



The Cost of Cloud Report 2024

A comprehensive analysis of cloud computing costs and future strategies

CIVO

White Paper | 2024

Has cloud computing lost its way to complexity and cost?

Our latest report, *The Cost of Cloud 2024*, seeks to address the emerging challenges of cloud computing and prompts questions about the industry's current trajectory.

Through a survey of 500+ industry professionals, we have gained an understanding of the current trends, challenges, and strategic responses shaping the future of cloud economics, especially in relation to the dominance of the "Big 3" cloud providers - Azure, Amazon Web Services (AWS), and Google Cloud Platform (GCP).

This report serves not only as a reflection of where the industry stands but also as a call to action for what needs to be done next. The findings underscore a pressing need to pivot towards the core values that once defined cloud computing: transparency, customer-centricity, and a relentless pursuit of value.

We welcome you to join us in reimagining the cloud. Together, a course can be charted toward a cloud computing future that is equitable, sustainable, and true to its original promise.



Mark Boost, CEO & Co-founder of Civo






Introduction	1
The cloud provider landscape	2
Leveraging GPUs in cloud computing	6
Envisioning the future of cloud computing	8
Optimizing and managing cloud costs	10
Conclusion	12
References	13

Introduction

Since our previous cost of cloud report two years ago⁽¹⁾, concerns surrounding the complexity and unpredictability of cloud costs have continued to grow, posing significant challenges for organizations of all sizes.

When first introduced, cloud computing promised to revolutionize how users access, manage, and deploy technology. Its initial allure was undeniable: scalability, flexibility, and cost-efficiency. These benefits propelled the cloud to become a cornerstone of modern IT strategies. However, as the industry matured, it has become increasingly apparent that emerging challenges have clouded this promise. Rising costs, complex billing models, and concerns over data sovereignty and privacy have revealed the critical areas for reflection and recalibration.

Back in 2012, Andy Jassy (current CEO of Amazon) spoke at re: Invent⁽²⁾ about the benefits of the cloud, promising a future of low-margin business.



"AWS is able to take its very large scale and pass on those savings to customers in the form of lower prices."

Andy Jassy (current CEO of Amazon)

Fast forward just over 10 years later, London Technology researchers Dark Matter⁽³⁾ have concluded that cloud costs are spiraling, challenging 82% of organizations.

Market power among the "Big 3" cloud providers, Azure, AWS, and GCP, has led to increased concerns about cost, complexity, and vendor lock-in. After speaking with 500 industry professionals, it is clear that the cloud is broken.

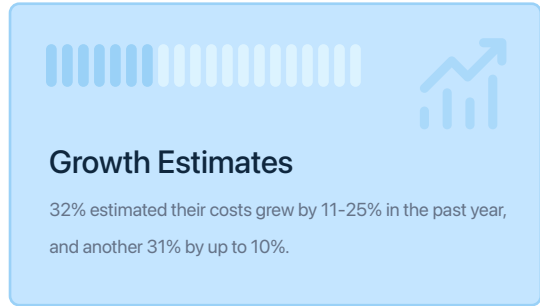
Amid these challenges, alternatives like Civo are emerging, advocating for a shift back to the core principles of cloud computing: transparency, fairness, and simplicity. Organizations such as Civo recognize the importance of fostering an ecosystem where technology empowers developers rather than restricts them. Through fair pricing, simplified architecture, and community engagement, Civo is laying the groundwork for a future where the cloud is accessible to all.

This report not only aims to map the terrain of cloud costs but also to chart a course toward a more equitable and sustainable future in cloud computing.

The cloud provider landscape

According to our research, 77.4% of the 500 industry professionals surveyed are using one of the Big 3 hyperscalers. This dominance underscores their critical role in the industry yet also spotlights the issues surrounding market saturation, cost complexity, and vendor lock-in.

The market’s consolidation around Azure, AWS, and GCP has prompted a discussion on the implications of such dominance:



These figures underscore the challenges organizations have been facing while managing and predicting cloud expenses in the current market landscape. This contradicts Andy Jassy’s earlier promise, as **37% of organizations from our research believe that the cloud has failed to live up to its earlier promise of cost-effectiveness**. Consequently, organizations are grappling with the multifaceted challenge of cloud cost management and the false promises once made.

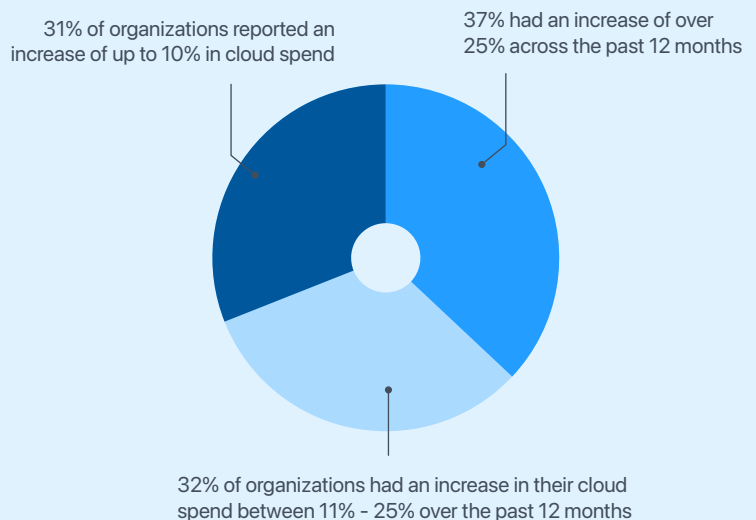
The cost of cloud over last 12 months

Research according to a survey of 500+ industry professionals



Breakdown of cost increase

Based on the 59% of 500+ industry professionals that faced an increase in their cloud spend over the past 12 months



Despite the many benefits of the cloud, a study by IDC⁽⁴⁾ found that 71% of respondents expected to partly or fully migrate public cloud workloads into a dedicated IT environment over the next two years.

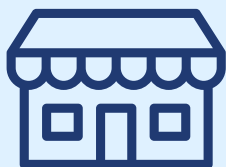
Additionally, when looking at this year's Flexera⁽⁵⁾ study, organizations initially ranked security as their top cloud challenge. However, within a year, 'managing cloud spend' became the primary concern, highlighting the growing issue of hyperscaler dominance that continues to escalate annually.

A recent publication from The Guardian⁽⁶⁾ further supports this by highlighting how the hyperscalers have increased their hosting and storage fees anywhere from 11% to as much as 50% compared to a year ago. Our research uncovers a broad spectrum of cloud spending across organizations, reflecting the diverse adoption and reliance on cloud services:



Small organizations

32% of organizations with less than 50 employees find themselves with monthly cloud bills over \$10,000, with some ranging up to \$5,000,000.



Small to medium-sized businesses (SMBs)

54% of SMBs with between 100-499 employees have a monthly cloud expenditure ranging from \$10k - \$499k.



Large organizations

43.6% of large organizations (500+ employees) incur monthly expenses exceeding \$1,000,000, with 10% exceeding \$10,000,000 each month.

This is further enhanced by the clear challenge in forecasting cloud expenses, as **26% of those surveyed noted that they have had unexpectedly large cloud bills.**

Comparing the pricing for large CPU instances from 2022 to 2024 reveals significant changes and highlights the evolving cost dynamics across major cloud providers.

2022 Pricing

Provider	Instance Type	Compute	Storage	Control Plane	Data Transfer	Total	Civo Saving
CIVO	Civo Kubernetes <i>g4s.kube.large</i>	\$480.00	\$20.00	Free	Free	\$500.00	-
aws	AWS <i>t4g.2xlarge</i>	\$588.66	\$20.00	\$73.00	\$460.80	\$1,142.46	\$642.46
Google GKE	Google GKE <i>e2-standard-8</i>	\$587.01	\$24.00	\$71.60	\$425.00	\$1,107.61	\$607.61
Microsoft (AKS)	Microsoft (AKS) <i>d8 v5</i>	\$840.96	\$38.02	Free	£399.60	\$1,278.58	\$778.58

2024 Pricing

Provider	Instance Type	Compute	Storage	Control Plane	Data Transfer	IP Address	Total	Civo Saving
CIVO	Civo Kubernetes <i>g4s.kube.large</i>	\$480.00	\$20.00	Free	Free & unlimited	Free	\$500.00	-
aws	AWS <i>t4g.2xlarge</i>	\$579.43	\$38.56	\$72.00	\$460.80	\$10.80	\$1,234.59*	\$734.59
Google GKE	Google GKE <i>e2-standard-8</i>	\$750.14	\$40.80	\$72.00	\$378.81	\$8.64	\$1,250.35*	\$750.35
Microsoft (AKS)	Microsoft (AKS) <i>d8 v5</i>	\$1,032.40	\$21.32	Free	£392.00	\$12.96	\$1,458.68*	\$958.68

* Based on an equivalent size 3-node cluster with 200GB Persistent Volume and 5TB data transfer. Prices are per month. Prices are correct as of May 2024.

While base pricing typically covers compute, storage, and control planes, several hidden fees are not always transparently listed, such as IP addresses, data transfer fees, and additional services such as security and monitoring, including data encryption and managed security services. These hidden fees make financial planning and cost management more complex for organizations relying on cloud services. This cost comparison analysis highlights the importance of carefully evaluating and comparing different cloud providers to manage costs effectively and avoid unexpected expenses.

66%

Organizations can save up to 66% by choosing Civo over traditional hyperscalers.



This has led to organizations being motivated by the desire to mitigate costs and avoid vendor lock-in, which has prompted them to seek alternatives to traditional cloud services. Our research indicates a growing interest in alternative cloud providers, with organizations seeking solutions that offer simplicity, cost efficiency, and flexibility to avoid vendor lock-in.



Cost management actions

53% of organizations have taken actions to manage or reduce their cloud service costs, utilizing a mix of software tools (30%), internal strategies (42%), and specialized consultants (15%).



High costs concern

47% believe the cost of cloud services is too high, reflecting a growing interest in alternatives that offer better control and cost efficiency.



Organizations are not only grappling with the challenges posed by the dominance of the Big 3 but are also actively exploring alternatives that align more closely with their needs for transparency, flexibility, and cost efficiency. As the market continues to evolve, the responses from cloud providers and the decisions made by organizations will shape the future trajectory of cloud computing, potentially leading to a more diversified and competitive ecosystem.

Leveraging GPUs in cloud computing

As artificial intelligence (AI) and machine learning (ML) applications surge within the cloud landscape, the demand for powerful computational resources, particularly GPUs, has grown exponentially.

According to the BBC⁽⁷⁾, **data center power use is expected to surge six-fold in the next ten years, with demand on the grid forecasted to double by 2050 as heat, transport, and industry continue to electrify.** Additionally, the International Energy Agency⁽⁸⁾ predicts that AI will consume ten times as much power in 2026 as it did in 2023, with data centers projected to use as much energy as Japan by that year. Nick Harris, Founder and CEO of Lightmatter, highlights that the computing power required for AI doubles every three months, a pace that outstrips even Moore's law, potentially straining companies and economies.

“Every cloud decision has an economic cost and an environmental cost. These two costs are inextricably linked, however they aren’t often discussed in the same room, by the same people.”

A Dark Matter Film: **Clouded II, Does Cloud Cost the Earth**



With the growth in AI/ML applications, the cost associated with GPU instances is increasing. Hyperscalers charge twice as much as alternatives like Civo, further restricting access to critical AI/ML resources. This disparity is evident when comparing the prices of NVIDIA A100 GPU offerings from hyperscalers and Civo:

NVIDIA A100 80GB GPU monthly pricing (compute only)

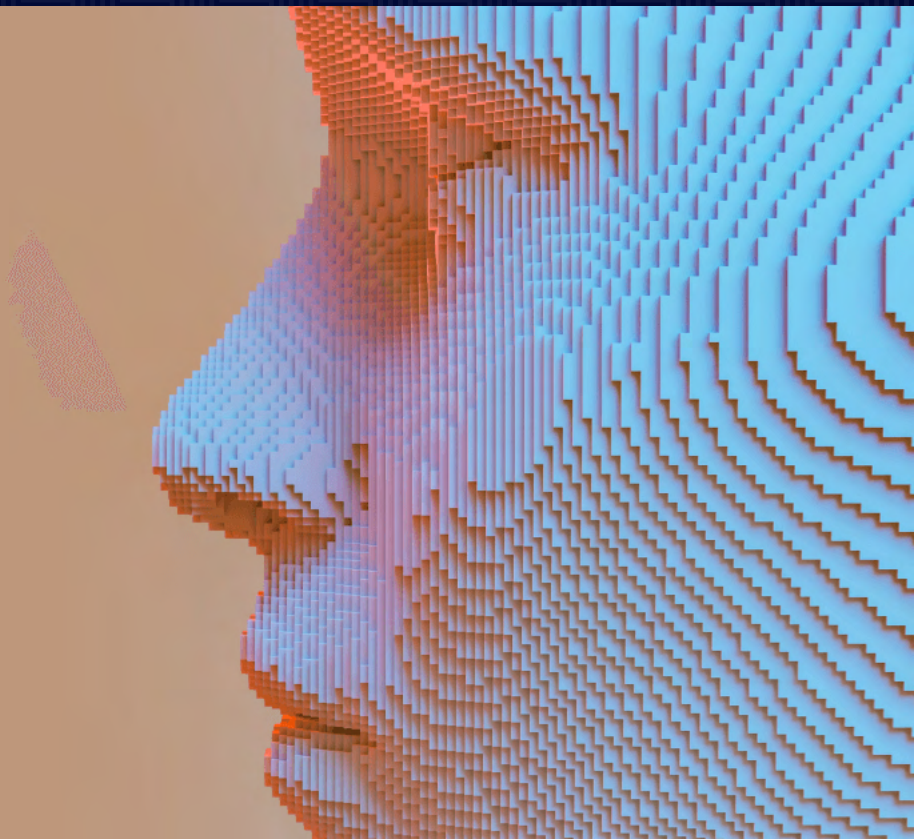
Provider	Instance Type	Total Monthly Cost	Civo Saving
Civo	Civo Kubernetes g4g.xlarge	\$12,800	-
AWS	AWS p4d.24xlarge	\$23,594	£10,794
Google GKE	Google GKE a2-highgpu-8g	\$21,889	£9,089
Microsoft (AKS)	Microsoft (AKS) NDm A100 v4-series	\$30,960	£18,160

Prices are correct as of July 2024.



83%

of respondents reported they either have already deployed or are currently piloting generative AI (GenAI) within their organizations.



A Gartner survey⁽⁹⁾ conducted from October to December 2023 found that 83% of respondents reported they either have already deployed or are currently piloting generative AI (GenAI) within their organizations. As more organizations adopt GenAI and other AI technologies, the demand for powerful and affordable GPU resources has become crucial. The pricing disparities among these providers highlight the financial challenges many organizations face in accessing the GPU power necessary for advanced AI and ML applications. Compared to hyperscalers, alternative cloud providers can offer a competitive pricing model that stands out as a cost-effective alternative, promoting broader access to high-performance computing resources.

The rising costs and added fees associated with traditional hyperscalers highlight the necessity for organizations to meticulously evaluate their cloud service providers. As we look towards the future, the cloud computing landscape continues to evolve. Organizations must navigate these changes strategically to optimize their cloud investments and ensure sustainable growth. This transition sets the stage for envisioning the future of cloud computing, where innovative solutions and equitable practices will play a crucial role.

Notes

Free tiers or offers are excluded from the comparison of NVIDIA A100 80GB GPU monthly pricing.

Envisioning the future of cloud computing

The economics of cloud computing is constantly being scrutinized⁽¹⁰⁾, and based on our cost comparison findings, it is no surprise organizations are beginning to reconsider their reliance on cloud services. As organizations struggle with the rising costs reflected in the 23% average price increase by hyperscalers like AWS in the past year⁽¹¹⁾, it's clear that alternative infrastructures are becoming increasingly attractive.

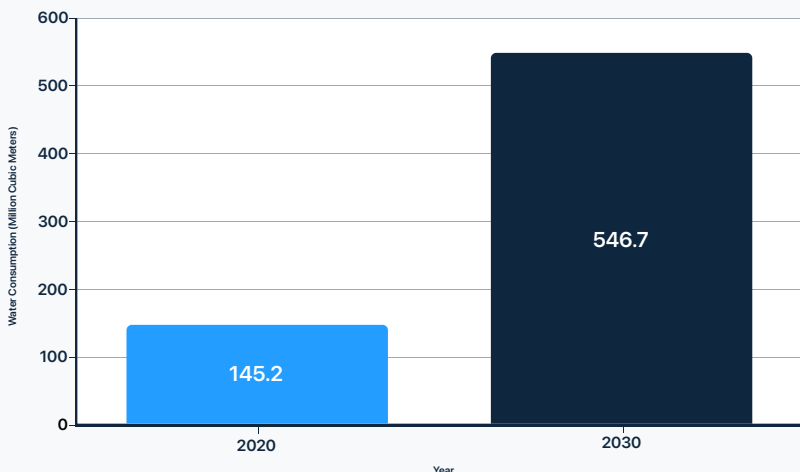
The decision to move beyond traditional cloud services is reflected in our research, revealing a diverse response to cloud migration:

37% of organizations are considering a move to alternative infrastructures, albeit without firm plans, while 7% are actively in the process of moving away from the hyperscalers.

This journey towards alternative infrastructures has its challenges, requiring careful consideration of the trade-offs between cost, control, scalability, and technical capabilities.

Vertice⁽¹²⁾ found in their cloud spend report that cloud costs were among the top cost-cutting priorities for CFOs in 2024. With this in mind, it's crucial for organizations to explore more transparent and cost-effective alternatives to traditional cloud providers.

Projected Increase In European Data Center Water Consumption (2020-2030)



In addition to financial considerations, the environmental impact of cloud computing is a growing concern. Fardan and Lohrmann⁽¹³⁾ outlined how European data center water consumption will increase from 145.2 million cubic meters in 2020 to 546.7 in 2030 - almost a 300% increase in 10 years. This surge underscores the need for sustainable practices in the industry.

Some companies are exploring innovative solutions to address the environmental impact. For instance, [Deep Green](#) is focusing on supporting waste heat utilization from data centers, which is one of the IEA's core recommendations to help combat the increasing carbon emissions. Deep Green's system runs on zero-carbon electricity, and the swimming pools heated by the system can reuse 90% of the energy, meaning each kWh of energy is effectively used twice.

"Energy is expensive, and precious. So many organizations and communities struggle to afford the heat they need to exist. Yet so much energy, so much heat, is wasted everyday by data centers. That's what Deep Green is fixing, and in the process dramatically reducing overall carbon emissions. For too long, compute infrastructure has contributed to the problem of climate change. Now it can be part of the solution."



Mark Bjornsgaard
CEO of Deep Green



The future of cloud computing lies in a dual focus on cost efficiency and environmental sustainability. As organizations navigate the evolving landscape, integrating cost-effective practices with innovative environmental solutions will be key. By prioritizing transparent financial practices and sustainable technologies, the industry can move towards a more balanced and responsible future. This vision not only addresses the immediate financial concerns but also contributes to the broader goal of mitigating climate change.

Optimizing and managing cloud costs

The challenges of managing and optimizing cloud costs are becoming increasingly prominent in the modern landscape. According to AAG⁽¹⁴⁾, managing cloud spend has become the top priority for both enterprises and SMBs (82%).

By leveraging a combination of open source tools, proprietary software tools, internal practices, specialized consulting, and the principles of FinOps, organizations can navigate the complexities of cloud costs more effectively, ensuring their cloud investments deliver the desired value and performance.

Picking a cloud provider that allows users the flexibility to choose the tools best suited to their use case ensures that organizations can effectively manage and optimize their cloud costs. Alternative cloud providers such as Civo allow users to choose from various observability tools without locking them in.

40%

According to our research, 40% of organizations have yet to take any actions to manage or reduce their cloud service costs.

The following table highlights some of the available open-source and proprietary observability tools that can be used with Civo:

Open source tools



Proprietary tools



As organizations begin exploring alternatives, several considerations emerge, such as the initial investment and maintenance costs, scalability and flexibility, and the technical expertise required to manage and maintain change.

Another effective strategy for optimizing and managing cloud costs is through the use of a multi-cloud system. This approach allows users to implement multiple cloud service providers to meet their needs. While some cloud providers may prioritize performance, others emphasize cost-effectiveness, security, or data sovereignty.

By integrating services from multiple cloud providers, users can more effectively optimize performance, ensure cost predictability, and safeguard their data and information. In a Gartner⁽¹⁵⁾ webinar, they highlighted that by 2028, **more than 50% of enterprises would not get value from their multi-cloud implementations if they do not have a strategy in place.**

One way to implement an efficient cost of cloud strategy is to utilize the four key pillars: Observe, Enact, Automate, and Maintain Consistency:

Steps	Actions	Tools/Practices
Observe	<ul style="list-style-type: none"> Utilize observability tools to monitor resource utilization in real time. Conduct regular audits to identify waste and optimization opportunities. Identify and eliminate hidden costs by scaling down unnecessary instances when not needed. 	<ul style="list-style-type: none"> Tools: OpenTelemetry, Grafana, Prometheus Practices: Regular reviews of resource utilization, rightsizing instances, and eliminating unused resources.
Enact	<ul style="list-style-type: none"> Establish rules and thresholds to detect resource usage deviations. Adopt FinOps principles to align financial and operational goals for better cost management. 	<ul style="list-style-type: none"> Practices: Set proactive rules for resource usage and implement the FinOps framework for accountability and cost-efficiency.
Automate	<ul style="list-style-type: none"> Leverage cost management tools for detailed analysis and optimization. Implement autoscaling to adjust resources based on demand, preventing over-provisioning dynamically. 	<ul style="list-style-type: none"> Tools: Kubecost, Opencost, StormForge Practices: Automate routine tasks and use Karpenter for dynamic resource scaling.
Maintain Consistency	<ul style="list-style-type: none"> Use reserved instances and savings plans for substantial savings and better budget planning. Embrace multi-cloud strategies to balance cost, performance, and risk. Regularly rightsize instances to match workload requirements and avoid wasted resources. 	<ul style="list-style-type: none"> Practices: Take advantage of long-term commitment discounts, develop a clear multi-cloud strategy, and use tools for optimal instance sizing.

As the cloud computing landscape evolves, organizations must adapt and make informed decisions about their infrastructure commitments. This agility will be crucial in achieving both operational and financial objectives.

Conclusion

The journey ahead for the cloud computing industry is both challenging and full of potential. By embracing the insights and strategies outlined in this report, the industry can chart a path towards a future that not only leverages the full power of cloud technology but does so in a manner that is economically viable, socially and environmentally responsible, and universally accessible.

This calls not just for a roadmap for navigating the immediate challenges of cloud computing, but also a manifesto for reimagining the future of the cloud as a catalyst for innovation, growth, and positive change. The time is now for the industry to come together, embracing these principles, to ensure that cloud computing continues to be a force for good in the digital age, empowering organizations and individuals to achieve their highest potential.

As we move forward, it is imperative to prioritize transparency, fairness, and simplicity, steering away from the complexities and high costs that have marred the original promise of cloud computing. The findings of our report advocate for a redefined approach, one that champions customer-centric values and sustainable practices.

We hope you can join us on this journey of reimagining cloud computing. Together, we can reshape how we think about and engage with cloud technology.



Mark Boost, CEO & Co-founder of Civo

References

- ⁽¹⁾ Civo 2022 Cost of Cloud Report: <https://www.civo.com/the-cost-of-cloud>
- ⁽²⁾ re:Invent 2012 Keynote: <https://www.youtube.com/watch?v=8FJ5DBLSFe4>
- ⁽³⁾ Clouded II The True Cost of Cloud: <https://www.darkmatter.co.uk/featured>
- ⁽⁴⁾ IDC 2022 White Paper: https://www.supermicro.com/white_paper/IDC_On-Prem_Cloud_Success_Stories.pdf
- ⁽⁵⁾ Flexera 2023 State of Cloud Report: <https://info.flexera.com/CM-REPORT-State-of-the-Cloud>
- ⁽⁶⁾ The Guardian News Report: <https://www.theguardian.com/business/2023/nov/05/cloud-service-provider-consumer-prices-netflix-microsoft>
- ⁽⁷⁾ Data center power use 'to surge six-fold in 10 years': <https://www.bbc.co.uk/news/technology-68664182>
- ⁽⁸⁾ AI Needs Enormous Computing Power. Could Light-Based Chips Help?:
<https://www.quantamagazine.org/ai-needs-enormous-computing-power-could-light-based-chips-help-20240520>
- ⁽⁹⁾ Gartner Predicts More Than 30% of the Increase in Demand for APIs will Come From AI and Tools Using Large Language Models by 2026:
<https://www.gartner.com/en/newsroom/press-releases/2024-03-20-gartner-predicts-more-than-30-percent-of-the-increase-in-demand-for-apis-will-come-from-ai-and-tools-using-llms-by-2026>
- ⁽¹⁰⁾ Ofcom cloud report: What interventions could the CMA take against AWS and Microsoft?
<https://www.computerweekly.com/news/366554712/Ofcom-cloud-report-What-interventions-could-the-CMA-take-against-AWS-and-Microsoft>
- ⁽¹¹⁾ Liftr Insights News Release: <https://liftrinsights.com/news-releases/aws-and-azure-cloud-pricing-moving-in-different-directions-as-shown-by-liftr-insights-data>
- ⁽¹²⁾ Vertice Cloud Spend Report 2024: <https://www.vertice.one/inside-saas/reports/cloud-spend-report-2024>
- ⁽¹³⁾ Gone with the clouds: Estimating the electricity and water footprint of digital data services in Europe
<https://www.sciencedirect.com/science/article/pii/S019689042300571X>
- ⁽¹⁴⁾ AAG Cloud Computing Statistics: <https://aag-it.com/the-latest-cloud-computing-statistics>
- ⁽¹⁵⁾ Gartner The Future of Cloud 2028: <https://www.gartner.com/en/webinar/566448/1276680>



White Paper | 2024