

Cloud TCO and TCO Savings

As global cloud spending pushes past one trillion dollars and data centres expand at an unprecedented pace, **Total Cost of Ownership (TCO)** has become one of the most important measure of whether a cloud investment delivers value or drains it. This paper explains what TCO is, where TCO savings come from, and how to calculate them credibly.

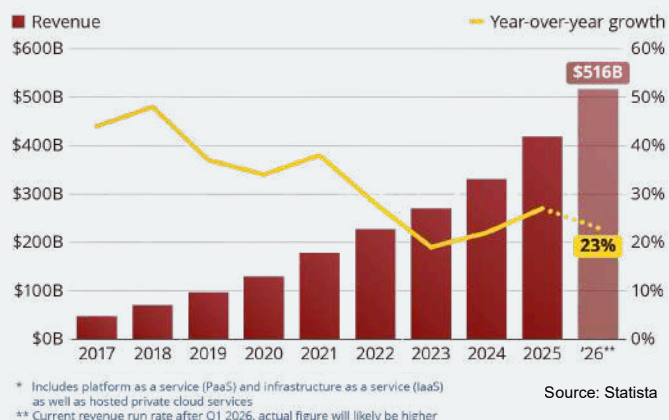
What is TCO?

Total Cost of Ownership (TCO) is the full cost of owning and running a technology asset across its useful life, not just the price paid to buy it or subscribe to it. For an on-premises server, TCO includes the hardware itself, but also the electricity it consumes, the diesel that keeps it running when the grid fails, the cooling, the floor space, the people who maintain it, the licences it needs, the security tooling that protects it, and the audits that prove it complies. For a cloud workload, TCO includes the monthly run-rate, the data that moves in and out, the people who manage and govern it, and the one-time cost of getting there in the first place.

Comparing only the visible costs such as a server's purchase price against a cloud monthly bill, almost always leads to the wrong answer, which is precisely why a thorough TCO study matters. Most of what an IT system actually costs is invisible at the moment of purchase. A credible TCO study makes the invisible visible and turns a guess into a decision leadership can defend.


Why We Need TCO




Cloud and data centres are reshaping the global economy. Spending, capacity and policy attention are all rising together, and every organisation participating faces the same financial question: what will this actually cost us, in full, over the years we will rely on it?



Global Cloud Infrastructure revenue

- Global public-cloud spending reached **USD 723 billion in 2025, up 21.5 per cent year on year**, with growth continuing through 2026 (Gartner press release, 19 November 2024; IDC Worldwide Public Cloud Services Tracker, 2025).
- Hyperscale data centres passed **1,136 operational facilities** by the end of 2024, **doubling in five years**; three companies (Amazon Web Services, Microsoft Azure and Google Cloud) hold roughly 59 per cent of global hyperscale capacity (Synergy Research Group, 2025).
- **Governments are racing alongside enterprises:**

Government	Initiative	Scale / Commitment
 Saudi Arabia	Humain (PIF subsidiary)	1.9 GW of data centre capacity targeted by 2030; \$77 billion investment plan; partnerships with NVIDIA, AMD, AWS and Google Cloud

Government	Initiative	Scale / Commitment
 United Arab Emirates	UAE-US AI Campus (G42 / Stargate UAE)	5 GW AI campus in Abu Dhabi, the largest outside the United States; first 1 GW phase under construction; part of \$200 billion US-UAE investment agreements
 India	GI Cloud (MeghRaj), operated by NIC / MeitY	2,170 central and state ministries/departments hosting cloud applications on MeghRaj; 26 empanelled cloud service providers
 Singapore	National data centre ecosystem ; IMDA-regulated capacity	Ranked #8 globally and #1 in Southeast Asia by installed capacity; anchor for regional cloud and AI workloads across ASEAN

- **Without a credible TCO, a cloud business case is guesswork.** With one, leadership can compare options, set realistic budgets, manage currency and tariff risk, and defend the decision when the numbers are challenged later.

These numbers matter because the stakes of a poor decision scale with the spending. At cloud budgets of hundreds of millions, the difference between a well-constructed TCO and a back-of-envelope estimate can be material. It is equally important to be clear about what a TCO analysis is not: it is not an argument for migration. A credible TCO can and sometimes should conclude that moving to the cloud is not justified — that on-premises operation, optimisation in place, or a hybrid approach delivers better value. That conclusion is not a weakness of the framework; it is the framework working exactly as intended.

The Key Categories of Cost

A credible TCO study captures costs across three distinct groups: what the organisation currently spends on-premises, what the future cloud state will cost to run, and what it costs to get from one to the other. These groups are not equivalent. Some entries are recurring cost buckets; others are one-off transition expenses; and some, foreign exchange movement, tariffs and policy change, are risk factors that must be modelled separately rather than treated as fixed line items. The taxonomy below is drawn from the AWS Well-Architected Framework (AWS, 2024), the FinOps Foundation Cost Categorisation Standard (FinOps Foundation, 2025), and Gartner's IT cost taxonomy (Gartner, 2024).



Current State (On-Premises)

- **Hardware and software:** Servers, storage, networking, OS licences, on-premises software
- **Facilities and power:** Data-centre floor space, electricity, cooling, diesel backup
- **People and skills:** IT operations staff, management overhead, third-party support contracts



Future State (Cloud Run-Rate)

- **Cloud compute and storage:** IaaS, PaaS, SaaS subscriptions; reserved instances; egress
- **Shared services:** Networking, identity, monitoring, service desk allocated costs
- **Governance, security and compliance:** Cloud security tooling, audit, compliance, policy enforcement



One-Off Migration Cost

Migration execution: Migration tooling and consultancy fees; parallel-run period to validate parity; application refactoring where workloads require re-architecture for cloud compatibility

The Eight Categories of a complete TCO Cloud Deployment Modes



TCO Savings: A Definition

TCO Savings is the net financial benefit of migrating, expressed in present-value terms:

$$\text{Net Saving} = \text{PV (on-prem continued)} - \text{PV (future-state cloud)} - \text{migration one-time cost}$$

All three terms are discounted over a chosen horizon (typically five years) at the firm's weighted average cost of capital. The mathematics of net present value is standard corporate finance (Brealey, Myers and Allen, Principles of Corporate Finance, 14th edition, McGraw-Hill, 2025, chapters 5 and 6), and is the same calculation Microsoft's Azure Migrate uses to build its business cases, defaulting WACC to 7 per cent (Microsoft Learn, 2025).

How Workloads Move to the Cloud

Most organisations end up running workloads in more than one place. The four basic destinations are well established: keep a workload on-premises, build a private cloud, adopt a hybrid model, or move to a public hyperscaler. The World Bank's 2022 guide on government cloud migration concludes that hybrid is "the path most governments will take, reflecting the optimal tradeoff between costs, benefits, and risks." The private sector tells the same story. Flexera's 2026 State of the Cloud Report finds that 73 per cent of organisations now operate hybrid estates.

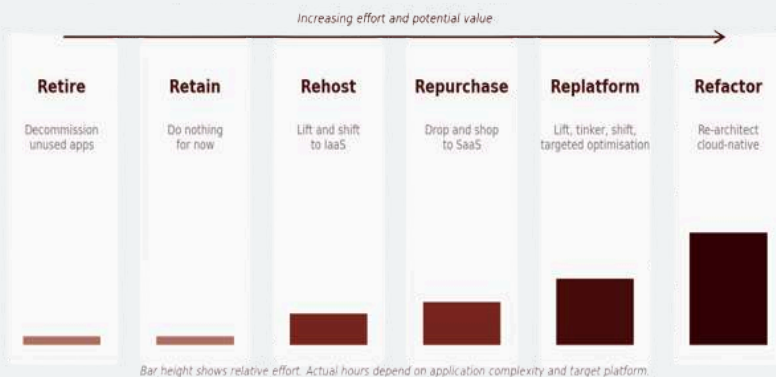
Item	On-premises	Private cloud	Hybrid	Public / hyperscaler
Cost profile	High CapEx	High	Mixed	OpEx
Control	Full	High	Balanced	Shared
Scalability	Low	Medium	High	Very high
Data residency	Full	Full	Configurable	Region-dependent
Typical use case	Legacy & sensitive	Regulated BFSI, defence	Most enterprises & govt	Cloud-native, bursty

Sources: World Bank (2022); Flexera, State of the Cloud 2026.

Cloud Deployment Modes

Within those destinations, individual applications take one of six paths. The original taxonomy of five came from Gartner in 2011 (Richard Watson, Migrating Applications to the Cloud), and AWS added the sixth in 2016 with Stephen Orban's "6 Strategies for Migrating Applications to the Cloud". The six options, in order of effort, are: Retire (decommission unused applications), Retain (do nothing for now), Rehost (a lift-and-shift to infrastructure-as-a-service), Repurchase (drop and shop to software-as-a-service), Replatform (lift, tinker and shift, with targeted optimisation), and Refactor (a full re-architecture as cloud-native). The order matters. Orban himself cautions that "refactoring sometimes takes 20 times longer versus rehosting or replatforming approaches." A migration cost estimate built on a flat percentage will almost always miss the truth; it has to price each application by its disposition and the target it is heading to.

Bringing Market Demand and Execution Discipline Together



Important Things to Consider When Calculating TCO Savings

Five principles, drawn from the FinOps Foundation framework and McKinsey's analysis of more than USD 3 billion in cloud spending, separate a credible TCO study from a hopeful one.

1 Visibility before optimisation.

The FinOps Foundation's framework places "understand cost and usage" before any optimisation activity. Most organisations cannot save what they cannot see. Workload optimisation and waste reduction remained the number one FinOps priority worldwide in 2025 (FinOps Foundation, State of FinOps 2025).

2 Ownership is the cultural lock.

Decentralised teams must own their cloud usage, not just consume it. McKinsey's review of more than USD 3 billion in cloud spending found 10 to 20 per cent of additional savings sits beyond a central FinOps team's reach until ownership is distributed to product and engineering teams (McKinsey, Everything is better as code, 2025).

3 Right-size to actual utilisation, not provisioned capacity.

Cloud waste was estimated at 29 per cent of IaaS and PaaS spending in 2026, the first increase in five years, driven by AI workloads and pricing complexity (Flexera, State of the Cloud 2026). Right-sizing to real utilisation, not what was originally provisioned, is the single largest lever any organisation can pull.

4 Use commitments to lower unit prices, but ramp them carefully.

Reserved instances, savings plans and committed-use discounts can cut bills by 15 to 40 per cent, but only on capacity an organisation is sure it will use. Over-committing locks in future cost as surely as it locks in savings, especially in markets where currency or demand is volatile.

5 Treat business value, not cost cuts, as the goal.

McKinsey's cloud value research is explicit on this point: "the value cloud generates from enabling businesses to innovate is worth more than five times what is possible by simply reducing IT costs" (McKinsey, In search of cloud value, 2024). A TCO that ignores the value side of the equation is incomplete and tends to under-justify the investment.

Important Things to Consider When Calculating TCO Savings

The three largest cloud providers each publish a calculator and an enterprise-grade business-case tool to support TCO analysis. They are free, well documented, and widely used. They are also the right place to begin any TCO study.

- **Amazon Web Services offers the AWS Pricing Calculator (calculator.aws)** for pricing a defined target state, and AWS Migration Evaluator (formerly TSO Logic) for a discovery-driven business case that includes net present value, return on investment, payback period and three-to-five-year cash-flow analyses, at no cost (AWS Migration Evaluator documentation, 2025).
- **Microsoft offers the Azure TCO Calculator** for quick comparisons and Azure Migrate for an end-to-end business case with 95th-percentile right-sizing, year-on-year cash flow, NPV at 7 per cent WACC, and Azure Hybrid Benefit modelling (Microsoft Learn, 2025). Google Cloud offers

- **Google Cloud Migration Center** for discovery, asset grouping and cost estimation, with explicit handling of egress to other regions and to the public internet (Google Cloud Migration Center documentation, 2025). The three tools share the same underlying method, but their defaults, discount mechanics and outputs differ in ways that matter.

Each of these tools is also produced by a vendor with a direct commercial interest in the outcome of the analysis. Their default assumptions, discount rates, utilisation benchmarks, on-premises cost baselines, are set by organisations that benefit when the migration case is strong. An independent TCO model corrects for those assumptions, applies local market rates, and reaches a conclusion that is not anchored to any provider's pricing sheet. That independence is precisely the position that VTT Global's parameterised model occupies: built on international hyperscaler methodology, but calibrated to the actual cost structures that its clients face.

Items			
Primary tool	Pricing Calculator, Migration Evaluator	TCO Calculator, Azure Migrate	Migration Center
Discovery method	Agentless collector	Migrate appliance	Migration Center discovery client
Right-sizing approach	Utilisation benchmark, zombie detection	95th-percentile, comfort factor	Best-practice optimised configs
Discounts modelled	Reserved Instances, Savings Plans	Reserved, Hybrid Benefit, ESU	Committed-Use Discounts
Output	PPTX and Excel business case	NPV business case (7% WACC default)	Cost estimates, export-ready

About Us

VTT Global is a management consulting firm advising governments, multilateral institutions, and large enterprises on technology strategy, cloud economics, and digital infrastructure. With over 260 professionals across South Asia, the Middle East, and international markets, we combine global reach with hands-on delivery experience.

Our cloud advisory practice covers end-to-end TCO modelling, cloud migration strategy, vendor assessment, FinOps governance, and public-sector cloud programme design, giving us a practitioner's understanding of what separates a credible cost case from a vendor estimate.

Why VTT Global?

- Global footprint with local expertise
- 200+ subject matter experts
- 260+ professionals worldwide

Our Focus Industries

- Technology, Media and Telecom (TMT)
- Energy, Climate & Sustainability
- Public Policy and Governance

Author:



Rizwan Shoukat
CEO and Managing Partner, VTT Global

Over three decades of experience spanning telecom operations, network transformation, and digital infrastructure advisory, with engagements across South Asia, the Middle East, and international markets. A track record of advising telecom operators, governments, and large enterprises on cloud strategy, data centre economics, and large-scale digital transformation programmes.