The on-demand tech paradox

Balancing speed and spend

Capgemini
RESEARCH INSTITUTE

#GetTheFutureYouWant

Table of contents
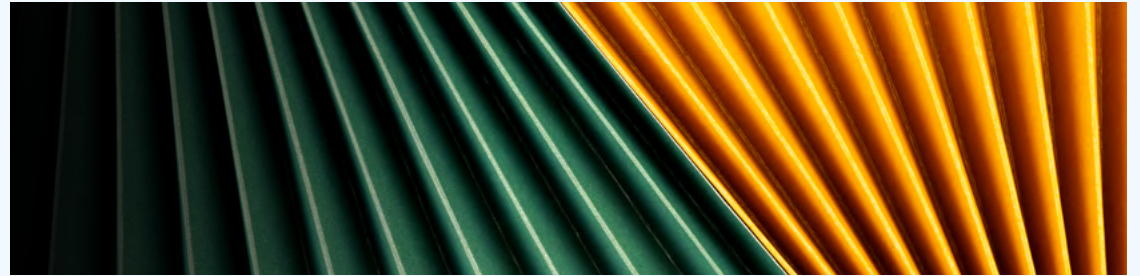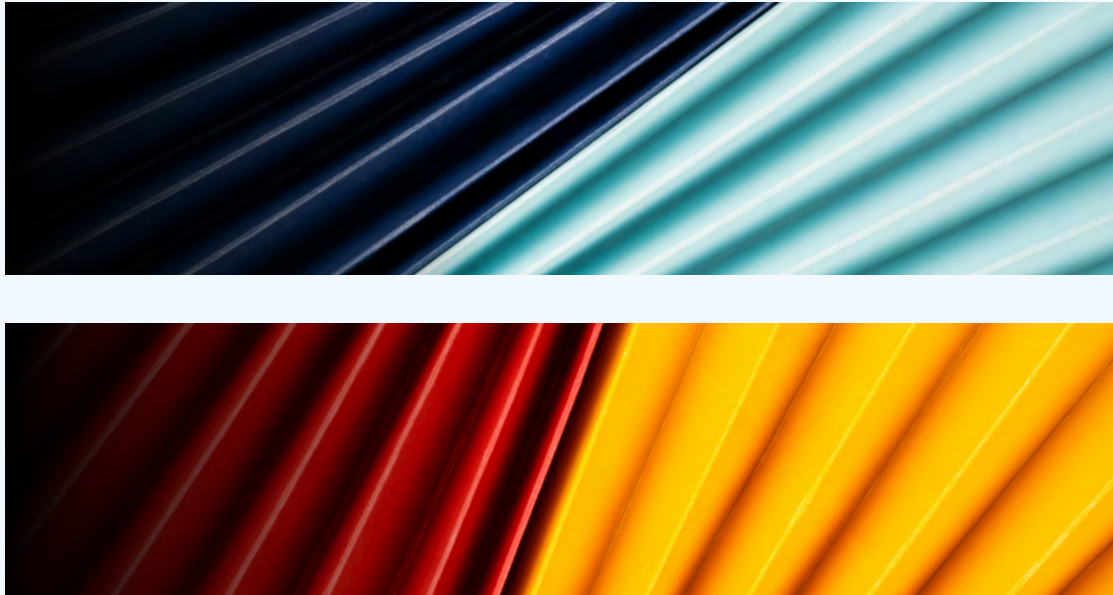
Who should read this report and why?

The economics of On-Demand technologies focuses on controlling costs and maximizing value from cloud services, software-as-a-service (SaaS), and generative AI (Gen AI). Transparency and control over operational expenditure (OpEx) are critical. The evolution from cloud FinOps to On-Demand consumption FinOps demands a holistic approach to cost management that will drive efficiency, reduce total cost of ownership (TCO) and carbon emissions, and unlock greater value from every digital investment and operational decision.

How can we optimize cloud spend without compromising AI performance or innovation? And what is the true cost of scaling AI workloads in the cloud? And how do we measure ROI for increasing Gen AI adoption and siloed SaaS spending across functions? On-Demand tech is here to stay, so how can organizations optimize implementation?

This report will attempt to answer these questions. It will be highly useful to stakeholders across technology, finance, business, and FinOps functions. CFOs, COOs, CIOs, CTOs, business unit heads, and FinOps heads will find insights that support strategic decision-making, financial agility, cloud cost optimization and value enablement.

The report draws on comprehensive analysis of the results of a survey of 1,000 leadership executives (CXOs, vice presidents, and directors) at organizations with annual revenue above $1 billion in 14 countries: Australia, Brazil, Canada, France, Germany, India, Italy, Japan, Netherlands, Singapore, Spain, Sweden, the UK, and the US. The survey spans 12 key industries and sectors: aerospace and defense, automotive, life sciences, manufacturing, consumer products, retail, energy and utilities, telecom, high tech, banking and capital markets, insurance, and the public sector. The report also includes qualitative findings from 10 industry leaders.

Executive summary

On-Demand tech is gaining traction and unlocking benefits

As digital transformation accelerates, organizations are increasingly adopting On-Demand technologies such as cloud, software-as-a-service (SaaS), and Gen AI. These serve to scale innovation, improve agility, and support competitiveness.

Our global survey of 1,000 organizations, each with over $1 billion in annual revenue, reveals that 77% of executives view cloud scalability and performance as critical to business growth and differentiation. IT/tech spending is projected to rise from 4.3% to 5.9% of revenue over the next year, with a notable pivot from maintenance (run) to innovation (build). In this period, the share of On-Demand tech in IT budgets is expected to grow from 29% to 41%. Advanced organizations, whose IT environment is all or predominantly in cloud, are making cost savings, accelerating product innovation, amplifying operational productivity, and improving quality of service.

But there are challenges

Costs and complexity are rising

Despite these benefits, 82% of executives report significant increases in cloud, SaaS, and Gen AI costs. More than six in 10 (61%) say this is a drag on profitability. Inflation, AI/Gen AI/agentic AI adoption, and demand for digital infrastructure emerge as cost drivers. Complexity in pricing and limited visibility exacerbate the issue. We see 55% of organizations planning to relocate workloads across public clouds and 45% considering moving workloads to private clouds, driven by cost, compliance, and sovereignty concerns.

Geopolitical tensions, evolving regulations, and concerns over data control, have prompted nearly half (46%) of organizations to embed cloud sovereignty into their cloud strategies. While this shift often raises operational costs, organizations see it as essential to managing regulatory risk, avoiding penalties, and ensuring long-term resilience. Notably, 42% are definitely and 37% tentatively willing to pay an average 11% premium for sovereign cloud.

Executive summary

Organizations are overrunning their budgets

Cloud and other On-Demand technologies such as SaaS and Gen AI are driving innovation, but they're also pushing up costs. As organizations scale usage, many face cost visibility, governance, and resource-optimization issues. Three-quarters (76%) exceeded public cloud budgets (10% average overrun); 68% overspent on Gen AI; and 52% on SaaS (11% average overrun) in the past 12 months due to underutilized resources, vendor pricing, and decentralized procurement.

"Unmanaged" IT is creating inefficiencies, security risks, and inflated costs

Business units now drive 59% of Gen AI and 48% of SaaS spending, with 12% of the latter unmanaged/unsanctioned. This decentralization, driven by the need for speed, flexibility, and control, leads to duplicative purchases, budget overruns, and security risks. Nearly all (98%) business leaders admit to bypassing IT for tech purchases (8% frequently, 58% occasionally, and 31% on rare occasions).

As a result of this On-Demand tech sprawl, organizations are facing visibility, transparency, and predictability issues:

- 64% say they are unable to accurately forecast cloud budgets

- 59% say cloud waste is a big challenge

- 58% say their organization's On-Demand tech costs are "a big black hole"

- 56% say they face bill shocks due to unpredictable spikes in cloud usage

Investments in On-Demand tech are falling short

Despite significant investments in cloud, SaaS, and Gen AI technologies, only a minority of organizations are achieving anticipated gains:

- 29% say they "fully or mostly" achieved the expected cost savings from SaaS

- 33% say they achieved the expected quality of service-related outcomes from public cloud investments

- 38% say they achieved the expected faster product innovation with Gen AI

Executive summary

Key barriers include poor cost management (71% cite this), underutilized or overprovisioned resources (68%), and lack of standardized ROI metrics (58%).

There are gaps in cost management and FinOps maturity

Without integrated governance, automation, and cross-functional collaboration, gaps in cost management and FinOps maturity hinder the ability to control rising On-Demand tech costs and realize full value from investments.

Cost considerations come too late

The savvy CIO approach is that if security is a Day 0 job, FinOps is a Day 0.5 job.[1] However, more than half (54%) of executives say they adopt cloud-first strategies without cost planning. This is more pronounced in the public (65%), aerospace (63%), and manufacturing (61%) sectors.

FinOps is growing, but remains limited in reach and impact

More than three-quarters (76%) say they either have a dedicated FinOps team already (29%) or will build one in the next 12 months (47%). But within FinOps, 51% focus only on cloud and 38% include SaaS; just 2% cover all of cloud, SaaS, and Gen AI. Moreover, most FinOps teams (63%) focus on operational tasks, not strategic initiatives. Only 42% say FinOps influences business decision-making.

Cost management tools are underutilized

While 60% use cloud cost management tools, only 37% evaluate their effectiveness or act on insights.

Sustainable FinOps is at an early stage

Although 53% agree that sub-optimal On-Demand tech usage leads to excessive energy consumption and increased carbon emissions, just 27% say they measure the environmental impact of cloud. Only 28% have dashboards showcasing cost and carbon for cloud and other On-Demand technologies; and only 36% have a strategy for integrating sustainability into FinOps.

Executive summary

Recommendations for optimizing costs and elevating On-Demand tech value

To fully realize the value of On-Demand tech, organizations must go beyond adoption and focus on optimization. Rising costs, budget overruns, and underwhelming returns highlight the need for a more strategic approach that harnesses smarter architecture, empowered FinOps, integrated governance, and AI-driven automation.

Stategy

- Develop a "cloud-smart" strategy aligned with cloud economics
- Formulate ecosystem partnership models aligned with business outcomes
- Equip finance, business, and tech leaders to arrive at a shared "language of value"

Architecture

- Design to cost/value
- Engineer scalable architecture for efficiency
- Use modular architecture
- Build fit-for-purpose architecture
- Adopt frugal AI architecture
- Implement cost-aware architecture that limits egress charges

Executive summary

FinOps governance and culture

- Start small, showcase quick wins, and evolve cloud FinOps into a strategic capability
- Expand the scope of FinOps to include SaaS and AI/Gen AI
- Foster a culture of shared accountability
- Bridge the skills gap

FinOps processes

- Delete idle resources through a robust tagging process
- Right-size overprovisioned instances and choose the appropriate storage type
- Set usage limits with role-based access
- Schedule resources to deactivate when not required
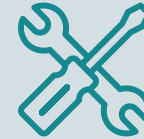
Tools

- Use tools and automation to optimize costs
- Harness the power of AI/Gen AI for FinOps

Sustainable FinOps

- Merge FinOps and GreenOps principles to reduce energy consumption, lower carbon emissions, and drive long-term efficiency while enhancing sustainability credentials

We'd also like to thank the many industry executives who shared their valuable insights with us.

Gerhard Schauer
Vice President, Global IT
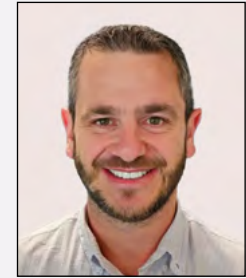Workplace Services at ZF Group

Tonino Greco
Head of Cloud, Infrastructure,
and Operations at River Island

Anna Kopp
Digital Lead Germany
at Microsoft

Pathik Sharma
Cloud FinOps Cost Optimization
Lead at Google

Cléber Alexandre Agazzi
Head of Infrastructure and IT
Operations at Sicredi

Definitions

Our definition of **On-Demand tech** spans public cloud, SaaS, and Gen AI on cloud. On-Demand tech enables businesses to scale rapidly and securely.

Public cloud: A public cloud is a cloud computing model where third-party providers deliver computing resources such as servers, storage, networking, and platforms for applications deployment over the internet to multiple customers on a shared infrastructure. Examples include Amazon Web Services (AWS), Microsoft Azure, and Google Cloud Platform (GCP).

Software-as-a-service (SaaS): SaaS is a cloud-based software delivery model where applications are hosted by a third-party service provider and are accessed by users through a web browser, application programming interface (API), or dedicated desktop client. Examples include Salesforce, Atlassian, Slack, and Google Workplace.

Gen AI on cloud: While Gen AI models can be built and trained on on-premises infrastructure, many organizations tap into cloud's wealth of computational resources, datasets, and tools to accelerate their Gen AI's modeling, training, and fine-tuning, etc. Examples include GitHub Copilot on cloud integrated development environments (IDEs), Microsoft Copilot Studio, and OpenAI ChatGPT on Azure OpenAI Service.

01 | On-Demand tech spending is on the rise – and driving impact

Usage of On-Demand tech is accelerating

As digital disruption accelerates, the ability to adapt and scale swiftly has become a defining advantage. On-Demand technologies such as cloud, SaaS and Gen AI are foundational to scaling innovation, cutting time-to-market and staying competitive.

Gartner estimates that, by 2028, cloud computing will be a business necessity.[2] Our survey of 1,000 global organizations, each with annual revenue over $1 billion, also reveals that three-quarters (77%) of executives agree that the scalability and performance of cloud services is central to business growth and competitive differentiation.

Organizations are pivoting from capital-intensive IT investments to flexible, consumption-based models and On-Demand technologies. Gartner estimates worldwide end-user spending on public cloud services to increase from $595 billion in 2024 to around $723 billion in 2025.

SaaS continues to be the largest spend segment, projected to reach $300 billion in 2025.[3] Additionally, Flexera's 2025 State of the Cloud report highlights that 40% of organizational cloud customers spend over $12 million annually on public cloud, while another 32% spend $2.4 million–$12 million.[4]

Our research also shows:

- In the coming 12 months, organizations are ramping up their overall IT/tech spend from 4.3% of their annual revenue on average currently to an expected 5.9%.

- Most (66%) IT/tech spend currently goes to maintaining existing systems (run), and 34% is allocated to new technologies (build). In the coming 12 months, this proportion is expected to shift to 62% and 38%, respectively, indicating a growing appetite for innovation.

- As a result of this strategic pivot, executives in our research expect the share of On-Demand tech in the overall IT/Tech spend to increase from 29% currently to 41% in the next 12 months (see **Figure 1**).

These figures align with long-term forecasts by analysts. Gartner expects public cloud spending to nearly double in the next four years – from $723 billion in 2025[5] to $1.47 trillion by 2029.[6] Goldman Sachs forecasts Gen AI to account for 10–15% of cloud spending ($200 billion to $300 billion) by the end of the decade.[7]

77%

of executives agree that the scalability and performance of cloud services is central to business growth and competitive differentiation.

Figure 1.

On-Demand tech will claim a larger share of IT/Tech budgets

Current and expected share of On-Demand technologies as a percentage of overall IT/tech spend (average)

■ Current spend as a % of IT/tech spend   ■ Expected spend as a % of IT/tech budget (in the next 12 months)
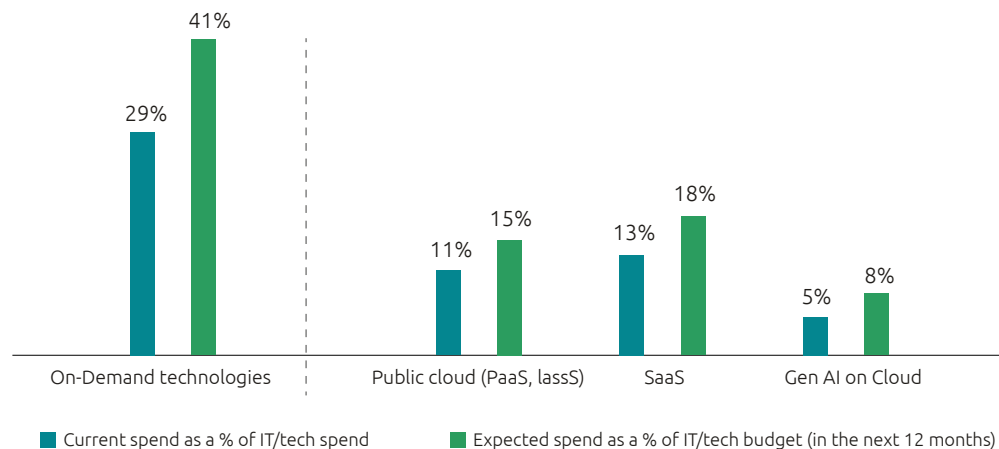
Note: IT/tech spend/budget includes costs for run, build, and maintenance technology, covering hardware, software, cloud, cybersecurity, infrastructure, and IT staff to support business operations and innovation.

Source: Capgemini Research Institute, Cloud and On-Demand tech economics, May 2025, N = 750 executives from technology and finance functions; N = 720 executives from technology and finance functions aware about or involved in public cloud-related activities/spend in their organization; N = 750 executives from technology and finance functions aware about or involved in Gen AI on cloud-related activities/spend in their organizations; N = 749 executives from technology and finance functions aware about or involved in SaaS-related activities/spend in their organizations.

Spending on On-Demand tech has consistently increased and is expected to continue rising.

Public cloud: For the past two years, 77% of executives from technology and finance functions reported an average 8% increase in public cloud spending, with 29% reporting a rise of over 10%. Additionally, 30% of fully public cloud organizations saw a 15–20% spending increase. This indicates growing demand for cloud-native technologies such as AI/machine learning (ML), big data analytics, or container orchestration.[8] More than eight out of 10 (81%) executives from technology and finance functions expect public cloud spending to increase by 11% on average in the next two years.

SaaS: SaaS adoption offers a path to modernization without heavy upfront upfront capital expenditure (capex). Nearly three-quarters (72%) of executives from business functions say their SaaS spend has increased by 6% on average in the past two years. Two-thirds of business executives expect SaaS spending to increase, by 8% on average, in the next two years.

Gen AI on cloud: A significant 63% of executives from technology and finance functions report an average 6% increase in Gen AI spending over the past two years. Looking ahead, 88% of technology and finance executives anticipate an average rise of 8% over the next two years.

On-Demand tech has driven a range of positive outcomes

As **Figure 2** shows, organizations whose IT environment is either predominantly or entirely in cloud see greater cost savings, accelerated product innovation, increased operational productivity, and improved quality of service. Irrespective of SaaS usage maturity, most say their organization has seen greater scalability, improved customer satisfaction, and enhanced sustainability from their SaaS investments.

Figure 2.

Organizations with advanced cloud usage are experiencing positive outcomes

Percentage of organizations seeing positive impact from their public cloud (platform-as-a-service, infrastructure-as-a-service) investments in the past year
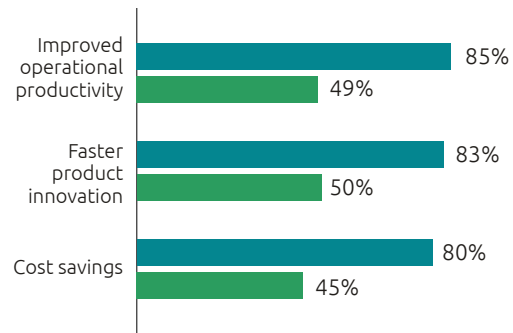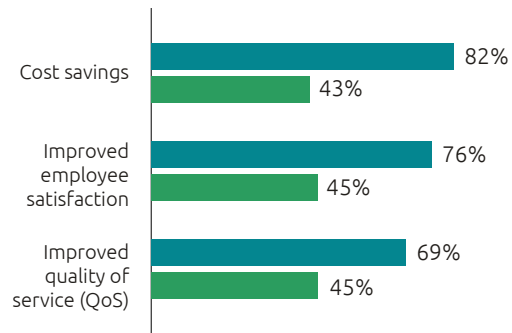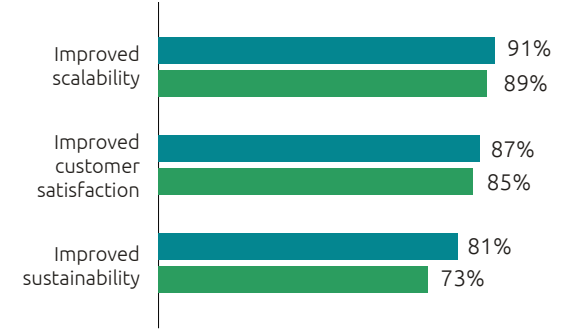
Improved operational productivity
- 85%
- 49%

Faster product innovation
- 83%
- 50%

Cost savings
- 80%
- 45%

■ Organizations with all cloud/mostly cloud in their IT environments
■ Organizations with some cloud and mostly in-house/on-premises IT

Percentage of organizations seeing positive impact from their Gen AI on cloud investments in the past year

Cost savings
- 82%
- 43%

Improved employee satisfaction
- 76%
- 45%

Improved quality of service (QoS)
- 69%
- 45%

■ Organizations with all cloud/mostly cloud in their IT environments
■ Organizations with some cloud and mostly in-house/on-premises IT

Percentage of organizations that have achieved positive impact from their SaaS investments in the past year

Improved scalability
- 91%
- 89%

Improved customer satisfaction
- 87%
- 85%

Improved sustainability
- 81%
- 73%

■ Organizations with SaaS usage across multiple areas
■ Organizations with SaaS usage in a few key areas

Note: The chart highlights the percentage of executives who say their organization has seen >5% change in the above parameters in the past 12 months.

Source: Capgemini Research Institute, Cloud and On-Demand tech economics, May 2025, N = 1,000 executives; N = 615 executives from organizations with all cloud/mostly cloud in their IT environments; N = 385 executives from organizations with some cloud and mostly in-house/on-premises IT; N = 823 executives from organizations using SaaS extensively across multiple areas; N = 177 executives from organizations using SaaS across a few key areas.

We see many examples emerging:

Cost savings: German telecom organization **Deutsche Telekom** uses a platform-based approach that harnesses cloud-native public cloud to ramp up its IT productivity. CEO Laurent Donnay says: *"The real benefits of the cloud are a 30% reduction in infrastructure costs and a drastic improvement in resilience, stability, and time to market … Also, a lot of the native AI features you only get when you run on hyperscalers."* [9]

Scalability: New York City's largest academic medical system, **Mount Sinai Health System**, uses Microsoft Azure Large Instances, a solution designed to run large-scale databases such as the Epic electronic health record (EHR). The system can provide up to 50 million database accesses per second. [10]

Innovation: US-based financial services organization **BNY's** 80% developer community uses GitHub Copilot to increase the speed of code development. A virtual assistant, Eliza, supports employee innovation and workflow management. [11]

Customer experience (CX): Spanish banking group **BBVA** migrated its customer service function to cloud with the intention of offering a more personalized CX. Since its 2019 migration, it has reduced customer waiting times by 42% and shortened response times by 45% in its Peru operation. The bank's Spanish operation can now resolve urgent customer requests in around an hour on average, when it used to take as long as a day. [12]

Productivity: A FinTech organization implemented an error pattern detection AI agent that highlighted specific problematic code blocks causing a spike in payment-processing errors. In three months, this cut debugging time from 12 hours to under two per incident, almost halving overall error rates. [13]

Sustainability: Japanese telecom organization **NTT** has a SaaS system that helps the manufacturing and transportation sectors reduce their carbon footprints. NTT's software combines various technologies, including internet of things (IoT), private 5G, edge compute, digital twins, and its own NTT Smart Solutions platform, to create a customized product that tracks net-zero goals and supply chain footprint. [14]

"The real benefits of the cloud are a 30% reduction in infrastructure costs and a drastic improvement in resilience, stability, and time to market."

Laurent Donnay
CEO,
Deutsche Telekom

02 Organizations struggle to harness On-Demand tech

Costs of On-Demand tech are accelerating

Surging AI adoption and demand for digital infrastructure, as well as inflation, are pushing up cloud costs across sectors. Major SaaS productivity players such as Google Workspace have implemented a 20% price hike to their subscription plans, for example.[15] Similarly, Microsoft 365 pushed through a 9% rise earlier in 2023.[16] The industry also faces "SaaS shrinkflation,"[17] where vendors offer reduced functionality at the same rates.
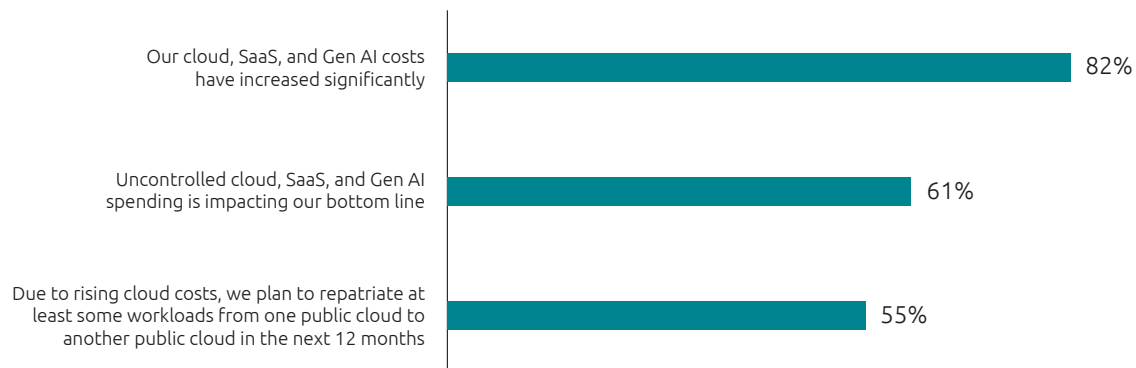
The complexity of cloud pricing models, coupled with limited visibility of usage and value realization, has amplified the need for better financial governance and performance tracking. As **Figure 3** highlights:

- 82% of executives agree that their cloud, SaaS, and Gen AI costs have increased significantly

- 61% say uncontrolled costs are impacting their profitability

- More than half (55%) say they plan to repatriate at least some workloads from one public cloud to another – suggesting a '"cloud arbitrage" trend[18]

Figure 3.

More than four out of five executives report rising On-Demand tech costs

Percentage of executives who agree with the statements below

Our cloud, SaaS, and Gen AI costs have increased significantly — 82%

Uncontrolled cloud, SaaS, and Gen AI spending is impacting our bottom line — 61%

Due to rising cloud costs, we plan to repatriate at least some workloads from one public cloud to another public cloud in the next 12 months — 55%

Source: Capgemini Research Institute, Cloud and On-Demand tech economics, May 2025, N = 1,000 executives; N = 989 executives are aware of or involved in public cloud-related activities/spend in their organization.

More than four in 10 (45%) organizations are even considering repatriating some workloads to private cloud. This is driven not only by cost optimization, but also AI/Gen AI, security, and compliance requirements, as well as data, operational, and technical sovereignty needs.

82%

of executives agree that their cloud, SaaS, and Gen AI costs have increased significantly

On-Demand tech expenses have burst through budgets

As On-Demand tech spending continues to increase, many organizations are facing cost overruns and unexpected expenses. According to Gartner's 2024 *Cloud Spending report*, 69% of IT leaders reported budget overruns in their organizations' cloud spending.[19] **Figure 4**, below, confirms these findings.

Of those organizations that have a dedicated On-Demand technology budget:

- 76% say they have exceeded their public cloud budget in the past 12 months, by 10% on average. Nearly half (48%) of organizations say they have overspent by more than 10%.

- 68% say their Gen AI spend and 52% their SaaS spend overshot their budgets, each by an average 11%. Nearly 20% of organizations have overspent by more than 15% on SaaS.
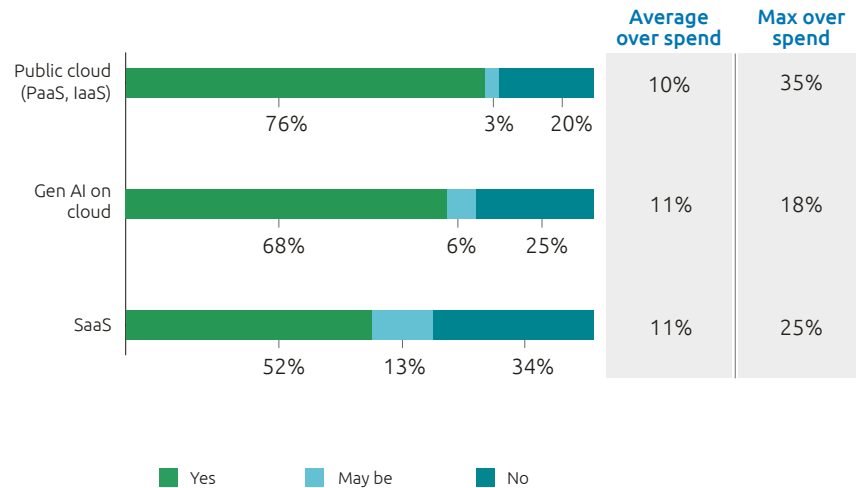
Scott Sellers, CEO of Azul, a US-based Java platform organization, talks about the "Jevons Paradox"[20] in cloud spending: *"The fact that cloud is inherently more cost-effective encourages consumption and leads to budget overruns."*[21] In our research, more than half (53%) of respondents said that IT costs are now a CEO-level discussion topic. The top reasons for budget overruns include:

- Lack of cost visibility (65% of those reporting budget overrun for On-Demand technologies rank this among the top five factors)

- Absence of cost governance (56%)

- Underutilized and/or overprovisioned resources (50%)

- Price increase from vendors (50%)
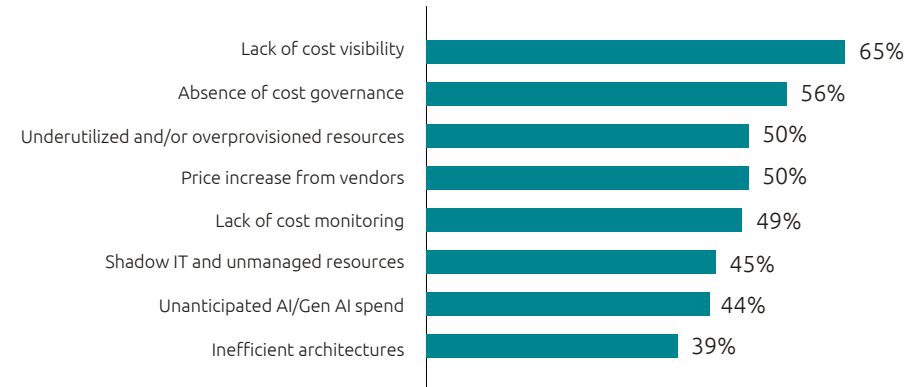
- Lack of cost monitoring (49%)

Figure 4.

Most organizations struggle with On-Demand tech costs, with poor visibility and governance being the top reasons for the spike

Percentage of organizations that have overrun their On-Demand tech budgets in the past year

	Average over spend	Max over spend
Public cloud (PaaS, IaaS) — Yes 76%, May be 3%, No 20%	10%	35%
Gen AI on cloud — Yes 68%, May be 6%, No 25%	11%	18%
SaaS — Yes 52%, May be 13%, No 34%	11%	25%

Legend: ■ Yes ■ May be ■ No

Percentage of executives citing the below as top reasons for their On-Demand tech budgets overrun

Reason	%
Lack of cost visibility	65%
Absence of cost governance	56%
Underutilized and/or overprovisioned resources	50%
Price increase from vendors	50%
Lack of cost monitoring	49%
Shadow IT and unmanaged resources	45%
Unanticipated AI/Gen AI spend	44%
Inefficient architectures	39%

Many organizations still treat cloud infrastructure like on-premises setups, leading to uncontrolled On-Demand tech costs. This stems from a legacy mindset – teams lack real-time visibility and governance, and change management still follows outdated models. Without adapting to cloud-native thinking, businesses risk inefficiencies and budget overruns in their digital transformation journeys. On-Demand tech cost accountability should span every level – from CFOs and heads of infrastructure to DevOps engineers.

Organizations are investing more in AI technologies including Gen AI, AI agents, and agentic AI – which are inherently resource-intensive. They are driving the demand for high-performance compute and real-time data processing. Models such as pay-per-task are fueling the next wave of public cloud spending. Nearly three-quarters (74%) of executives in our research agree that Gen AI has led to an unexpected surge in cloud consumption costs. A recent study highlights that 72% of IT and financial leaders believe that Gen AI-driven cloud spending is becoming unmanageable.[22]

Rising "unmanaged" IT is fueling cost overruns and operational inefficiencies

Business, not IT, is now steering SaaS and Gen AI spend

Technology investment is no longer solely the domain of IT. As **Figure 5** shows, 59% of Gen AI spend and 48% of SaaS spend now comes from business functions. Moreover, an additional 12% of SaaS spend is "unmanaged" or "unsanctioned" – i.e., the IT team has no direct knowledge of or involvement with it.

Figure 5.

IT drives less than 40% of SaaS and Gen AI spend

Estimated split of On-Demand technology spend by function*

Public cloud (PaaS, IaaS): 60% / 32% / 8%
Gen AI on cloud: 32% / 59% / 8%
SaaS: 40% / 48% / 12%

■ Spend by IT function   ■ Spend by business units   ■ Spend by unmanaged IT
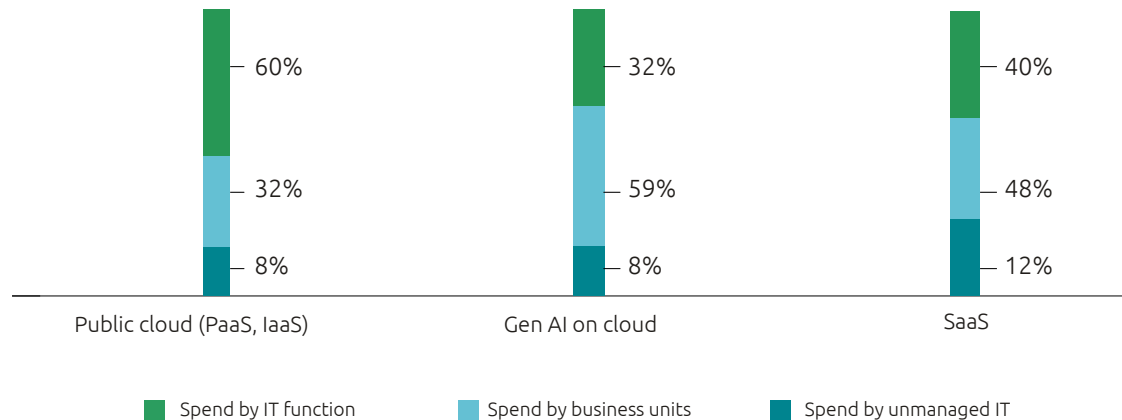
Note: Spend by business units includes technology purchases made by the business unit through approved procurement process and serviced/managed by the IT team. Spend by unmanaged IT includes technology purchases within a team or business unit without the direct knowledge/awareness of IT team, and that is not serviced/managed by the IT team.
* As per the best estimate provided by executives from technology and finance functions.
Source: Capgemini Research Institute, Cloud and On-Demand tech economics, May 2025, N = 616 executives from technology and finance functions who say they are aware of the split of their organization's public cloud spend between IT, business and unmanaged IT; N = 494 executives from technology and finance functions who say they are aware of the split of their organization's Gen AI on cloud spend between IT, business and unmanaged IT; N = 535 executives from technology and finance functions who say they are aware of the split of their organization's SaaS spend between IT, business, and unmanaged IT.

Business is procuring On-Demand tech bypassing IT, resulting in increased costs, and visibility and security challenges

Gartner estimates that 41% of employees acquired, modified, or created technology outside of IT's visibility in 2022. This number is expected to grow to 75% by 2027.[23] Our previous research on Gen AI at work highlighted that nearly 46% of employees and 33% of leaders and managers use Gen AI in a personal capacity, either in ignorance or defiance of their organization's policy.[24] In this research, 98% of business executives acknowledge that they purchase technology directly, rather than via a central IT team – of which 67% do it either frequently or occasionally (see **Figure 6**).

A technical architect for cloud FinOps at an India-based large bank told us: *"Shadow IT emerges when business units consume cloud independently. Finance often sees only the bill, not the waste behind it. Without centralized visibility, organizations risk duplication, inefficiencies, and missed opportunities for optimization."*

Figure 6.
Two-thirds of business executives bypass IT teams to get On-Demand tech

Percentage of business executives purchasing On-Demand technology (public cloud, Gen AI on cloud, or SaaS) directly as opposed to via IT teams[25]

Legend:
- Yes, very frequently
- Yes, sometimes
- Yes, very rarely
- No

2%
8%
31%
58%

The most common reasons for bypassing IT for On-Demand tech purchases are desire for faster solutions (80% of business function executives agree); desire for flexible and adaptive solutions (77%); lengthy approval processes (72%); and desire for greater control over tech budgets (72%) (see **Figure 7**). *"Employees seek out unauthorized software to make their work easier, more efficient, better, or all three,"* notes Uzi Dvir, CIO at US-based SaaS provider WalkMe.[25]

98%

of business leaders admit to bypassing IT for tech purchases (8% frequently, 58% occasionally, and 31% on rare occasions)

Source: Capgemini Research Institute, Cloud and On-Demand tech economics, May 2025, N = 250 executives from business functions.

Figure 7.

Desire for agility, adaptability, and autonomy are driving business units to bypass IT teams for On-Demand tech procurement

Percentage of business executives who cite the below as reasons for their business unit to purchase On-Demand tech directly as opposed to going via central IT

Reason	Percentage
Desire for faster solutions	80%
Desire to have more flexible solutions that can adapt to changing business needs	77%
Lengthy approval process with IT/administrative reasons	72%
Desire to have greater control over business function's tech budgets	72%
Desire to have more innovative solutions	64%
Desire to have greater control and autonomy over tech implementation	58%
Desire for specific/tailored solutions	55%

Source: Capgemini Research Institute, Cloud and On-Demand tech economics, May 2025, N = 244 executives from business functions who say they purchase On-Demand technology directly (either frequently, occasionally or rarely) as opposed to going via an IT team.
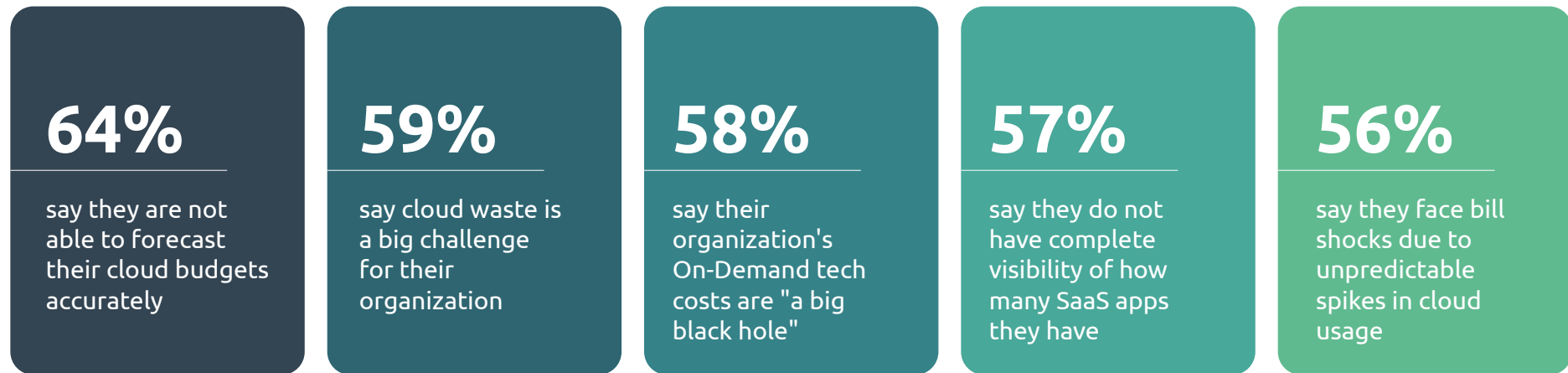
Three-quarters (75%) of technology and finance executives in our research say that IT without centralized oversight leads to duplicative spending; 67% say that it leads to higher overall tech costs; and 67% also say it leads to increased security and compliance violations.

Charlie Livingston, Head of Infrastructure and Security at UK-based financial well-being platform Wagestream, says: *"There are two points about shadow IT from a security standpoint. The first is the SaaS spend and a lot of people saying: 'Oh, I just need a little tool in my browser to do text-to-speech.' It costs $10 a month. Across 200 employees and 20 different platforms, it gets expensive quickly. The second point is dealing with: What is that browser tool actually reading? Have you read the terms of service? Is it cheap? Oh, it's a free tool. Okay, well why is it free? What data are they selling?"*[26]

From our research, some concerning stats emerge:

Figure 8.

On-Demand tech sprawl is creating challenges

64%

say they are not able to forecast their cloud budgets accurately

59%

say cloud waste is a big challenge for their organization

58%

say their organization's On-Demand tech costs are "a big black hole"

57%

say they do not have complete visibility of how many SaaS apps they have

56%

say they face bill shocks due to unpredictable spikes in cloud usage

Source: Capgemini Research Institute, Cloud and On-Demand tech economics, May 2025, N = 1,000 executives; N = 989 executives aware of or involved in public cloud-related activities/spend in their organization; N = 999 executives aware of SaaS-related activities/spend in their organizations.

"There are two points about shadow IT from a security standpoint. The first is the SaaS spend and a lot of people saying: 'Oh, I just need a little tool in my browser to do text-to-speech.' It costs $10 a month. Across 200 employees and 20 different platforms, it gets expensive quickly. The second point is dealing with: What is that browser tool actually reading? Have you read the terms of service? Is it cheap? Oh, it's a free tool. Okay, well why is it free? What data are they selling?"

Charlie Livingston

Head of Infrastructure and Security
Wagestream

# Despite positive outcomes, On-Demand tech falls short of expectations

On-Demand tech has driven positive outcomes for organizations, but it nevertheless falls short in some respects (see **Figure 9**). For example, only 33% of executives achieved the expected quality of service-related outcomes from public cloud investments. Only 38% achieved the expected faster product innovation with Gen AI; and only 29% say they "fully or mostly" achieved expected cost savings from SaaS.

Executives cite poor cost management (71%), inefficiencies such as idle or underused resources (62%), and over-provisioning (58%) among the top reasons hindering expected outcomes from On-Demand tech.

Figure 9.

A minority of executives say they have achieved the expected quality of service, innovation, or cost savings outcomes from their On-Demand tech investments

Percentage of executives who say they have either "fully" or "mostly" achieved the expected outcomes from On-Demand technologies
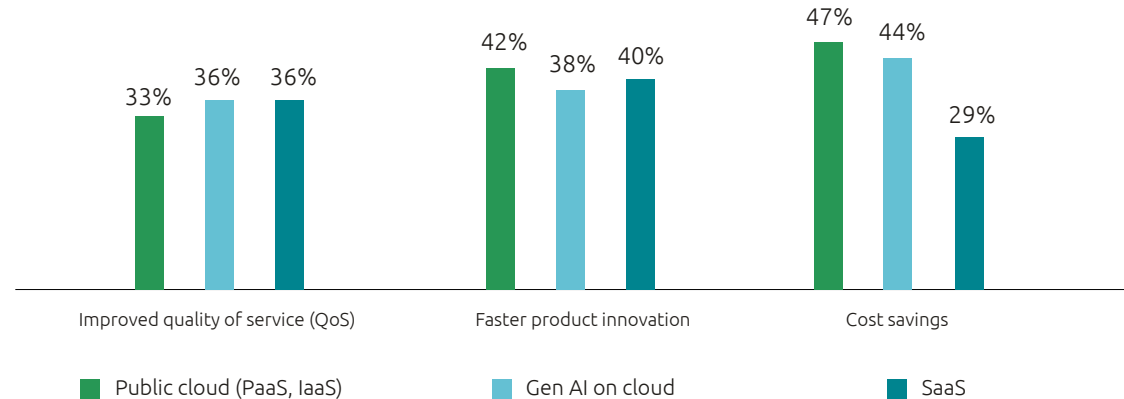
Source: Capgemini Research Institute, Cloud and On-Demand tech economics, May 2025, N = 1,000 executives, N = 938 executives who say their organization has achieved benefits from public cloud investments in the past year; N = 977 executives who say their organization has achieved benefits from Gen AI on cloud investments in the past year; N = 973 executives who say their organization has achieved benefits from SaaS investments in the past year.

Moreover, nearly six in 10 executives say they struggle to gauge ROI of public cloud (58%), Gen AI on cloud (58%), and SaaS (56%) On-Demand tech, often due to unclear business cases and lack of standardized metrics. To align investments with business value, organizations need better tools, frameworks, and practices, such as FinOps, cross-team collaboration, and a "cloud-smart" (not "cloud-first") mindset.[27]

Only **33**^%

of executives achieved the expected quality of service-related outcomes from their public cloud investments

Navigating cloud sovereignty

Concerns around cloud sovereignty – including data, operational, and technical issues – have been mounting over recent years.

But the technology is coming under increasing scrutiny amid geopolitical tensions; the shifting regulatory and compliance landscape; the dominant role of cloud players concentrated in a few regions; and heightened concerns around where critical data is stored and processed, and by whom.

As a result, governments and organizations are re-evaluating their external exposure to maintain physical and digital control over strategic assets, including data, algorithms, and critical software.
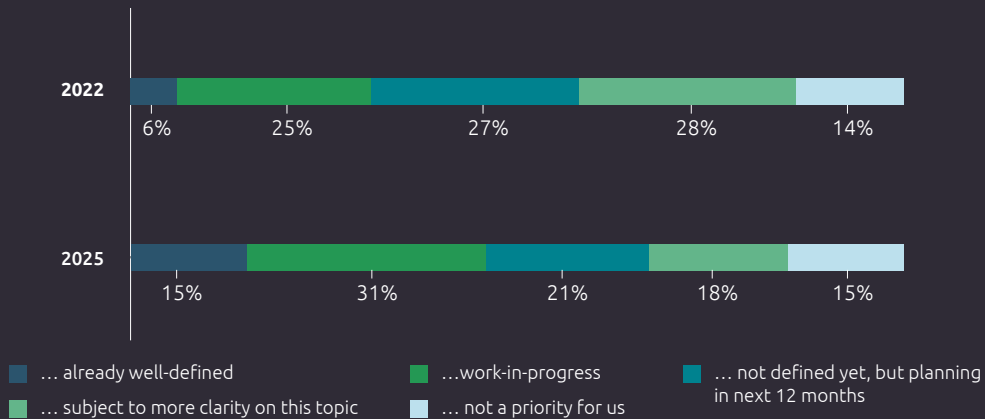
IDC estimates that global spending on sovereign cloud will rise to more than $250 billion in 2027.[28] Our research shows that **46% of organizations are already embedding cloud sovereignty in overall cloud strategy** (up from 31% in 2022),[29] and 21% expect to start in the next 12 months (see **Figure 10**). These numbers are even higher for Europe, where 50% of organizations say they either already have or are currently working on a sovereign cloud strategy. Analyzing by sector, industrial manufacturing (60%), public sector (57%), and financial services (53%) lead in terms of well-defined or work-in-progress (WIP) sovereign cloud strategies.

46%

of organizations are already embedding cloud sovereignty in overall cloud strategy

Figure 10.

Organizations are actively investigating cloud sovereignty

Percentage of executives responding to the current state of their organzation's cloud sovereignty strategy

2022
6%   25%   27%   28%   14%

2025
15%   31%   21%   18%   15%

- … already well-defined
- …work-in-progress
- … not defined yet, but planning in next 12 months
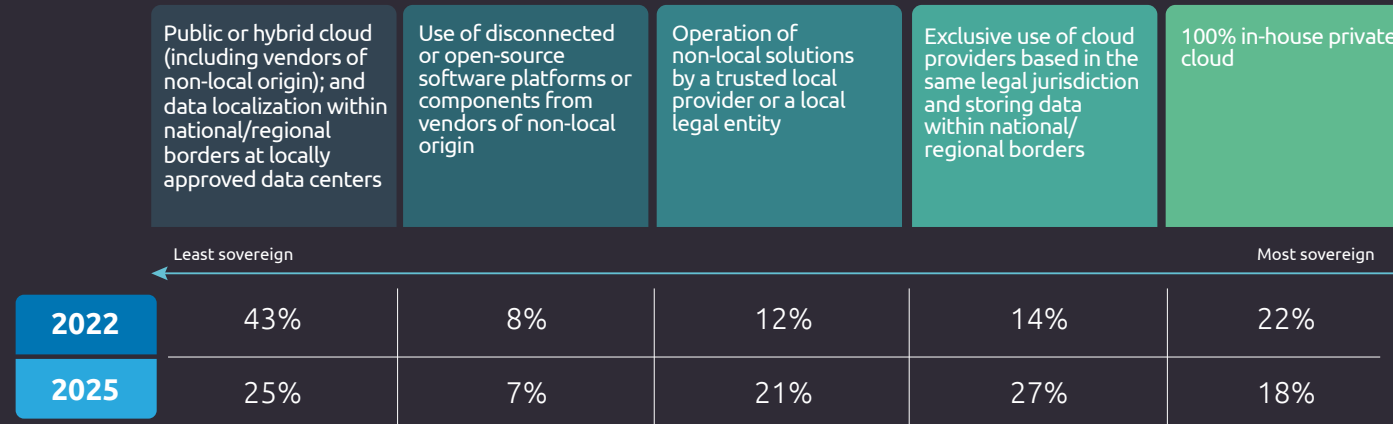- … subject to more clarity on this topic
- … not a priority for us

Source: Capgemini Research Institute, Cloud and On-Demand tech economics, May 2025, N = 750 executives from technology and finance functions; Capgemini Research Institute, Cloud sovereignty research, July 2022, N = 1,000 executives.

The definition of cloud sovereignty has evolved. We asked the same question in 2022, and 43% of executives said they focus on "data localization."[30] In 2025 only 25% limit cloud sovereignty to data localization, while 27% define it as exclusive use of cloud providers based in the same legal jurisdiction and storing data within a country/regional border. A further 21% emphasize the presence of a local legal entity operating non-local cloud solutions (see **Figure 11**). Capgemini and Orange's joint-venture Bleu, for example, engages with select French public and private organizations to offer "***cloud de confiance***" services based on Microsoft technology.[31]

Figure 11.

Organizations now associate cloud sovereignty with either cloud providers based in the same legal jurisdiction or a local legal entity of a hyperscaler

What does your organization most closely associate with cloud sovereignty?

	Public or hybrid cloud (including vendors of non-local origin); and data localization within national/regional borders at locally approved data centers	Use of disconnected or open-source software platforms or components from vendors of non-local origin	Operation of non-local solutions by a trusted local provider or a local legal entity	Exclusive use of cloud providers based in the same legal jurisdiction and storing data within national/regional borders	100% in-house private cloud
	Least sovereign →				← Most sovereign
2022	43%	8%	12%	14%	22%
2025	25%	7%	21%	27%	18%

Source: Capgemini Research Institute, Cloud and On-Demand tech economics, May 2025, N = 750 executives from technology and finance functions; Capgemini Research Institute, Cloud sovereignty research, July 2022, N = 1,000 executives.

A few interesting regional and sectoral variances emerge here:

- Only 17% of US-headquartered organizations define cloud sovereignty as "exclusive use of cloud providers based in the same legal jurisdiction and storing data within national/ regional borders" – in comparison with 25% of Europe-headquartered and 39% of APAC-headquartered organizations.

- More than half of organizations from energy and utilities (54%), aerospace and defense (56%), and life sciences (57%) define cloud sovereignty either as "exclusive use of cloud providers based in the same legal jurisdiction and storing data within national/ regional borders" or "operation of non-local solutions by a trusted local provider or a local legal entity."

The shift often increases operational costs due to the limitations of localized infrastructure, stricter compliance measures, and limited vendor flexibility. Moreover, half (51%) of executives who have well-defined, or WIP sovereign cloud strategies face higher operational and maintenance costs.

Despite the expense, there is a strong imperative to retain control over data; manage regulatory risk; adapt to different country/ regional policy environments; avoid fines and penalties; manage reputational risk; and achieve long-term strategic resilience. *"Sovereign cloud isn't a cost decision. It's a compliance mandate. We tag data by region, so we can move quickly if regulations change. You can't build resilience after the rule arrives,"* says an executive from a large pharmaceutical organization.

In our research, as **Figure 12** shows, 42% are definitely and 37% tentatively willing to pay a price premium for sovereignty – of 11% on average.
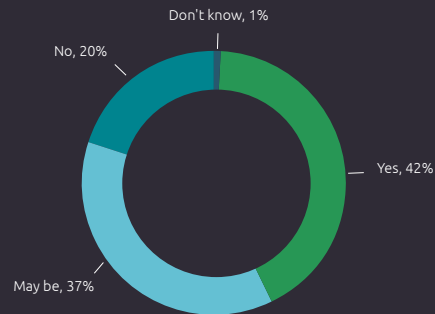
42%

of executives are definitely willing to pay a price premium for sovereignty

Figure 12.

More than four in 10 executives are willing to pay a premium for sovereign cloud

Percentage of executives responding to their willingness to pay a price premium for sovereign cloud

Don't know, 1%

No, 20%

Yes, 42%

May be, 37%

Average price premium (over current cloud costs) that organizations are willing to pay for sovereign cloud

11%

Source: Capgemini Research Institute, Cloud and On-Demand tech economics, May 2025, N = 348 executives from technology and finance functions who say they have either a well-defined or work-in-progress sovereign cloud strategy.

Source: Capgemini Research Institute, Cloud and On-Demand tech economics, May 2025, N = 276 executives from technology and finance functions who say they are definitely or tentatively willing to pay a premium for sovereign cloud.

Of those organizations who have well-defined or WIP sovereign cloud strategies:

52%
adopt sovereign cloud solutions selectively, based on data classification and criticality

58%
conduct a cost-benefit analysis to balance sovereignty needs with cost efficiency

68%
integrate sovereign cloud solutions into their existing infrastructure to manage costs

03 Managing the cost of On-Demand tech
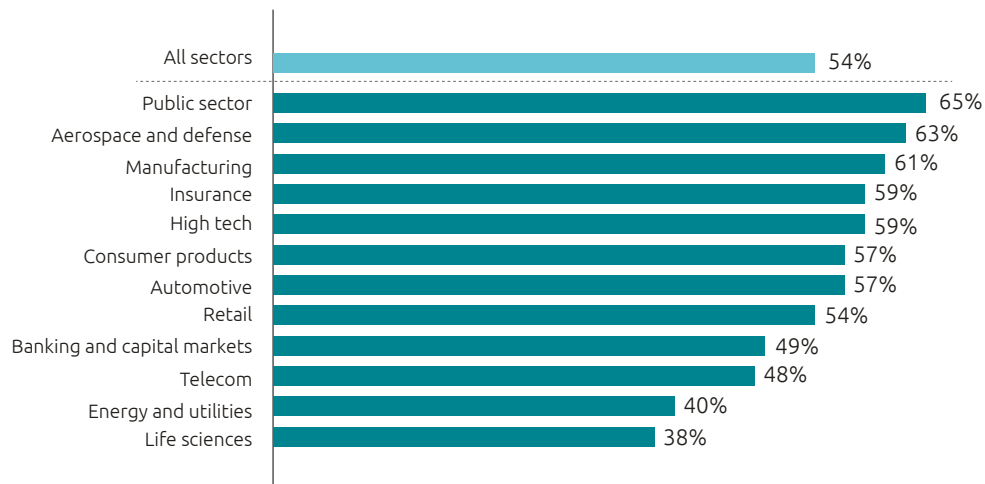
On-Demand tech costs are an afterthought

As **Figure 13** highlights, many organizations (54%) first choose cloud, then think about costs – this pattern is especially evident in public sector (65%), aerospace and defense (63%) and manufacturing (61%) organizations.

J.R. Storment, Executive Director of the FinOps Foundation, highlights: *"Only considering cost after deployment can lead to unwelcome outcomes, such as surprisingly high cloud bills, lower than acceptable product margins, and reduced options for cost optimization due to earlier design decisions."* [32]

Figure 13.
For more than half, On-Demand tech costs are an afterthought

Percentage of executives who agree with the statement: "We usually think about cloud and other On-Demand tech costs after a product has been built and launched"

Sector	Percentage
All sectors	54%
Public sector	65%
Aerospace and defense	63%
Manufacturing	61%
Insurance	59%
High tech	59%
Consumer products	57%
Automotive	57%
Retail	54%
Banking and capital markets	49%
Telecom	48%
Energy and utilities	40%
Life sciences	38%

Source: Capgemini Research Institute, On-Demand tech economics, May 2025, N = 1,000 executives.

We believe that shifting left is key to enhancing visibility, reducing cloud bill shocks, improving efficiency and resource allocation, enriching quality, reducing security risks, as well as decreasing technical debt.

Despite growing adoption, FinOps remains limited in reach and impact

The FinOps Foundation defines FinOps as *"an operational framework and cultural practice that maximizes the business value of cloud and technology, enables timely data-driven decision-making, and creates financial accountability through collaboration between engineering, finance, and business teams."*[33]

FinOps enables organizations to answer two key questions: Where are the costs generated? And what business value does that spending provide? FinOps facilitates greater flexibility and agility for investments in On-Demand tech in terms of how they spend, how and when they should shut resources down, how long they need to occupy a given environment, and what the working hours for that environment should be. A large luxury retailer identified 20% operational savings using Capgemini's FinOps business case analyses.[34] A vacation experience organization working with Capgemini achieved 30% cost savings on overlooked AWS assets within 10 days, and 10% overall AWS cost optimization with FinOps.

In our survey, as **Figure 14** shows, 76% of executives say they either have a dedicated FinOps team (29%) or are building one in the next 12 months (47%).[35]
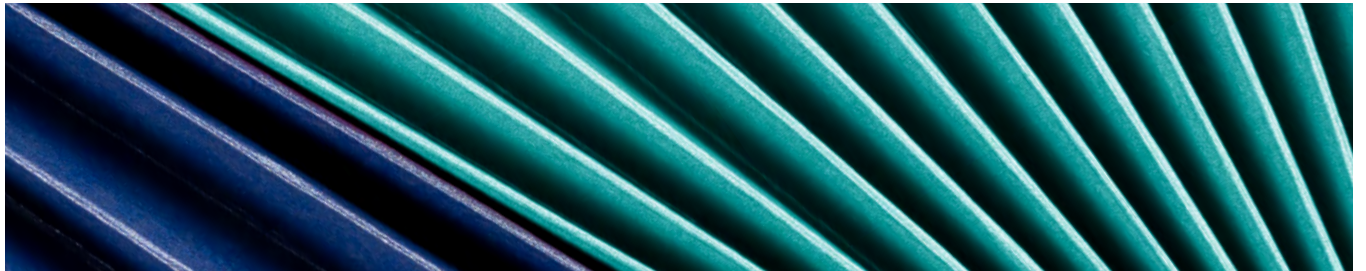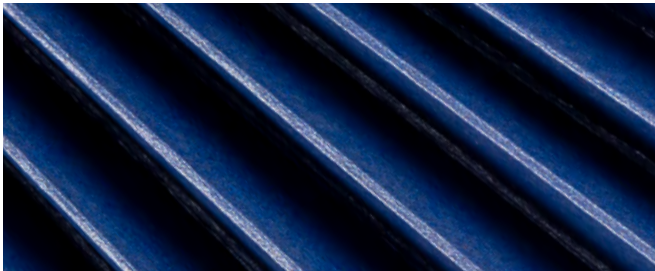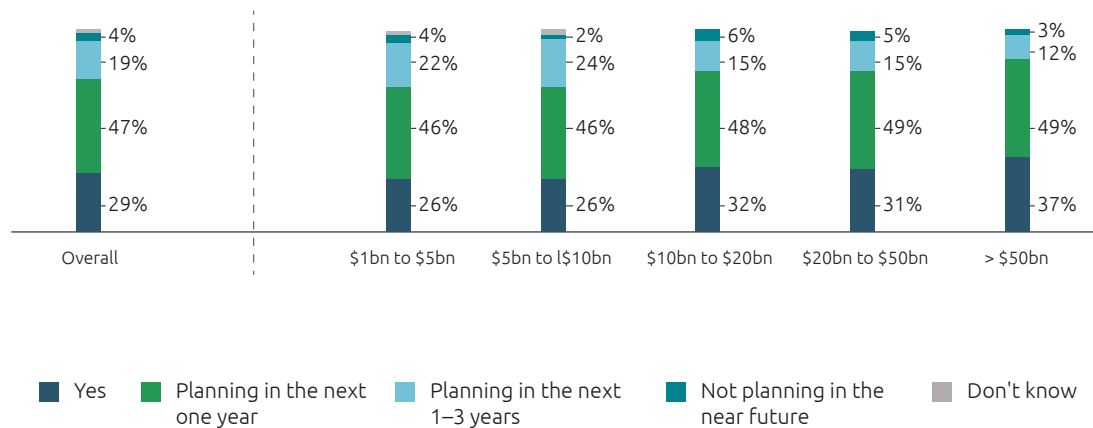
Figure 14.

Three-quarters (76%) of executives say they either have or plan to have a dedicated FinOps team

Percentage of executives responding to the question: "Does your organization have a dedicated FinOps function?"

- Yes
- Planning in the next one year
- Planning in the next 1–3 years
- Not planning in the near future
- Don't know

Source: Capgemini Research Institute, Cloud and On-Demand tech economics, May 2025, N = 750 executives from technology and finance functions.

76%

of executives say they either have a dedicated FinOps team or are building one in the next 12 months

However, FinOps practices have run into a few challenges:

FinOps has limited scope

Organizations use an ever-increasing variety of consumption-based technologies. The FinOps Foundation is also widening its scope of practice to include SaaS and AI,[36][37]albeit these concepts are in their infancy. Moreover, many SaaS vendors are not adopting the FinOps Foundation's FinOps Open Cost and Usage Specification (FOCUS) guidelines to standardize billing data.

Among those organizations in our research who have a dedicated FinOps function, more than half (51%) say FinOps covers cloud only. Nearly four in 10 say it covers SaaS as well, while only 2% say it covers all AI in addition to cloud and SaaS (see **Figure 15**).

A technical architect for cloud FinOps at an India-based large bank adds: *"FinOps is no longer just about cloud. It's now 'cloud plus' – covering SaaS, PaaS, licensing, and Gen AI. The FinOps Foundation has expanded its scope, and organizations must adapt to manage all on-demand tech spend holistically."*

Figure 15.
Half of FinOps teams are focused only on cloud services

Percentage of executives responding to the question: "Which areas does your FinOps team currently cover?"
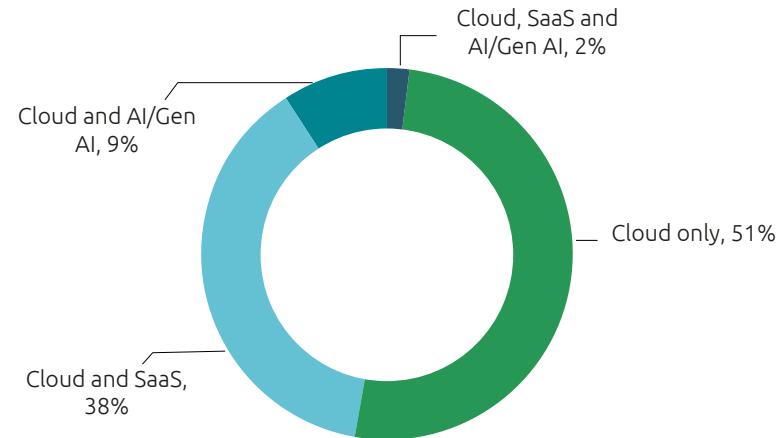
Cloud, SaaS and AI/Gen AI, 2%

Cloud and AI/Gen AI, 9%

Cloud only, 51%

Cloud and SaaS, 38%

Source: Capgemini Research Institute, Cloud and On-Demand tech economics, May 2025, N = 219 executives from technology and finance functions who say they have a dedicated FinOps function currently.

FinOps teams focus on operational initiatives, rather than strategic ones

Among organizations that have a dedicated FinOps team in place, a majority (63%) focus on operational and tactical activities. Only 28% say their FinOps team focuses on strategic activities such as providing unit economics[38] delivering forecasts, and driving change management (see **Figure 16**). Flexera's State of the Cloud 2025 report highlights that only 40% of all organizations surveyed use a unit economics model for cloud cost analysis.[39] Gerhard Schauer, Vice-President, Global IT Workplace Services at Germany-based ZF Group, a supplier of advanced mobility systems for automakers, says: *"About 30% of our FinOps efforts are now strategic, focused on forecasting, planning, and validating pilot use cases, especially in R&D and engineering."*

Figure 16.

Fewer than three in 10 organizations say their FinOps team drives strategic initiatives

Percentage of executives who agree with the below statements
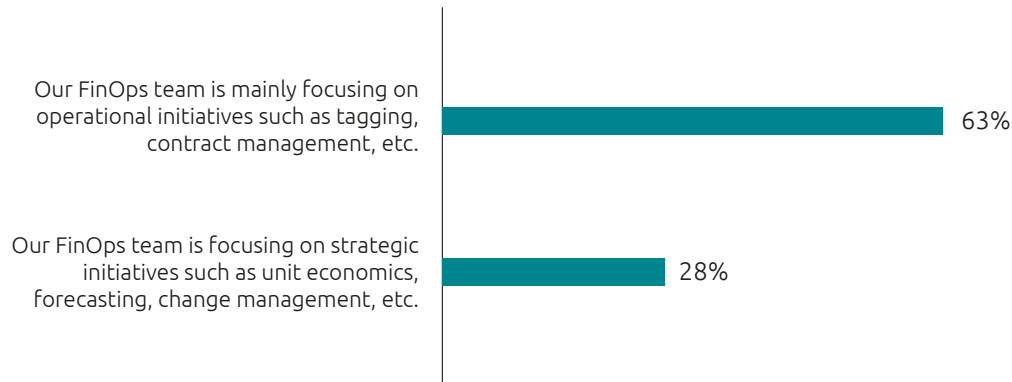
Source: Capgemini Research Institute, Cloud and On-Demand tech economics, May 2025, N = 219 executives from technology and finance functions who say they have a dedicated FinOps function currently.

"About 30% of our FinOps efforts are now strategic, focused on forecasting, planning, and validating pilot use cases, especially in R&D and engineering."

Gerhard Schauer

Vice-President,
Global IT Workplace Services
ZF Group

FinOps teams are unable to influence business decision-making

Car retailer CarMax has built a recommendation engine into its developer portal that shows cloud usage trends, calls out cost anomalies, and makes daily recommendations. But CITO Shamim Mohammad points out: *"One challenge for this year is to ensure the organization follows up on those recommendations, and that compute costs are part of the calculation when evaluating new initiatives."*[40] As **Figure 17** shows, among organizations with a dedicated FinOps team, only 42% say theirs can influence business decision making and implementation of the recommendations it makes.

Figure 17.

Only four in 10 organizations say their FinOps team drives implementation of recommended optimizations

Percentage of executives who agree with the statement: "Our FinOps team is able to influence business decision-making"
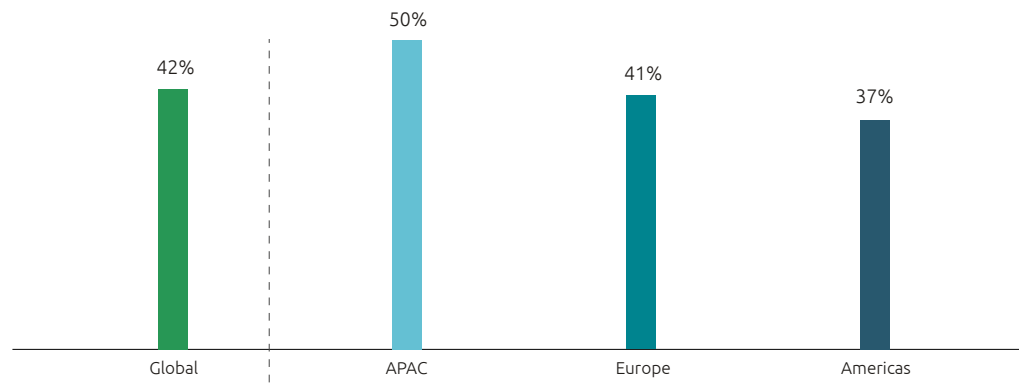
Global	APAC	Europe	Americas
42%	50%	41%	37%

Source: Capgemini Research Institute, Cloud and On-Demand tech economics, May 2025, N = 219 executives from technology and finance functions who say they have a dedicated FinOps function currently.

Some principal reasons emerge for this:

- Even the most successful cloud FinOps capabilities are **limited in their influence on demand management.**

- Secondly, **FinOps teams don't usually report to the C-Suite**. This limits their influence on strategic decision-making. J.R. Storment from the FinOps Foundation emphasizes: *"Executive buy-in has a massive influence in building successful FinOps practices, and the role of a CIO or CFO is often to ensure that the FinOps strategy and deployment are well crafted and carried out."*[41] Placing the FinOps team under the COO hierarchy, with a matrix reporting into both the CFO and the CTO/CIO office could ensure stronger accountability.

- Next, there is a **lack of organizational alignment across various stakeholders from finance, IT, procurement, engineering, and business teams.** In our research, 60% of executives say lack of internal collaboration is a prominent reason for their organization failing to achieve expected outcomes from On-Demand tech investments. To counter this, CarMax has built a cross-functional FinOps team, reporting to the CIO.[42]

- Moreover, FinOps teams are **unable to articulate the business value –** in terms of KPIs such as customer acquisition cost, time-to-market, or revenue impact. And, without strong business acumen, stakeholder engagement, and storytelling skills, FinOps teams struggle to influence decision-makers.

- And finally, **KPIs for the FinOps teams need to be linked to spend optimization and not just opportunity discovery.** But only 45% of organizations in our research say they have defined KPIs to measure On-Demand technologies' cost-optimization success.

- Summing up, FinOps needs more than just cost control. It requires a mindset shift, along with holistic sponsorship and governance across all functions, including architecture, sourcing, finance, and cloud usage. It's not just a business or IT concern; it's a system-wide approach to influence technology choices and drive financial accountability.

"FinOps teams' influence is often limited to the technology vertical. Many teams are strong at highlighting cost implications and discovering efficiencies, and even help shape cost-effective solutions, but they rarely influence business outcomes directly. Once a FinOps team can truly embrace and deliver the cultural and behavioral science of change, they can have much bigger and lasting impact within the business."

Jez Back

Cloud Economist and Global Offer Leader
for Cloud Consumption On-Demand
Capgemini

Most use cloud cost management tools, but don't make the most of them

The global cloud cost management tools market was valued at $9.8 billion in 2024 and is projected to grow at a CAGR of 17.2% between 2025 and 2034.[43] Our research shows that six in 10 executives from finance and technology functions use them. These offer wide visibility of cloud spend; aid in budgeting and forecasting; minimize waste by identifying idle, unattached, underused, or unused instances; right-size overprovisioned instances; and even recommend appropriate commitment-based usage discounts.
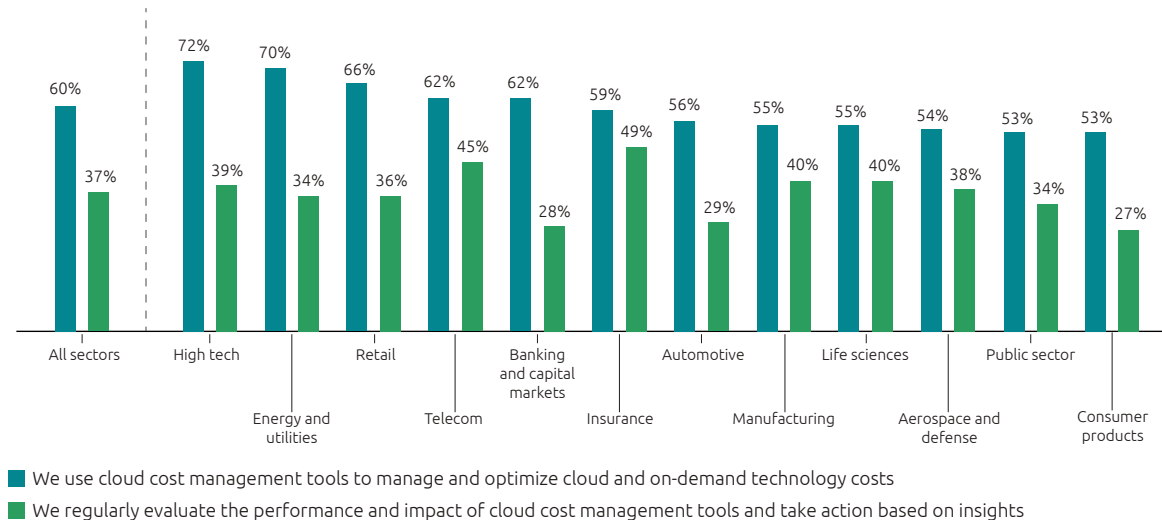
But only 37% of executives evaluate the impact of the tools and act upon the insights they provide (see **Figure 18**). This could be due to lack of ownership, tool complexity, siloed operations, or lack of well-defined success metrics.

Investing in cloud cost management tools isn't enough. Organizations must evolve. True transformation lies in governance: embedding accountability, visibility, and strategic alignment across every layer of cloud operations.

Figure 18.

Most use cloud cost management tools but fail to act on the insights gained

Percentage of executives who agree with the statements below

- We use cloud cost management tools to manage and optimize cloud and on-demand technology costs
- We regularly evaluate the performance and impact of cloud cost management tools and take action based on insights

Source: Capgemini Research Institute, Cloud and On-Demand tech economics, May 2025, N = 750 executives from technology and finance functions.

"An often-overlooked aspect of governance is the tracking and management of the implementation of recommended optimization. In the absence of such mechanisms and KPIs, recommendations are often moved to the bottom of the priority list and may never see the light of day."

Thomas Sarrazin

Global FinOps Offer Lead
Capgemini

Sustainable FinOps is currently overlooked

Estimates suggest the ICT sector's share of global carbon emissions ranges between 1.5% and 4%.[44] Public cloud has a larger carbon footprint than the airline industry.[45] Growing demand for cloud services, accelerated by the Gen AI and agentic AI revolution, has increased carbon footprint concerns. Recent research forecasts that global power demand from data centers will increase by 50% by 2027, and by as much as 165% by the end of the decade (compared with 2023 levels).[46]

It is crucial to understand that wasted cloud and other On-Demand resources also come with a financial price tag. Dr. Werner Vogels, VP and CTO at Amazon, speaking at AWS re:Invent 2023, said: *"Cost is a close proxy for sustainability."* In our research, more than half (53%) of executives agree that sub-optimal On-Demand tech usage leads to excessive energy consumption and increased carbon emissions (see **Figure 19**).

53%

of executives agree that sub-optimal On-Demand tech usage leads to excessive energy consumption and increased carbon emissions

Figure 19.

Over half of executives link inefficient On-Demand tech usage to higher carbon emissions

Percentage of executives who agree with the statement "Sub-optimal cloud and on-demand tech usage lead to excess energy consumption and increased carbon emissions for our organization"
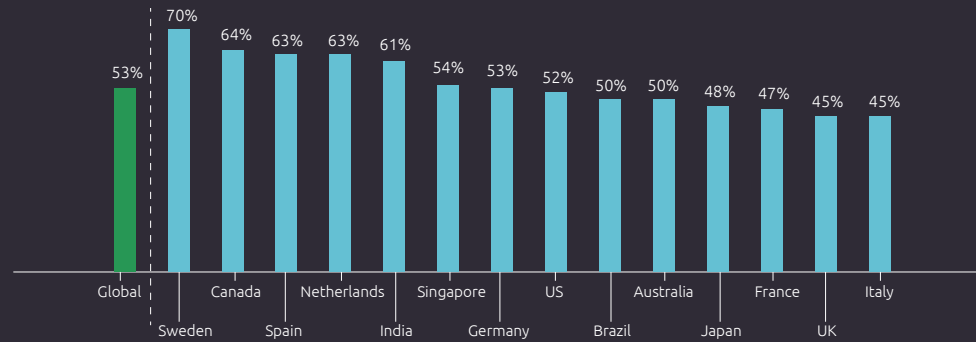
Source: Capgemini Research Institute, On-Demand tech economics, May 2025, N = 1,000 executives.

Today's fast-paced and dynamic digital landscape calls for a dual focus on cost and carbon efficiency through "sustainable FinOps" — where GreenOps is inherently aligned with FinOps.

- FinOps, when implemented correctly, is inherently sustainable. It promotes optimal usage of cloud resources, supporting cost efficiency, energy savings, and carbon-emissions reduction.

- GreenOps focuses on sustainable practices such as developing energy-efficient architectures, selecting regions that use renewable energy, deploying intelligent and cloud-native solutions such as serverless[47] and container solutions, and using code that requires less energy to run. These initiatives also contribute to cost-efficiency such as right-sizing, choosing a region closer to usage, switching off idle resources, and optimizing computing and storage.[48]

Pathik Sharma, Cloud FinOps Cost Optimization Lead at Google, adds: *"Our customers are actively discussing overlapping practices between FinOps and GreenOps. Often, we see when teams think about green and optimizing carbon, they also save on costs and vice versa. Also, with Google's Carbon Assessment tool, our clients can measure, assess, and reduce the carbon footprint of their cloud usage."*
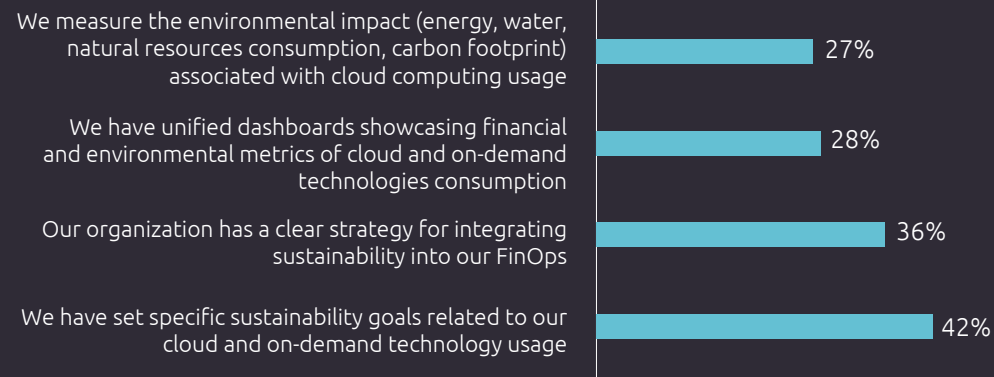
However, our research shows that current efforts fall short in this space. As Figure 20 shows:

- Only 27% of executives say their organization measures the environmental impact of cloud

- Only 28% have dashboards showcasing cost and carbon for On-Demand tech

- Only 36% have a strategy integrating sustainability in FinOps

- And just 42% have set targets related to sustainable On-Demand technology

Figure 20.

The adoption of sustainable FinOps lags across the board

Percentage of executives who agree with the below statements

We measure the environmental impact (energy, water, natural resources consumption, carbon footprint) associated with cloud computing usage — 27%

We have unified dashboards showcasing financial and environmental metrics of cloud and on-demand technologies consumption — 28%

Our organization has a clear strategy for integrating sustainability into our FinOps — 36%

We have set specific sustainability goals related to our cloud and on-demand technology usage — 42%

Source: Capgemini Research Institute, On-Demand tech economics, May 2025, N = 1,000 executives.

"Our customers are actively discussing overlapping practices between FinOps and GreenOps. Often, we see when teams think about green and optimizing carbon, they also save on costs and vice versa. Also, with Google's Carbon Assessment tool, our clients can measure, assess, and reduce the carbon footprint of their cloud usage."

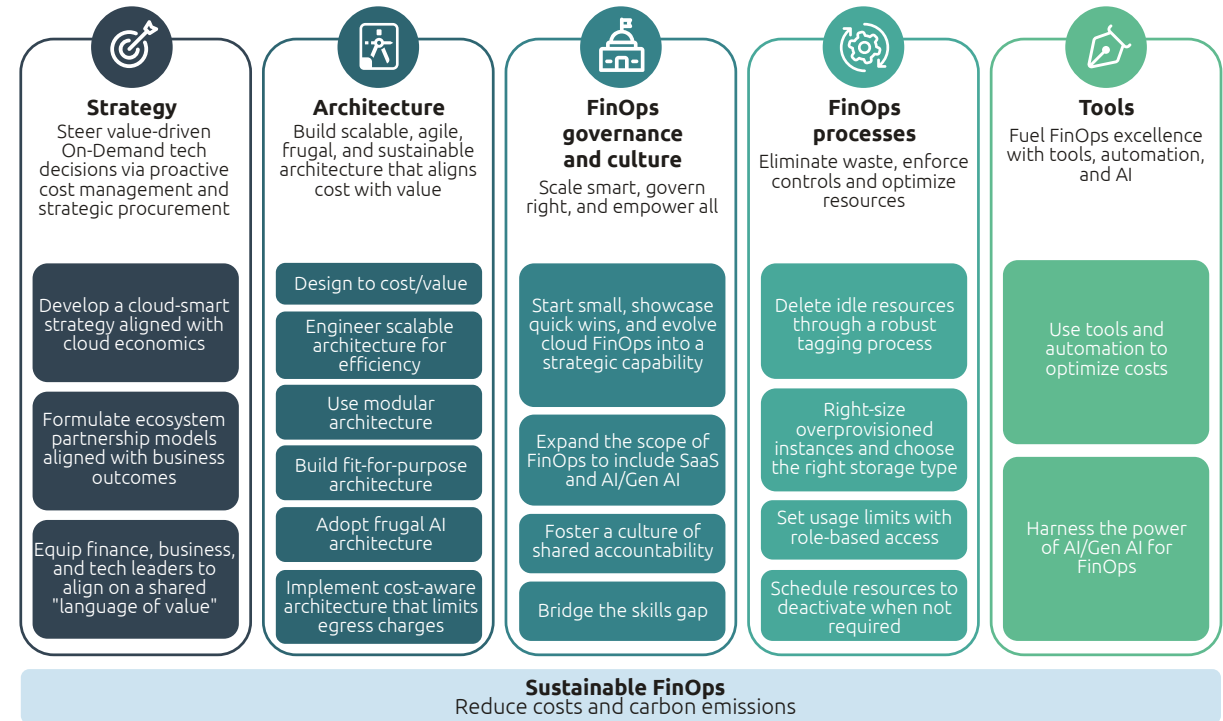**Pathik Sharma**

Cloud FinOps Cost Optimization Lead
Google
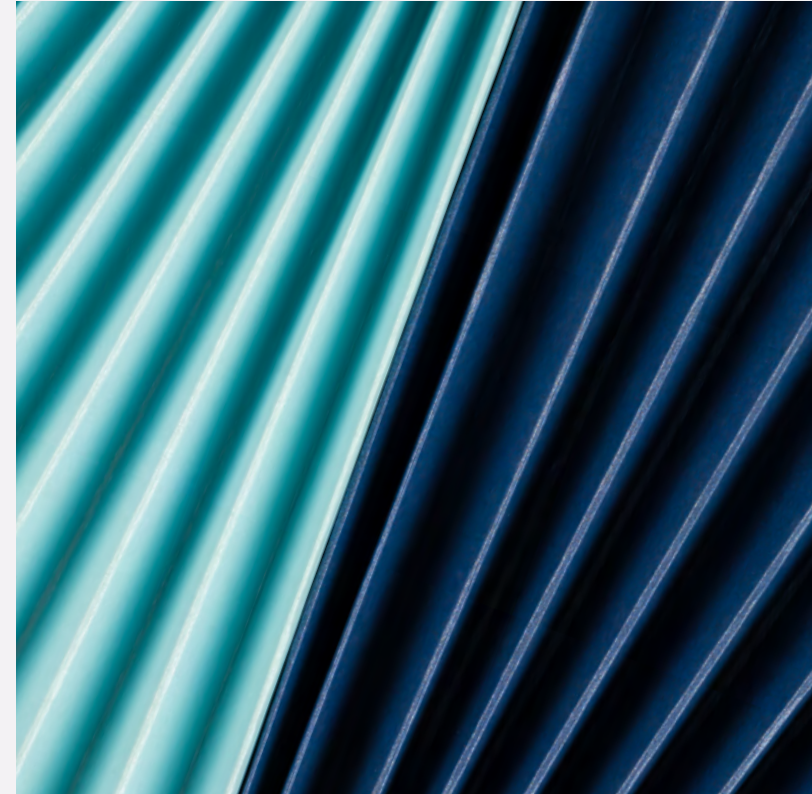
04 Optimizing costs and elevating the value of On-Demand Tech

Historically, cloud and other On-Demand tech have been measured as cost, but now, it's about value and driving revenue, resilience, responsiveness, and business re-engineering. To unlock the full economic potential of cloud and On-Demand tech, organizations must adopt a holistic approach that takes in strategy, architecture, governance, culture, tools, and processes. It is crucial to understand that on-demand tech economics is broadly a management problem, rather than just a tech resolution.

Figure 21.
Pillars of effective On-Demand tech

Strategy	Architecture	FinOps governance and culture	FinOps processes	Tools
Steer value-driven On-Demand tech decisions via proactive cost management and strategic procurement	Build scalable, agile, frugal, and sustainable architecture that aligns cost with value	Scale smart, govern right, and empower all	Eliminate waste, enforce controls and optimize resources	Fuel FinOps excellence with tools, automation, and AI
Develop a cloud-smart strategy aligned with cloud economics	Design to cost/value	Start small, showcase quick wins, and evolve cloud FinOps into a strategic capability	Delete idle resources through a robust tagging process	Use tools and automation to optimize costs
Formulate ecosystem partnership models aligned with business outcomes	Engineer scalable architecture for efficiency	Expand the scope of FinOps to include SaaS and AI/Gen AI	Right-size overprovisioned instances and choose the right storage type	Harness the power of AI/Gen AI for FinOps
Equip finance, business, and tech leaders to align on a shared "language of value"	Use modular architecture	Foster a culture of shared accountability	Set usage limits with role-based access	
	Build fit-for-purpose architecture	Bridge the skills gap	Schedule resources to deactivate when not required	
	Adopt frugal AI architecture			
	Implement cost-aware architecture that limits egress charges			

Sustainable FinOps
Reduce costs and carbon emissions

Source: Capgemini Research Institute analysis.

Strategy:
Steer value-driven On-Demand tech decisions via proactive cost management and strategic procurement

Develop a cloud-smart strategy aligned with cloud economics

Developing a cloud-smart strategy requires a holistic approach that tightly integrates cloud architecture with cloud economics to unlock business value. In defining your cloud strategy, it is critical to ask questions such as: Why are you moving to the cloud? What advantages does it bring? How do you ensure you meet your business needs? Once you have answers to these questions, it is important to build an organization-level cloud adoption plan that is technically and financially viable.

- **Analyze the business case for cloud adoption:**
  This should include metrics such as TCO, ROI, net present value (NPV), internal rate of return (IRR), total economic impact (TEI), total economic value (TEV), etc. The cloud business case should consider all necessary IT expenses, including the cost of hardware refresh and elimination of technical debt. In our research, a large majority (64%) of executives have an approved business case for cloud investments. Crucially, cloud business cases must be grounded in realistic assumptions about cloud utilization, based on the current and target state of cloud architecture.

  It is crucial to align On-Demand tech adoption with broader business goals through a well-defined target operating model (TOM). Our research shows that only 45% of executives today say their IT strategy is regularly re-aligned with business strategy.

- **Build an adoption roadmap that details the move-to-cloud in a technically and financially viable manner:**
  Many organizations focus on a "lift and shift" approach of moving workloads into the cloud without modifying them, migrating technical debt and operational inefficiencies as well. Rehosting an unmodified existing application can raise cloud costs by 15%.[49] Since lifted and shifted workloads are not cloud-optimized,

overprovisioning compute resources can be a costly and recurring challenge. Organizations should assess each app: retire, retain, rehost, re-platform, refactor, or repurchase for real cloud value using code-level analysis tools. Anna Kopp, Digital Lead Germany at Microsoft says, *"Some companies still have thousands of shadow applications running on-premises. You have to prioritize what brings value and what to switch off when moving toward cloud, which is a strategic decision. Cloud transformation isn't a Friday afternoon hobby. It needs clear ownership and strategic leadership."*

- **Shift-left for On-demand tech cost management:**
  The early days of cloud journey are the perfect time to bake-in cost discipline. Shifting left in On-Demand tech cost management involves:

  – Integrating cost awareness and optimization efforts early in the development lifecycle, including cost-aware architecture reviews

  – Empowering developers with cost insights

  – Embedding cost-aware practices in a continuous integration/continuous delivery (CI/CD) pipeline

  – Fostering a culture of cost awareness

"Some companies still have thousands of shadow applications running on-premise. You have to prioritize what brings value and what to switch off when moving towards coud, which is a strategic decision. Cloud transformation isn't a Friday afternoon hobby. It needs clear ownership and strategic leadership."

Anna Kopp

Digital Lead Germany
Microsoft

Formulate ecosystem partnership models aligned with business outcomes

- **Explore multi-cloud solutions:** Nine in 10 (91%) executives in our research say their organization currently has a multi-cloud[50] environment for their public cloud. But an earlier study found that 43% of organizations adopted multi-cloud in an ad-hoc, rather than planned, manner.[51] With the right framework, organizations can take full advantage of the cloud at lower cost and with simplified governance, while ensuring elasticity and resilience. Nearly half (49%) of executives in our research say they explore multi-cloud solutions to take advantage of cost differences across providers.

- **Evaluate "gain-share" models to foster shared accountability and to incentivize continuous optimization:** Below, we highlight a few key models to consider. By strategically selecting and combining these models, organizations can drive financial discipline, foster innovation, and ensure mutual value creation with their cloud service providers and their technology partners – ensuring ecosystem-wide alignment to unlock shared incentives, spend optimization, and maximation of value.

  – **Pay-for-savings only:** Clients pay providers a percentage of actual cost savings achieved. This model ensures zero risk for the client and strong alignment of incentives.

  – **Fixed fee and gain-share:** Combining a predictable base fee with a performance-based savings share, this model balances cost certainty with motivation for sustained efficiency.

  – **Tiered gain-share:** Incentives increase as savings thresholds are met, encouraging deeper and ongoing optimization. This model is best for long-term engagements with ambitious cost-reduction goals.

  – **Outcome-based pricing:** Tying pricing to business KPIs (e.g., cost per transaction or user) aligns cloud spend with measurable value.

  – **Pay-as-you-go:** A pay-as-you-go model offers flexibility for variable workloads. However, without optimization, costs can escalate – making it a candidate for layering with gain-share incentives.

- **Maximize savings with smart instance and discount strategies:** Use strategies such as optimizing reserved instances,[52] considering commitment-based pricing,[53] and using spot instances [54] (for predictable or flexible workloads) and vendor negotiations to optimize costs. In our research, only 47% of executives say they regularly review and renegotiate contracts with cloud service providers to ensure cost efficiency.

  Around 66% of technology and finance executives in our research already use **reserved instances** to reduce On-Demand tech costs.

  Cléber Alexandre Agazzi, Head of Infrastructure and IT Operations at Sicredi, a large cooperative financial institution in Brazil, says, *"We use 70% spot instances in our environment. That architecture alone accounts for 60% of our cloud cost savings."* Uber's Michelangelo platform (AI) uses AWS spot instances to train machine learning models, significantly reducing cloud costs by leveraging unused compute capacity at lower prices without compromising performance.[55]

  Build a centralized team for deploying and managing these discount reservations/spending across the organization.

- **Establish joint steering committees** with vendors to review performance, cost trends, and optimization opportunities.

"We use 70% spot instances in our environment. That architecture alone accounts for 60% of our cloud cost savings."

Cléber Alexandre Agazzi

Head of Infrastructure and IT Operations
Sicredi

Equip finance, business, and tech leaders to align on a "shared language of value"

With the rapid adoption of On-Demand tech impacting operating expenses, tech and infrastructure leaders must become more finance-savvy. CIOs/CTOs should be able to articulate On-Demand tech's value in financial terms that resonate with CFOs/COOs/CEOs – thereby emerging as strong "value communicators." CIOs should start with financial fundamentals, understand the key business drivers and shape their On-Demand tech investment strategies aligned with the desired business outcomes.

However, CFOs and business leaders should be able to grasp On-Demand tech's agility, scalability, and cost dynamics to drive smarter decisions that drive real value and avoid cost surprises. They must shift from a capex-heavy mindset to an opex-oriented approach – embracing On-Demand tech not merely as cost centers, but as integral components of the business value stream that enhance agility, accelerate innovation, and deliver measurable outcomes.

Anna Kopp from Microsoft adds, *"FinOps today is short-term focused because shareholder-driven companies prioritize quarterly numbers over long-term strategy. Transformation is not just digital – it's cultural and process-driven. CIOs must engage with finance to align cloud investments with value."*

Architecture: *Build scalable, agile, frugal, and sustainable architecture that aligns cost with value*

Design to cost/value

Architecture is about translating business requirements and constraints into optimal technical requirements. Since cost is an important factor, "optimal" includes best value for the business.

Frédéric Chanfrau, CIO, Head of Technology, RBC Clear, US, adds: *"For CIOs and COOs advancing cloud adoption, begin with primary education and certifications to understand cloud architecture. Develop a long-term strategy by weighing the cost of adoption against the risk of falling behind."* [56]

"FinOps has conventionally been reactive, primarily focused on reducing public cloud spend only after it has spiraled out of control. With the continued rise in On-Demand spending and the growing need to extract more value from fewer resources, FinOps must evolve rapidly to establish a clear link between business value and cost."

Alison McIntyre
Director of Cloud Economics
Capgemini Invent

Our previous research highlights that switching to a green-cloud architecture and framework has delivered 19% cost savings among organizations that have been able to scale the solution organization-wide.[57] **Embedding cost, value, and sustainability as non-functional requirements is required from the start.**

Engineer scalable architecture for efficiency

Design systems with auto-scaling and cost-aware architecture patterns and integrate cost thresholds into architectural decision-making. As **Figure 22** shows, half (52%) of technology executives say they have scalable architectures to optimize resource usage as workloads fluctuate. Airbnb cut $63.5 million in costs by shifting to Kubernetes, automating scaling, optimizing storage, and fostering a cost-aware engineering culture to manage cloud workloads efficiently.[58]

Figure 22.

Half of organizations focus on scalable On-Demand tech architecture

Percentage of executives who agree with the below statements

We have scalable architectures to optimize resource usage as workloads fluctuate — 52%

Our architecture is agile, having variable spending resources and on-demand infrastructure — 50%
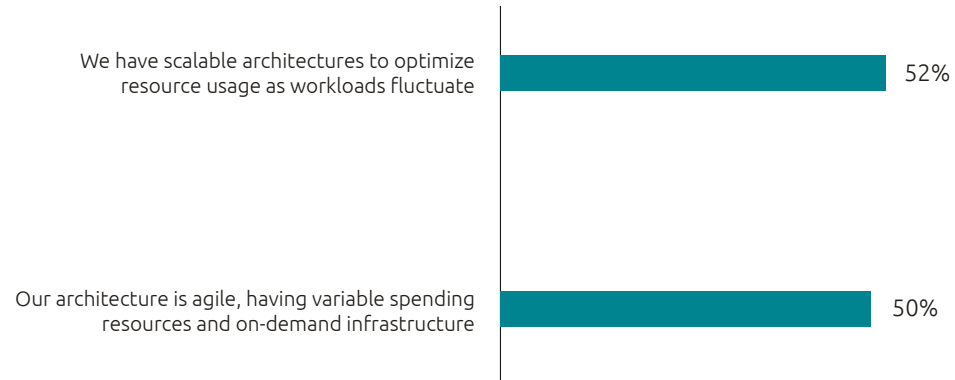
Source: Capgemini Research Institute, Cloud and On-Demand tech economics, May 2025, N=500 executives from technology function.

Use modular architecture

Modular, cloud-native designs – including microservices,[59] containerization,[60] and serverless architecture[61] – enhance agility, reduce technical debt, and optimize costs. The scalability of cloud-native designs enables organizations to allocate resources precisely. In our research, only 46% say they employ cloud-native designs.

A hybrid cloud architecture[62] also helps businesses to optimize spending and lower cost.

Build fit-for-purpose architecture

Cloud architecture must be fit-for-purpose – solving functional and operational business requirements in the most cost-effective way. Batch processing, as an example, typically does not need high availability and can often utilize spot instances[63] or other forms of inexpensive computing, especially if you can easily re-run a job in case of an interruption.[64]

Tonino Greco, Head of Cloud, Infrastructure, and Operations at UK-based fashion brand River Island, told us: *"Architectural changes are critical. We don't just replace systems – we assess the entire ecosystem to avoid technical debt and ensure long-term efficiency."*

Adopt frugal AI architecture

The growing investments in GPUs, data centers, and platforms signal an expected surge in future spending and operational costs to deliver ROI. To manage this, organizations must adopt frugal AI architecture: choosing efficient models, optimizing tech stacks, using techniques like caching, and designing with cost-efficiency at the core.

This architecture must be agile, enabling easy switching between models, adapting to emerging frameworks, and supporting low-cost alternatives.

Implement cost-aware architecture that limits egress charges

While managing cloud and other On-Demand tech costs, many organizations overlook the data transfer and other indirect charges due to poor egress architecture.[65] Organizations can reduce egress costs by minimizing data transfers, for example, by keeping data transfers within the same availability zone; ensuring data transfer within the same region; using caching options; or using edge network accelerators.
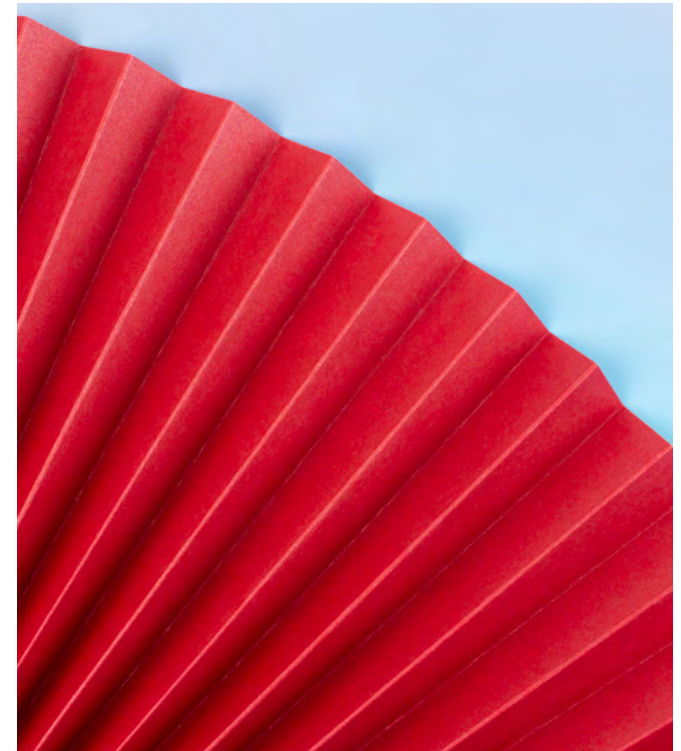
"Architectural changes are critical. We don't just replace systems – we assess the entire ecosystem to avoid technical debt and ensure long-term efficiency."

Tonino Greco

Head of Cloud, Infrastructure, and Operations
River Island

FinOps governance and culture: *Scale smart, govern right, and empower all*

FinOps helps to optimize On-Demand tech spend, improve budgeting and cost control, empower product teams, as well as reduce carbon emissions. To maximize the value of On-Demand technologies, organizations must evolve FinOps from a cost-monitoring function into a strategic enabler.

Start small, showcase quick wins, and evolve cloud FinOps into a strategic capability

Beginner organizations need to:

- Start by developing a minimal viable vision, forming a centralized team and establishing a governance framework

- Have C-suite (CEO, COO or the CFO) championing the FinOps function

- Focus on **quick wins** to build credibility, rally executive buy-in, and create momentum for broader FinOps adoption

- Enable **cross-functional collaboration** between cloud, IT operations, development, business, platform engineering, and finance; foster a culture of shared responsibility

- Equip FinOps teams with **cost management tools** that help forecast spend, organize spend by workload, and eliminate waste

- Define **KPIs to measure success**. In our research, only 45% of executives say their organization has clearly defined KPIs to measure On-Demand tech cost-optimization

Home Depot, the home improvement retailer, built a dedicated cloud cost team in 2022 and identified "tens of millions of dollars" in savings.[66]

Advanced organizations should:

- Evolve FinOps **from a tactical cost-saving function to a core strategic capability**, by embedding it in the TOM

- Focus on **tackling deep, structural inefficiencies and hidden waste**

- Integrate **cost control directly into infrastructure-as-code, CI/CD, and DevOps workflows.** This means all FinOps initiatives should be shared, understood, and deployed from the beginning of application development or cloud-modernization projects.

- Define value through **unit economics**.[67] For example, an insurance organization used unit economics to identify inefficiencies in compute and storage usage for data-intensive tasks such as risk assessment and fraud detection. This led to a 20% reduction in cloud costs per policy.[68]

- Enable **decentralized ownership, with a centralized enablement** through a FinOps center of excellence. Pepsico uses a two-tiered FinOps approach: a centralized governance team responsible for policy, KPI dashboards, and reporting tools enables smaller teams to monitor day-to-day costs across functions.[69]

Expand the scope of FinOps to include SaaS and AI/Gen AI

SaaS expenditures are rapidly approaching the scale of cloud spend. Zylo's SaaS Management Index found that the average organization has 275 SaaS applications,[70] With each application coming with its baseline, premium, and add-on costs, the complexity of contracts and renewals demands a proactive, structured approach. However, as previously seen in **Figure 15**, only four in 10 organizations include SaaS within FinOps scope.

Organizations should start by identifying all SaaS apps in use, centralize SaaS purchasing and governance while maintaining close collaboration with business units. They should use benchmarking tools before buying and negotiate every renewal using market and usage data. Establishing a system of record, integrating with single sign-on (SSO), and mapping spend to business value helps identify underutilized or redundant tools.

When it comes to AI technologies (including Gen AI and agentic AI), without financial governance, costs can quickly spiral. AI services are offered through various commercial models, each with its own pricing structure: per token, per case/instance, per user seat, or subscription. Understanding these models is crucial for effective cost management.

Adopt FinOps practices for AI such as (not an exhaustive list):

- Choosing the right model (pre-trained versus custom; small versus large)

- Choosing the right training type (fine-tuning versus prompt engineering)

- Right-sizing compute (through right-sizing instances, spot instances, multi-instance GPUs, serverless architecture, etc.)

- Model compression (through model pruning,[71] quantization,[72] knowledge distillation,[73] etc.)

- Optimizing data storage (through storage tier strategy)

- Optimizing data transfer (through placing data and compute resources in the same cloud region and using content delivery networks [CDNs]) [74]

- Optimizing inferences (through prompt caching, batching, token optimization, edge computing, etc.)

- Creating a governance framework with cross-functional engagement and accountability

Foster a culture of shared accountability

A clear, business-wide cost framework will help share accountability across functions. Kimberly Floss, Senior Director, Data and AI Project Management at Pepsico, says: *"Costs tied directly to a single source are tagged and go directly into the application team's budget code. They built it, they own it, they pay for it and, when they're done with it, they decommission it."* [75]

A charge-back or show-back model helps to identify the business units responsible for On-Demand tech consumption and fosters accountability. Business units must get involved early, align on budget flows, and educate their teams on the benefits.

Finally, IT cost awareness should be organization-wide. FinOps is everyone's job – from cloud and infrastructure heads to DevOps engineers. Andy Nallappan, President and COO, Cloud Software Group, and former CTO, CSO, and Head of Software Engineering at Broadcom, says: *"The biggest mistake you can make when you move to the cloud is not changing the culture. If you run your operations the same way as when you operated in data centers, the cloud will be three to five times as expensive."* [76]

Organizations must also prioritize employee engagement and capability-building to strengthen cloud cost optimization. The fact that only 38% report dedicated forums or communities of practice for cloud cost strategies indicates a missed opportunity. Developing a change management plan is key.

Bridge the skills gap

Implementing FinOps successfully requires individuals with a combination of financial and technical cloud knowledge. In our research, among organizations that have a dedicated FinOps team, most (61%) say it lacks skilled professionals who understand both finance and cloud complexities.

Organizations should identify and analyze their skills gaps, investing in upskilling and reskilling, hiring dedicated FinOps professionals, and partnering with managed service providers (MSPs) with FinOps expertise. To accelerate progress, organizations should also explore employing AI tools to augment human capabilities, automate insights, and support smarter, faster FinOps decision-making.

"Different parts of IT often operate with distinct priorities — CIOs focus on cost efficiency, while engineering prioritize speed and innovation and operations teams are goaled on stability and resilience. These differing goals can create misalignment when managing cloud resources and overall IT spending. Embedding a FinOps culture across all levels of the organization ensures a shared understanding of financial accountability, enabling better collaboration and more informed decision-making. Driving this cultural shift—both at the grassroots and leadership levels—requires innovative and adaptive adoption models."

Vikram Rajan

Vice President and Global Leader,
Cloud and Infrastructure Advisory
Capgemini

FinOps processes:
Eliminate waste, enforce controls and optimize resources

Organizations must focus on implementing processes and standards that consider the variability of On-Demand tech spending. Below, we highlight a few recommendations:

Figure 23.

About half of organizations are tagging, right-sizing, or scheduling On-Demand tech workloads to reduce costs

Percentage of executives who agree with the below statements
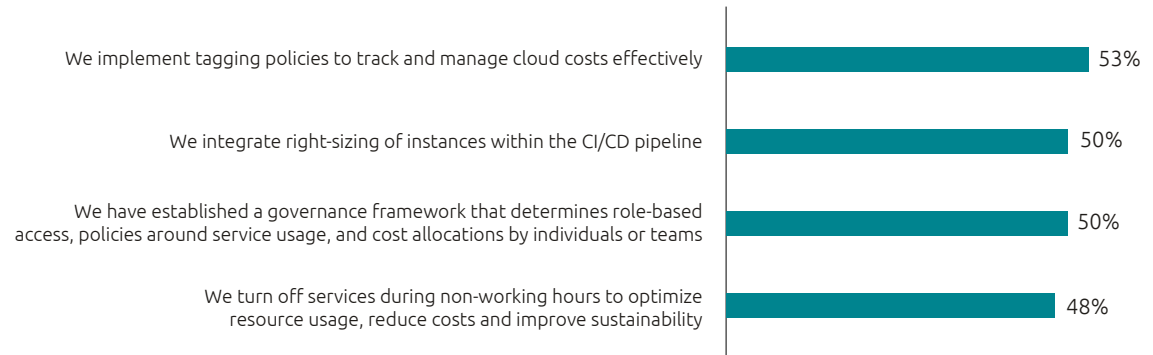
Statement	%
We implement tagging policies to track and manage cloud costs effectively	53%
We integrate right-sizing of instances within the CI/CD pipeline	50%
We have established a governance framework that determines role-based access, policies around service usage, and cost allocations by individuals or teams	50%
We turn off services during non-working hours to optimize resource usage, reduce costs and improve sustainability	48%

Source: Capgemini Research Institute, Cloud and On-Demand tech economics, May 2025, N = 750 executives from technology and finance functions.

Delete idle resources through a robust tagging process

Removing idle and unused resources is a way to make immediate savings without jeopardizing application performance. Tags created to address wasted resources also improve capacity and usage analysis.[77]

Netflix developed an automated tagging strategy to control costs and gain better visibility into their cloud resource usage. During peak traffic, this enabled engineers to reassign workloads efficiently, resulting in a 30% reduction in Amazon Elastic Compute Cloud (EC2) spending.[78]

In our research, only half (53%) of technology and finance executives say they implement tagging policies to track and manage cloud costs effectively (see **Figure 23**).

Right-size overprovisioned instances and choose the right storage type

Using larger instances that don't match your workload requirements creates cost overruns. But it is critical that right-sizing over-allocated resources should follow proper data analysis. For example, AWS recommends right-sizing instances whose CPU and memory usage falls below 40% for a period of four weeks.[79] Only 50% of technology and finance executives in our research integrate right-sizing of instances within the CI/CD pipeline (see **Figure 23**).

Set usage limits with role-based access

Role-based access control improves cost management by restricting access to services, reducing overspending risk, and enhancing cost allocation accuracy. Only 50% of executives in our research say their organization has established a governance framework that determines role-based access, policies around service usage, and cost allocations by individuals or teams (see **Figure 23**).

Schedule resources to deactivate when not required

Efficient scheduling can reduce costs as well as carbon by automatically deactivating cloud resources during periods of inactivity.

Carlsberg Group, a Danish brewer, implemented FinOps practices such as workload optimization and resource "snoozing" for its SAP development and sandbox environments. By using snoozing, the company does not incur costs during nights and weekends when resources are not in use. Carlsberg reports annual savings of 7–10% through the Azure Hybrid Benefit and policies such as snoozing.[80]

Only 48% of technology and finance executives in our research say they turn off services during non-working hours (see **Figure 23**).

Tools: *Fuel FinOps excellence with tools, automation, and AI*

Use tools and automation to optimize costs

Leaders need to **determine whether a third-party cost management tool, native services, or custom-developed methods and scripts are required.** These tools offer real-time visibility, automate budget enforcement, and optimize usage across cloud, SaaS, and Gen AI workloads. They also assist in revising strategy and process established during setup. The landscape of cloud cost management tools is very complex – as per our estimates, there are more than 150 tools in the market today. Organizations must collaborate across the ecosystem to choose effectively from an array of available tools.

Organizations should **automate cost controls using tools for auto-scaling, spot instances, and idle resource detection**. Razer, a gaming lifestyle brand, saved up to 90% in costs by using auto-scaling tools and advanced processors to manage resource scaling during peak hours,

avoiding underutilized capacity. This approach provided significant price-performance benefits across various workloads.[81]
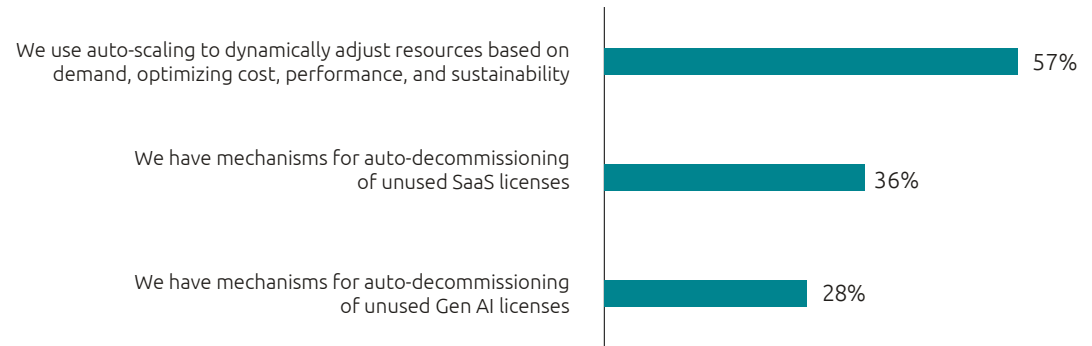
As **Figure 24** shows, more organizations (57%) are already using automation to control cloud consumption. SaaS and

Gen AI management tools can manage, optimize, govern, secure spending, and achieve greater visibility and accountability. However, as the figure shows, automation is nascent in these areas.

Figure 24.

One in three organizations has automation mechanisms for scaling or decommissioning SaaS and Gen AI

Percentage of executives who agree with the below statements

Statement	Percentage
We use auto-scaling to dynamically adjust resources based on demand, optimizing cost, performance, and sustainability	57%
We have mechanisms for auto-decommissioning of unused SaaS licenses	36%
We have mechanisms for auto-decommissioning of unused Gen AI licenses	28%

Source: Capgemini Research Institute, Cloud and On-Demand tech economics, May 2025, N=750 executives from technology and finance functions.

It is also essential to invest in enhancing or upgrading cloud cost management tools. However, our research shows that only 35% of organizations are doing this.

Harness the power of AI/Gen AI for FinOps

The use of AI/Gen AI in FinOps is at an early stage. Only 10% of executives from technology and finance functions in our research say they are using AI/Gen AI extensively for On-Demand tech cost monitoring and optimization. And 34% say they use it moderately. It should be noted that many commercially available cost management tools come with AI functionalities – which might be driving the responses of the executives here.

But among organizations using AI/Gen AI for FinOps (to whatever extent), top use cases include (see **Figure 25**):

- **Intelligent forecasting:** To analyze historical cloud usage, patterns, market trends, and business projections

- **Spend attribution and reporting:** To analyze unstructured data such as invoices and billing reports; categorize spend cost allocations; generate highly customized reports; and contribute to better visibility and accountability

- **AI-driven risk management:** To analyze data patterns, risk indicators, and usage metrics to identify potential compliance violations, security vulnerabilities, or unauthorized spending

- **Automated cost optimization:** To provide intelligent recommendations; identify waste and underutilization; and implement reserved instances[82] or saving plans, or even architectural changes

- **Anomaly detection:** To detect unusual spikes and unexpected changes in spend; identify potential overspending, billing errors, etc.

For example, Spotify uses AI and ML to analyze cloud usage, predict demand, and optimize workloads – cutting costs through smarter resource allocation, workload scheduling, and eliminating inefficiencies in real time.[83]

Pathik Sharma from Google adds, *"AI can transform FinOps – from anomaly detection and financial forecasting to democratizing insights via natural language prompts. Imagine an engineer seeing cost impact at the pull request stage – that's proactive FinOps."*
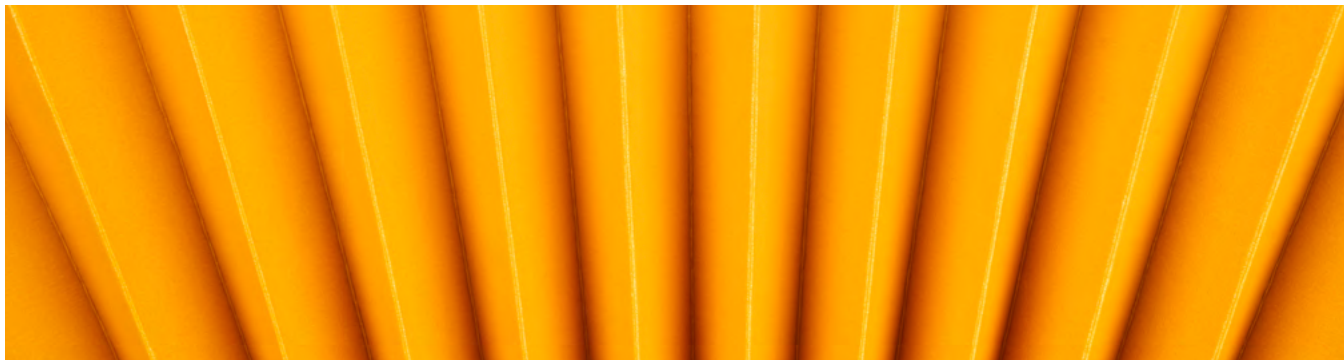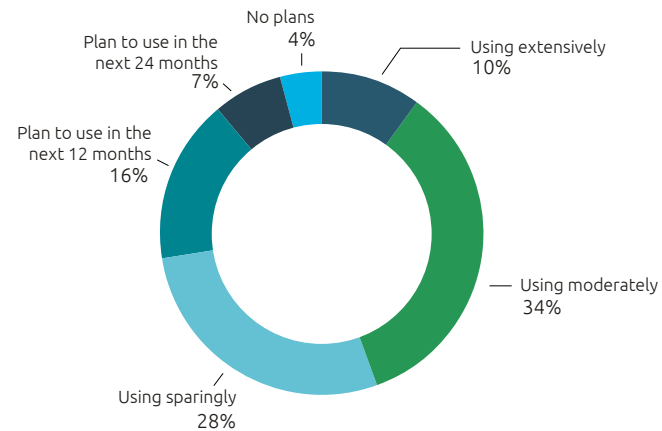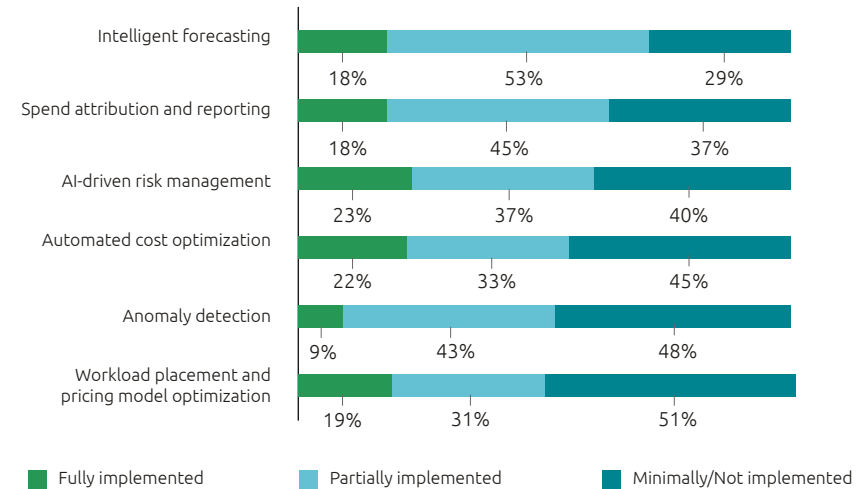
Figure 25.

Intelligent forecasting, spend attribution, and risk management are the top use cases where organizations are using AI/Gen AI in FinOps

Percentage of executives responding to the question: "Do you use AI/Gen AI for On-Demand tech cost monitoring and optimization?"

No plans 4%

Plan to use in the next 24 months 7%

Plan to use in the next 12 months 16%

Using sparingly 28%

Using extensively 10%

Using moderately 34%

Top use cases of AI/Gen AI for FinOps

	Fully implemented	Partially implemented	Minimally/Not implemented
Intelligent forecasting	18%	53%	29%
Spend attribution and reporting	18%	45%	37%
AI-driven risk management	23%	37%	40%
Automated cost optimization	22%	33%	45%
Anomaly detection	9%	43%	48%
Workload placement and pricing model optimization	19%	31%	51%

Source: Capgemini Research Institute, Cloud and On-Demand tech economics, May 2025, N = 750 executives from technology and finance functions.

Source: Capgemini Research Institute, Cloud and On-Demand tech economics, May 2025, N = 334 executives from technology and finance functions who say they are extensively/moderately using AI/Gen AI for cloud and On-Demand technologies cost monitoring and optimization.

To move forward with AI/Gen AI for FinOps, organizations should: [84]

- Build a strong FinOps culture, ensuring alignment of teams across IT, finance, and the business

- Recruit and develop skilled talent with both FinOps and AI/ML understanding

- Identify high-impact use cases where AI/Gen AI can have a tangible impact

- Start small and drive iterative implementation efforts, demonstrating value

- Define and then measure clear metrics for success

- Ensure continuous leadership and oversight

- Source and employ diverse and high-quality datasets

"Integrating AI/Gen AI into cloud FinOps practices unlocks smarter forecasting, proactive anomaly detection, and automated cost savings. A focus on data quality, explainability, and continuous learning will be essential to ensure optimal implementation and a better understanding of ROI."

Anne-Laure Thibaud

Executive Vice President,
AI First Business & Analytics Global Practice
Capgemini

Sustainable FinOps:
Reduce costs and emissions

The future of IT belongs to organizations that embrace both FinOps and GreenOps principles. By implementing sustainable cloud strategies, they can optimize costs, enhance brand image, and contribute to a greener future. We see many examples emerging:

- Through sustainable FinOps implementation, Capgemini helped a luxury fashion retailer to identify 30% savings in their Azure spend, with more than $1.3 million of overall savings in 2023. By optimizing storage and CPU environment, it saved 3,048 tCO_2.[85]

- Bodø kommune, a municipality in Norway, used sustainable FinOps to relocate virtual machines to hydropower-driven regions, cutting 4.59 tonnes of CO_2 emissions from cloud usage and reducing its Azure costs by 50%.[86]

- Netflix, a heavy user of AWS, implemented a predictive scaling mechanism to adjust its cloud resources based on demand forecasts. This not only optimized costs but also reduced unnecessary energy consumption.[87]

Below, we propose a three-step approach to sustainable FinOps:[88]

Inform

Collect On-Demand tech usage data

Analyze resource efficiency

Understand energy usage and CO_2 emissions

Identify opportunities to optimize usage and reduce CO_2 emissions

Optimize

Establish and implement strategy, governance, optimization levers, and processes

Design and build sustainable FinOps tooling platform and dashboards

Conduct employee awareness and training programs

Operate

Fine-tune cloud resource usage with right-sizing, autoscaling, scheduling, etc.

Minimize and optimize data transfer

Conduct monitoring and reporting

Drive continuous improvement proactively

Extend sustainable FinOps practices by connecting new cloud landscapes and adding new capabilities

"Sustainable FinOps is still very nascent. A cultural shift within IT is required to align business goals and ESG ambitions. Begin establishing the basics through sustainable, well-architected platforms and begin to experiment with tools in the market that fit your unique infrastructure and ESG needs." [89]

James Dunn

Global Cloud Portfolio Lead
Capgemini

Conclusion

The shift to On-Demand technologies has redefined how organizations innovate, scale, and compete. But while these technologies promise scalability and speed, they also introduce new layers of complexity, cost unpredictability, and governance challenges.

Many organizations have discovered that adoption is just the beginning. Strategic alignment, proactive cost planning, and empowered FinOps practices, are essential to extracting full value from these investments. Fragmented ownership, unmanaged tech sprawl, and underutilized tools are holding back the digital transformation.

To lead effectively in this new era, technology and finance leaders must move beyond reactive cost control.

They must embed cost intelligence into strategy and architecture; empower teams with automation and AI; and foster a culture of accountability and collaboration. Sustainable FinOps offers a path to optimize spend and align with broader environmental and operational goals. On-Demand tech economics is all about fundamentals: It's not just DevOps or FinOps. It's BizOps. It's a cultural shift. It's a leadership commitment.

Ultimately, the question is no longer whether you are investing in On-Demand tech, but whether you are truly in control of its value.
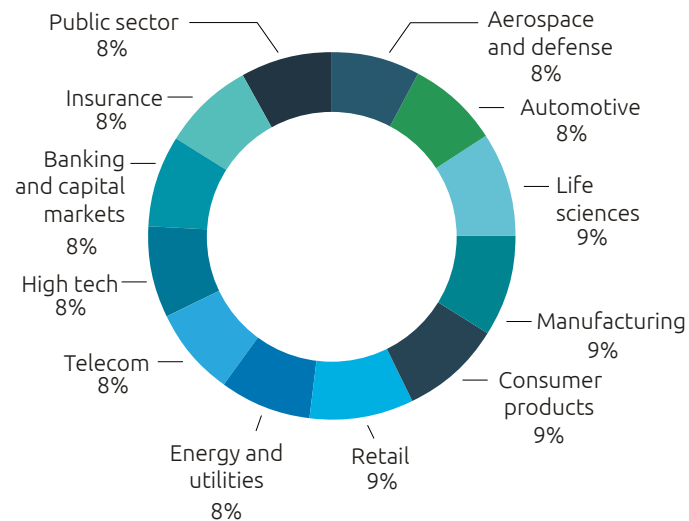
Research methodology

We surveyed 1,000 executives from global organizations with annual revenue over $1 billion who are consuming On-Demand technologies (public cloud, SaaS, and Gen AI on cloud) in 12 sectors and 14 countries across North America, Europe, and Asia-Pacific. We conducted the global survey in May 2025.

To supplement the survey findings, we also conducted in-depth discussions with 11 executives from organizations globally. The study findings reflect the views of the respondents to our online questionnaire for this research and are intended to provide directional guidance. Please contact one of the Capgemini experts listed at the end of the report to discuss specific implications.
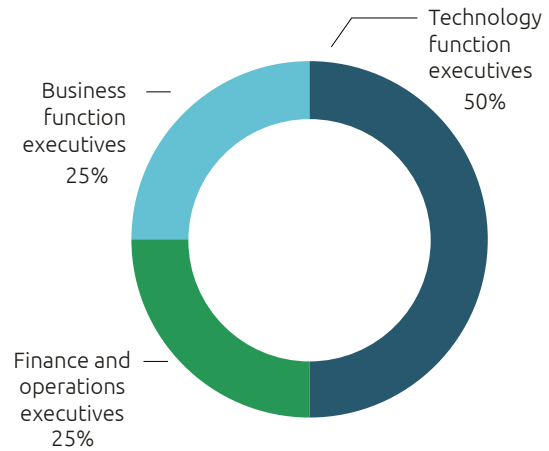
Organizations by country

- Japan 6%
- Singapore 5%
- Australia 12%
- India 7%
- Sweden 4%
- Netherlands 4%
- Spain 4%
- Italy 4%
- Germany 6%
- France 6%
- UK 10%
- Brazil 5%
- Canada 5%
- US 22%
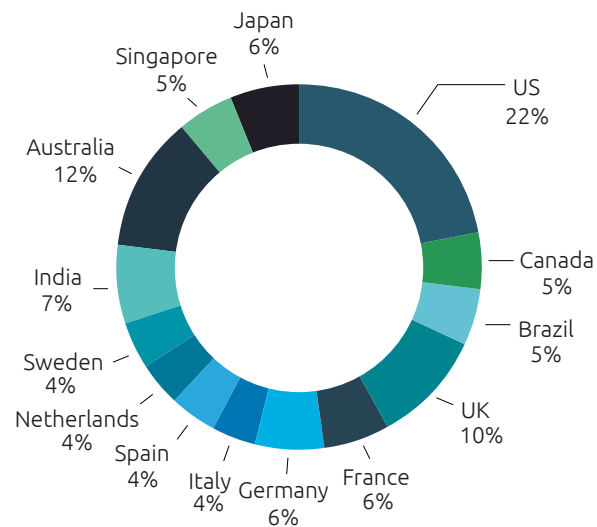
Organizations by sector

- Public sector 8%
- Insurance 8%
- Banking and capital markets 8%
- High tech 8%
- Telecom 8%
- Energy and utilities 8%
- Retail 9%
- Consumer products 9%
- Manufacturing 9%
- Life sciences 9%
- Automotive 8%
- Aerospace and defense 8%

Organizations by annual revenue

- $50bn+ 9%
- $20bn–$50 17%
- $10bn–$20 19%
- $5bn–$10 19%
- $1bn–$5 36%
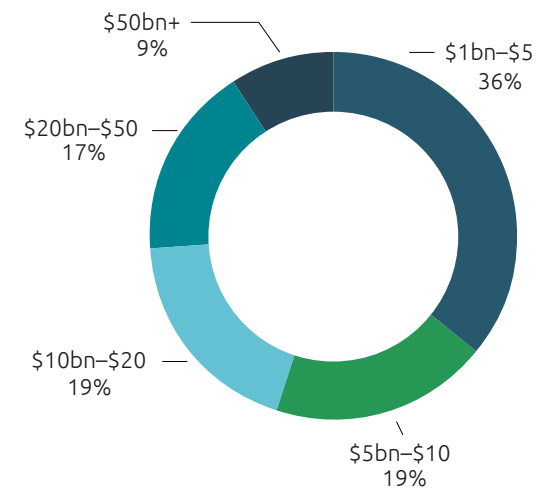
Respondents by current job title

C-level
5%

Head of
department
11%

Vice
president
24%

Director
60%
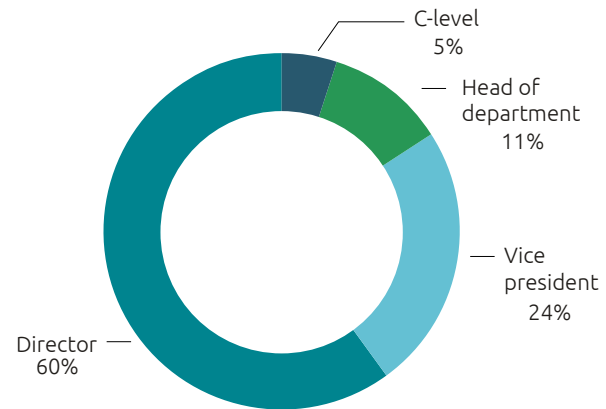
Respondents by function

Technology
function
executives
50%

Business
function
executives
25%

Finance and
operations
executives
25%

Source: Capgemini Research Institute, Cloud and On-Demand tech economics, May 2025, N = 1,000 executives.

References

1. IDC, "The era of FinOps: Focus is shifting from cloud features to cloud value," February 2023.

2. Gartner, "Gartner says cloud will become a business necessity by 2028," November 2023.

3. Gartner, "Gartner forecasts worldwide public cloud end-user spending to total $723 billion in 2025," November 2024.

4. Flexera, State of the cloud report, March 2025.

5. Gartner, Gartner forecasts worldwide public cloud.

6. Gartner, "Forecast: Public Cloud Services, Worldwide, 2023-2029, 2Q25 Update," June 2025.

7. Goldman Sachs, "Cloud revenues poised to reach $2 trillion by 2030 amid AI rollout," September 2024.

8. Containerization is a software deployment process that groups and isolates an application's code, libraries, and files into a single lightweight container to ensure consistency across different environments. It is essential for managing and deploying APIs.

9. TM Forum, "Today 80% of Deutsche Telekom's IT systems are hosted on cloud infrastructure," February 2025.

10. Microsoft, "Microsoft Azure meets growing needs of healthcare organizations, such as Mount Sinai Health System, with a highly scalable public cloud for Epic," August 2023.

11. Microsoft, "How real-world businesses are transforming with AI," April 2025.

12. BBVA, "BBVA migrates customer care services to the cloud to improve service," June 2022.

13. Relevance AI, "Error pattern detection AI agents," accessed April 2025.

14. Forbes, "Sustainability and SaaS: 3 ways to use tech to promote sustainability," May 2023.

15. Workspace.Google, "Pricing updates and more flexible payment options for Google Workspace," February 2023.

16. Cloudsight.Zendesk, "Microsoft 365 fee increase," January 2025.

17. PRnewswire, "Cost of software crisis: SaaS inflation running at more than double US consumer inflation," November 2023.

18. Cloud arbitrage refers to shifting workloads to cloud providers offering better pricing or performance for specific use cases; TechTarget, "When should my company use cloud arbitrage?" July 2020.

19. Gartner, 2024 Cloud spending: IT Balances costs with GenAI innovations, January 2024.

20. Jevons Paradox is an economic theory stating that efficiency gains in resource use can lead to higher overall consumption of that resource, rather than conservation; Mint, "What is Jevons Paradox? Satya Nadella explains, 'As AI becomes a commodity, its energy demands will skyrocket!'" January 2025.

21. CIO Dive, "CIOs grapple with cloud cost paradox," April 2025.

22. Information Week, "AI's impact on cloud spending: The hunger for capacity," March 2025.

23. Gartner, "Gartner unveils top eight cybersecurity predictions for 2023-2024," March 2023.

24. Capgemini Research Institute, Gen AI at work: Shaping the future of organizations, October 2024.

25. Forbes, "The benefits and drawbacks of shadow IT," May 2024.

26. 1Password, "Charlie Livingston talks shadow IT: 'We need to change to a collaborative model,' " May 2024.

27. Forbes, "Cloud ROI: Managing expectations and finding efficiencies," February 2024.

28. IDC, "IDC forecasts worldwide sovereign cloud spending to reach more than $250 billion in 2027," December 2023.

29. Capgemini Research Institute, "The journey to cloud sovereignty," July 2022.

30. Data localization entails keeping the data within a preferred jurisdiction, irrespective of whether the public or private cloud provider is of local origin

31. Capgemini, "Capgemini and Orange are pleased to announce the launch of commercial activities of Bleu, their future 'cloud de confiance' platform," January 2024.

32. FinOps Foundation, "FinOps X Europe 2024 – Day 1 Keynote," November 2024.

33. FinOps Foundation, January 2025.

34. Capgemini and IDC, "Rise of FinOps: Evolving beyond a discipline into a mindset, strategy, culture, and means to financial excellence," February 2023.

35. Capgemini, "Prepare your cloud for the future of FinOps: How to achieve a more efficient, cost-effective, and sustainable cloud," April 2023..

36. FinOps Foundation, "2024 changes to the definition of cloud FinOps," February 2024.

37. FinOps Foundation, "FinOps for AI Overview," accessed June 2025.

38. Unit economics in FinOps is a way of measuring the revenue and cost of cloud spending in relation to any one quantifiable unit that brings value to the business. It helps FinOps teams to correlate cloud spending growth to overall business growth, and to optimize profit based on objective metrics. Source: Flexera, "How FinOps and ITAM teams can work together in real-life scenarios," January 2025.

39. Flexera, "State of the cloud report," March 2025.

40. CIO.com, "FinOps breaks out of the cloud," March 2025.

41. CIO.com, "What is FinOps? Your guide to cloud cost management," December 2022.

42. CIO.com, "FinOps breaks out of the cloud," March 2025.

43. Global Market Insights, "Cloud cost management tools market size - By component, by deployment model, by organization size, by cloud, by end user, analysis, share, growth forecast, 2025 – 2034," November 2024.

44. World Bank, "Measuring the emissions and energy footprint of the ICT sector: Implications for climate action," March 2024.

45. The MIT Press Reader, "The staggering ecological impacts of computation and the cloud," February 2022.

46. Goldman Sachs, "AI to drive 165% increase in data center power demand by 2030," February 2025.

47. Serverless architecture is a cloud-native execution model that provides On-Demand computing resources and removes operating and managing overheads.

48. Nasscom, "FinOps gets sustainable edge with GreenOps," August 2024.

49. CloudZero, "What is lift and shift? Is it right for you?," October 2023.

50. Multi-cloud is the practice of using the services of multiple cloud providers to optimize workload performance, increase flexibility, and mitigate the risks of relying on any one vendor.

51. Capgemini, "How to make multi-cloud an enterprise game changer," November 2020.

52. Reserved instances (RIs) are prepaid compute instances that offer significant pricing discounts.

53. Commitment-based pricing is a pricing model where customers commit to using a certain amount of cloud resources or spending a specific amount for a predefined period (typically one or three years) in exchange for discounted prices compared with on-demand pricing.

54. Spot instances offer deep pricing discounts for cloud servers with the caveat that the servers may shut down at any time.

55. Cloudzero, "Use spot instances for AI workloads," March 2025.

56. Capgemini, World Report Series 2025: Cloud in Financial Services, November 2024.

57. Capgemini Research Institute, "Sustainable IT: Why it's time for a Green revolution for your organization's IT," May 2021.

58. The Stack, "Airbnb's AWS costs were getting out of hand. Here's how it tackled them," May 2022.

59. Microservices are small independent software components that collectively perform as a complete cloud-native software.

60. Containerization is a software deployment process that bundles an application's code with all the files and libraries it needs to run on any infrastructure

61. Serverless architecture is a software design approach where developers can build and manage applications without managing the underlying architecture. Serverless applications still run on servers, but the cloud service provider is responsible for provisioning, managing, and scaling all the cloud infrastructure.

62. Hybrid cloud architecture combines on-premises, private cloud, public cloud and edge settings to create a single, flexible managed IT infrastructure.

63. Spot instances offer deep pricing discounts for cloud servers with the caveat that the servers may shut down at any time.

64. Capgemini, Impact of planning and architecture on cloud economics, January 2022.

65. Egress costs are the network fees that cloud providers charge organizations to move their data out of the cloud.

66. The Wall Street Journal, "Amazon warns of weaker cloud sales as businesses cut spending," April 2023.

67. Unit economics refers to the process of calculating the cost of a single unit of a business that can show the business value of the cloud.

68. FinOut, "Unlock cloud cost efficiency with Finout's unit economics widget," October 2024.

69. CIO Dive, "How PepsiCo capped cloud overspend," September 2024.

70. Zylo, "2025 SaaS Management Index," Jan 2025.

71. Model pruning is removing unimportant parameters from a neural network model to reduce the model size.

72. Quantization is a model size reduction technique that converts model weights from high-precision floating-point representation (32-bit FP) to low-precision FP or integer representations (8-bit).

73. Knowledge distillation is the process of transferring knowledge from a large model to a smaller one.

74. CDNs are geographically distributed group of servers that caches content close to end users.

75. Capgemini, FinOps: Building the pillars for success, June 2023.

76. CIO.com, "Broadcom's Andy Nallappan on what cloud success really looks like," May 2023.

77. AWS, "How to optimize your AWS workload cost with Capgemini and Virtana," March 2019.

78. Substack, "AWS case study: Netflix's automated tagging strategy for cost optimization," October 2024.

79. Economize, "How to reduce cloud waste: 9 effective strategies for 2025," December 2024.

80. Microsoft, "Carlsberg Group increases cloud efficiency with FinOps and Azure solutions," October 2024.

81. Economize, "How 6 companies saved up to 80% cloud costs," December 2024.

82. Reserved instances (RIs) are prepaid compute instances that offer significant pricing discounts.

83. Ziffity, "Cloud cost control: AI & ML powerhouse for savings," May 2024.

84. Capgemini, "Want maximum value from FinOps in the cloud? Now's the time to leverage generative AI," July 2024.

85. Capgemini, "Reduce cloud costs and CO2 with Sustainable FinOps," July 2024.

86. Crayon, "Sustainable IT practices in 2025: Cost management for environmental and economic benefits," February 2025.

87. Quinnox, "Revolutionizing cloud cost management: A sustainable approach with GreenOps," July 2024.

88. Capgemini, Reduce cloud costs and CO2 with Sustainable FinOps.

89. Capgemini, "In cloud operations, agility and adaptability make the difference," May 2024.

Authors

Meet the experts

Arnaud Balssa
EVP, Global Business Technology leader, Capgemini
arnaud.balssa@capgemini.com

Arnaud apart from being the Head of Microsoft Invent Partnerships, also leads global digital transformation programs across manufacturing, energy, and finance. He previously headed France's Future of Technology entity and launched Inventive IT—a dual transformation approach for IT departments that enhances both operating and IT models to deliver business results. His work spans global clients like Michelin, Engie, and Société Générale.

Jez Back
Cloud Economist and Global Offer Leader for Cloud Consumption On-Demand, Capgemini
jez.back@capgemini.com

Jez is a Subject Matter Expert in On-Demand technology economics and FinOps with over 15 years of industry experience in this discipline. He previously owned a FinOps consultancy business, delivered transformation and advisory services at Deloitte, HP ES and EDS. He is the first Certified FinOps Professional globally and has featured in documentaries, TV and podcasts as well as delivering keynotes on FinOps and Technology Economics.

Thomas Sarrazin
Global FinOps Offer Lead, Capgemini
thomas.sarrazin@capgemini.com

For the past three years, Thomas has served as the Global FinOps Leader at Capgemini, where he has defined and implemented the FinOps strategy within the delivery teams. Prior to this role, Thomas held various positions as a technical leader and enterprise architect. He has an extensive experience in IT transformation programs, with a specialization in cloud-related topics.

Alison McIntyre
Director of Cloud Economics, Capgemini
alison.mcintyre@capgemini.com

Alison McIntyre is the Director of Cloud Economics in Capgemini Invent. Her function is currently the largest and most qualified FinOps and Cloud Economics consultancy team within the UK. She previously worked in industry, mostly Financial Services and she was the first Ambassador for the FinOps Foundation in the UK.

Authors

Meet the experts

Brian Hammond
Cloud Group Offer Leader for Capgemini;
Sogeti NA Portfolio Leader
brian.hammond@capgemini.com

Brian leads globally the Capgemini Cloud Group
Offer and Partnerships/Portfolio for Sogeti
in North America. He works with and advises
Capgemini clients on how they can leverage Cloud
and the Hyperscalers to rapidly achieve business
outcomes like enterprise agility, transformation,
innovation, sustainability and more.

Brian has over 20 years of experience in the
industry, the last 8 focusing on Cloud. Brian is a
regular speaker at events like Amazon Reinvent
and Microsoft Inspire.

Vikram Rajan
Vice President and Global Leader,
Cloud and Infrastructure Advisory,
Capgemini
vikram.rajan@capgemini.com

Vikram has over two decades of experience
in shaping technology strategy and
complex technology transformations for
clients across the globe. Vikram has been
helping clients navigate the ever-evolving
paradigms of cloud computing since its
early days of adoption.

Anne-Laure Thibaud
Executive Vice President – Head of AI First
Business & Analytics Global Practice, Capgemini
annelaure.thibaud@capgemini.com

Anne-Laure Thibaud leads a worldwide team
accelerating the adoption of Generative
and Agentic AI, helping organizations
unlock business value through AI-driven
transformation. With a focus on designing
and operating high-impact solutions in
collaboration with key technology partners,
Anne-Laure champions new ways of working
where human and AI agents collaborate to drive
meaningful and sustainable outcomes at scale.

Authors

Meet the Capgemini Research Institute

Jerome Buvat
Head,
Capgemini Research Institute
jerome.buvat@capgemini.com

Jerome is the head of the
Capgemini Research Institute. He
works closely with industry leaders
and academics to help organizations
understand the business impact of
emerging technologies.

Subrahmanyam KVJ
Senior Director,
Capgemini Research Institute
subrahmanyam.kvj@capgemini.com

Subrahmanyam is a Senior Director
at the Capgemini Research Institute.
He loves exploring the impact
of technology on business and
consumer behavior across industries
in a world being eaten by software.

Nancy Manchanda
Director,
Capgemini Research Institute
nancy.manchanda@capgemini.com

Nancy has over 16 years
of experience in market
research, consulting and digital
transformation. She collaborates
with business leaders to understand
challenges and opportunities
relating to technology, people and
sustainability transformation.

Siva Chidambaram Sathyanandan
Manager,
Capgemini Research Institute
siva.chidambaram-s@capgemini.com

Siva is a seasoned manager carrying
experience collaborating with industry
leaders to drive modern tech-enabled
solutions driven by data. He specializes
in strategic research, advisory, and
improving business operations
through business intelligence and
thought leadership that aids in staying
ahead amongst dynamic landscapes.

About the Capgemini Research Institute

The Capgemini Research Institute is Capgemini's in-house think tank on all things digital. The Institute publishes research on the impact of digital technologies on large traditional businesses. The team draws on the worldwide network of Capgemini experts and works closely with academic and technology partners. The Institute has dedicated research centers in India, Singapore, the United Kingdom, and the United States. The Institute was ranked #1 in the world for the quality of its research by independent analysts for six consecutive times - an industry first.

Visit us at www.capgemini.com/researchinstitute/
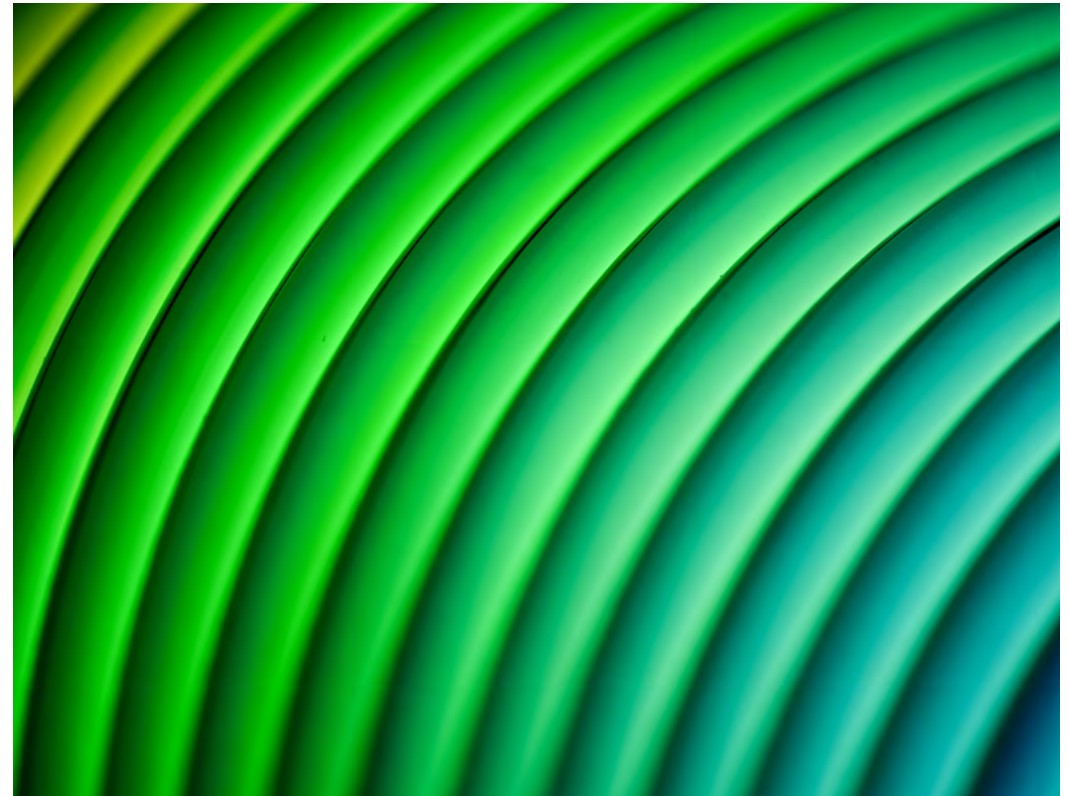
Why partner with Capgemini?

In today's on-demand world, technology consumption has moved beyond fixed models and static budgets. Enterprises now need cloud strategies that are as dynamic as the business landscape they operate in—flexible, insight-driven, and relentlessly efficient.

Capgemini's Cloud Consumption On-Demand (CCOD) offering is built for this new reality.

Going beyond traditional FinOps, CCOD is a strategic framework that enables enterprises to take control of their digital consumption, not just to cut costs, but to unlock performance, agility, and future-readiness.

Our two-phase, outcome-led approach enables organizations to:

- Assess current on-demand technology usage

- Identify quick wins and build a compelling business case for optimization

- Reduce costs through data-driven insights and governance

- Future-proof operations with frugal architecture and continuous efficiency

Whether navigating cloud, SaaS, or Gen AI ecosystems, CCOD helps you maximize the value of every click—translating insight into actions and spend into strategic advantage.

The future of FinOps isn't just about financial control, it's about unlocking smarter, leaner, and more adaptive digital operations.

Capgemini is your trusted partner in this journey, delivering transparency, efficiency, and sustained value.

Explore now: https://www.capgemini.com/in-en/solutions/cloud-consumption-on-demand/

Follow us on LinkedIn: Capgemini Cloud

For more information, please contact:

Global

James Dunn
Global Cloud Portfolio Lead, Capgemini
james.dunn@capgemini.com

Thomas Sarrazin
Global FinOps Offer Leader,
Cloud Infrastructure Services
thomas.sarrazin@capgemini.com

Jez Back
Cloud Economist & Global Offer Leader,
Capgemini Invent
jez.back@capgemini.com

Arnaud Balssa
Global Head of Business Technology,
Capgemini Invent
arnaud.balssa@capgemini.com

Regional

North America
Prerna Bhardwaj Pant
Managing Enterprise Architect
prerna.a.pant@capgemini.com

Pieter Pottie
Senior Director, Head of CIO Advisory
pieter.pottie@capgemini.com

UK
Ewan MacLeod
Cloud Enterprise Architect
ewan.macleod@capgemini.com

Jez Back
Cloud Economist & Global Offer Leader,
Capgemini Invent
jez.back@capgemini.com

Italy
Andrea Cenciarelli
Cloud Architect
andrea.cenciarelli@capgemini.com

Giovanni Vendramel
Principal Consultant,
Head of Cloud Transformation
giovanni.vendramel@capgemini.com

Carlo Pisani
Managing Consultant
carlo.pisani@capgemini.com

France
Zakia Queiroz
Manager FinOps Offer
zakia.queiroz@capgemini.com

Benjamin Piger
Director, Strategy and Transformation
benjamin.piger@capgemini.com

Germany
Lukas Kuckelkorn
Sustainable Cloud Economics Lead
lukas.kuckelkorn@capgemini.com

Holger Kuprian
Senior Manager, Business Technology
holger.kuprian@capgemini.com

Spain
Llorens Vanaclocha
South & Central Europe Cloud CoE Advisor
llorens.vanaclocha@capgemini.com

Jose Manuel Morales
Principal, Business Technology
jose-manuel.morales-menendez@capgemini.com

Netherlands
Jurjen Thie
NL Cloud COE, Principal Presales Solutioner
jurjen.thie@capgemini.com

Simon Van den Doel
Principal Consultant
simon.vanden.doel@capgemini.com

Bart Turk
Manager, Cloud Lead
bart.turk@capgemini.com

India
Jagdish Venkatesh
Cloud Delivery Lead
jagdish.venkatesh@capgemini.com

Balaji Venugopal
Senior Manager, Business Technology
balaji.venugopal@capgemini.com

APAC
Edouard Laroche- Joubert
Head of Cloud CoE
edouard.laroche-joubert@capgemini.com

Subbu Gandrala
Cloud Consumption on Demand Offer Lead
subbu.gandrala@capgemini.com

Nordics
Younes Bentahar
Head of Nordics Cloud Service Line
younes.bentahar@capgemini.com

Andreas Dencker
Director, Business Technology
andreas.dencker@capgemini.com

Oystein Gjerde
Manager
oystein.gjerde@capgemini.com

More Capgemini Research Institute publications

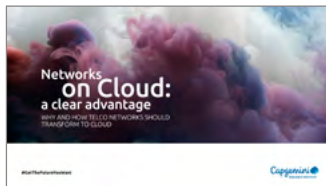**Rise of agentic AI: How trust is the key to human-AI collaboration**

Top Tech Trends of 2025: AI-powered everything

Cloud sovereignty: The road ahead

Future encrypted: Why post-quantum cryptography tops the new cybersecurity agenda

Networks on cloud: A clear advantage

Developing sustainable Gen AI

Subscribe to latest research
from the Capgemini Research Institute

Receive copies of our reports by scanning
the QR code or visiting
https://www.capgemini.com/capgemini-research-institute-subscription/

Capgemini Research Institute

Fields marked with an • are required

First Name •

Last Name •

Email •

By submitting this form, I understand that my data will be processed by Capgemini as indicated above and described in the Terms of use.

Submit

Notes

About Capgemini

Capgemini is a global business and technology transformation partner, helping organisations to accelerate their dual transition to a digital and sustainable world, while creating tangible impact for enterprises and society. It is a responsible and diverse group of 350,000 team members in more than 50 countries. With its strong over 55-year heritage, Capgemini is trusted by its clients to unlock the value of technology to address the entire breadth of their business needs. It delivers end-to-end services and solutions leveraging strengths from strategy and design to engineering, all fuelled by its market leading capabilities in AI, generative AI, cloud and data, combined with its deep industry expertise and partner ecosystem. The Group reported 2024 global revenues of €22.1 billion.

Get the Future You Want | **www.capgemini.com**