitigate attacks.

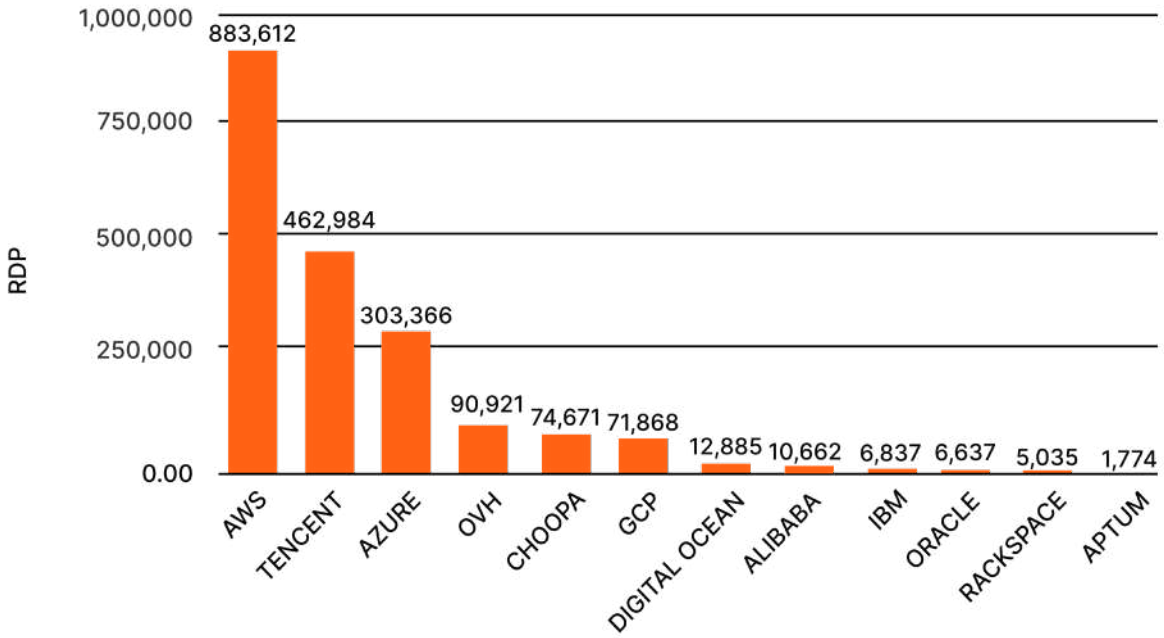

Figure 14, *Service Count of RDP Exposures by Cloud Provider*

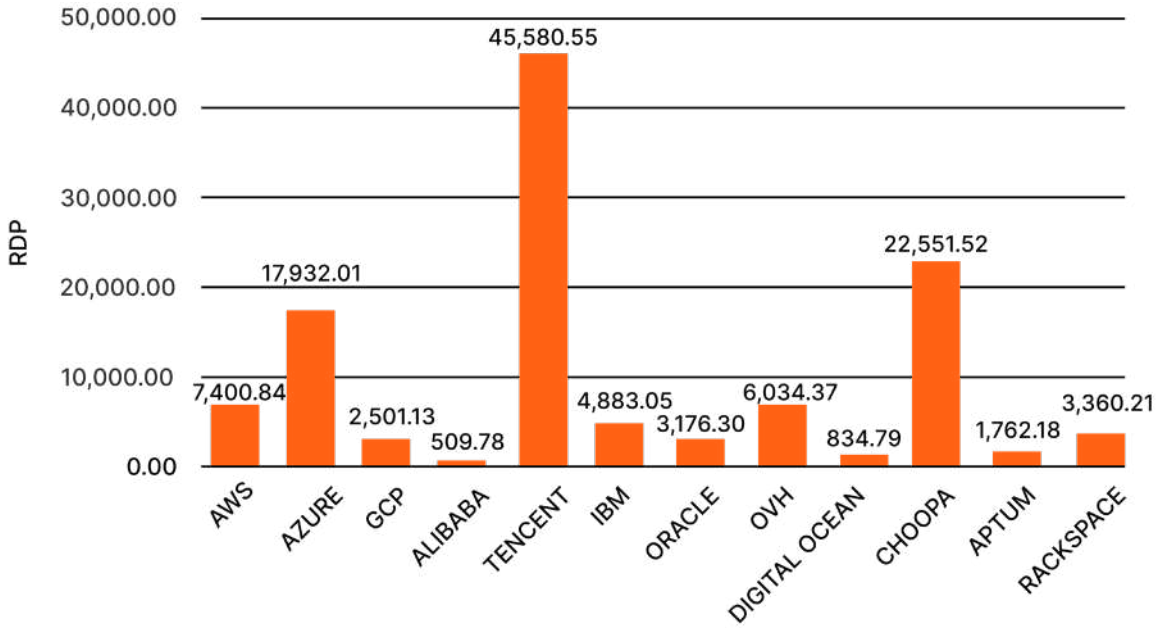

Figure 15, *Prevalence Rate of RDP Exposures per 100,000 hosts by Each Cloud Provider*

Figure 14 highlights RDP service counts by provider. After AWS, Tencent and Azure hold the next two highest counts of RDP services exposed in their infrastructure.

Figure 15 shows the prevalence rate of RDP per 100,000 hosts. At first glance it is clear that one provider has a significantly greater rate of exposure than any of the other providers we investigated. This provider is Tencent with over 45,000 RDP services per 100,000 hosts. This is two and half times greater than Azure. We reviewed Tencent documentation and found they recommend users connect to Windows instances via RDP which could explain the higher rate. Further explanation could also be related to customer usage bias, where perhaps more providers with higher prevalence rates are using more Windows services relative to others.

SMB

SMB, or Server Message Block, is a network protocol with a variety of uses on Windows systems such as file and printer sharing, as well as access to remote Windows services. SMB earned its fame from the notorious WannaCry ransomware that spread across the globe in 2017. The ransomware exploited an SMB vulnerability, EternalBlue, and remained one of the top 10 most exploited vulnerabilities from 2016-2019 according to CISA.

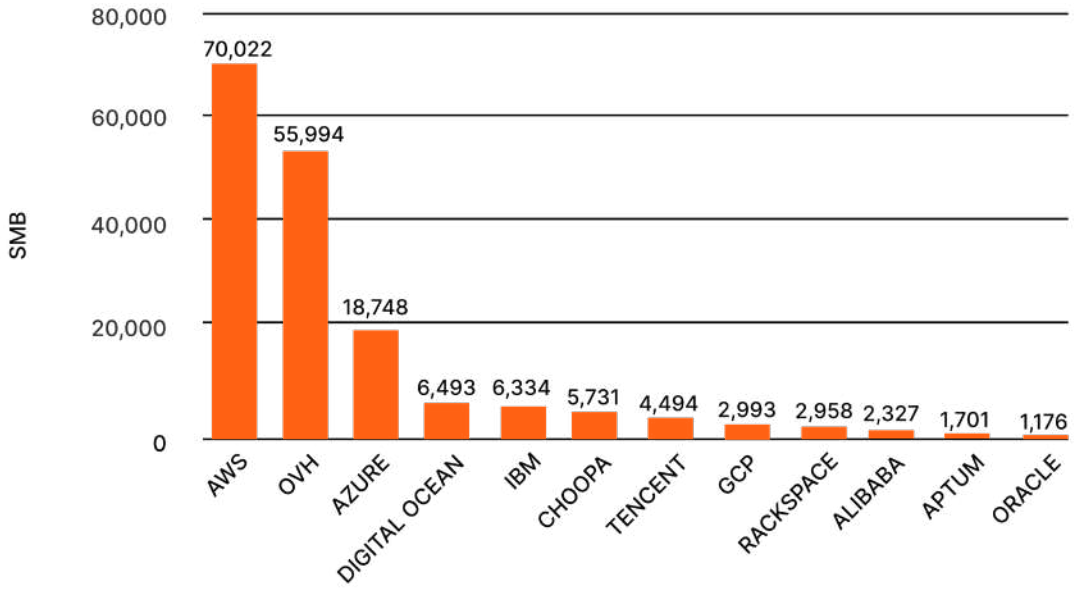

***Figure 16**, Service Count of SMB Exposures by Cloud Provider*

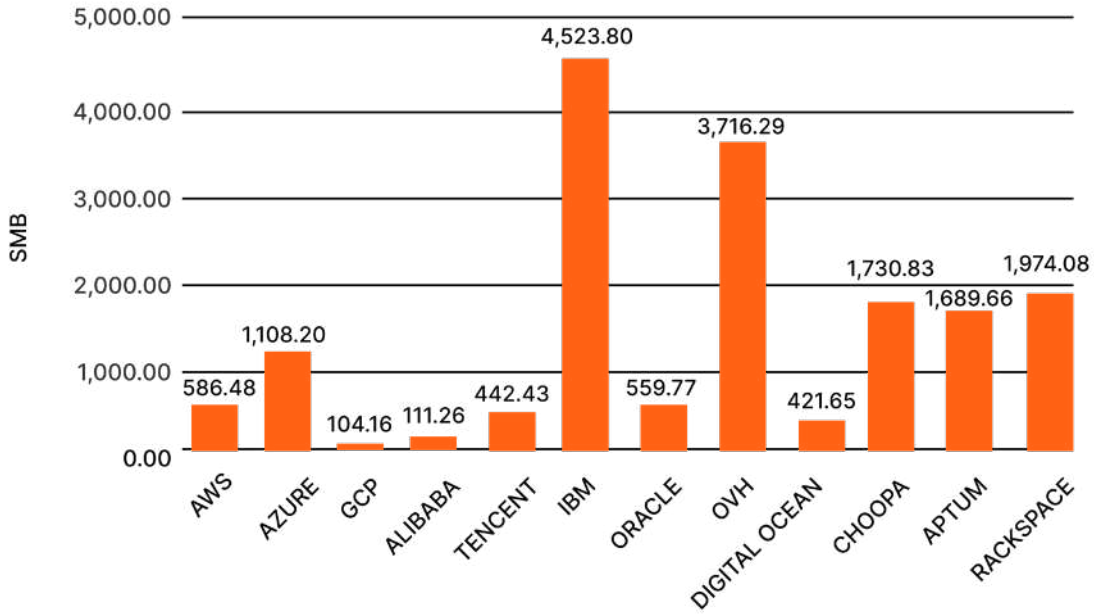

***Figure 17**, Prevalence Rate of SMB Exposures per 100,000 hosts by Each Cloud Provider*

Figure 16 highlights SMB service counts by provider. Unexpectedly, OVH surpasses Azure by a significant amount, encroaching on the AWS lead, despite the fact it has significantly less cloud infrastructure.

Figure 17 shows the prevalence rate of SMB service exposure per 100,000 hosts. Given the sheer number of SMB exposures we saw in OVH by service count, we expected to see a high rate of exposure on OVH, which we do. However, this was not the highest rate of SMB exposure across the dozen providers. In fact, IBM had the highest rate of SMB exposure than any other provider in our research - 4,524 per 100,000 hosts.

It is unclear from our initial research why users of IBM cloud infrastructure have higher rates of SMB exposure than other cloud providers. Interestingly, SMB was originally designed by IBM. We are unsure if this has any correlation, but when searching through IBM documentation, one IBM product that continued to show up was their General Parallel File System called IBM Spectrum Scale. This could point to usage bias, where customers of IBM may be more likely to be using a service that utilizes the SMB protocol than customers on other cloud providers.
As for OVH, we could find no obvious reason in their documentation that would explain why users of OVH cloud infrastructure would be more likely to expose an SMB service relative to users on other cloud providers. A possible explanation could be usage bias in that customers of OVH may be more likely to utilize services that leverage SMB relative to other providers.

Future Work and Potential Explanations

The research and analysis we conducted lays the foundation for future research, particularly around why these differences exist across cloud providers. A few explanations posited throughout the research to account for these variances in service exposure rates include things like:

- **Usage bias**: Customers on certain providers may be more likely to use services that utilize the services discussed in this report.

- **Default configurations**: Default configurations vary across providers and there is likely a difference in service exposure rate based on default configuration setup.

- **Cloud security controls, or maturity of controls, by provider**: Cloud service providers have varying security controls which could also account for exposure rate differences.

Investigating why these differences happen such as evaluating default service configurations by cloud provider or analyzing security tools that enable better security practices among practitioners are both areas of potential future work.

Conclusion

Our assumption at the beginning of this research was that unknown and unmanaged assets are more likely to have poorer security hygiene and therefore, more misconfigured services on the cloud. Our findings, however, suggest service exposure is more chaotic and prevalent across the board. With more and more multi cloud digital transformations projected in the near future, complete visibility of your entire attack surface is more important than ever.

To learn more about Censys Attack Surface Management Platform, please visit www.censys.io, or reach out at hello@censys.io.

To request a personalized demo, visit **https://censys.io/demo-request/**