Cloud and Data Center Market Snapshot – December 2025

Vladimir Galabov

Senior Research Director, Enterprise Infrastructure

askananalyst@omdia.com

Introduction

- This is the sixteenth edition of this report series from Omdia's Cloud and Data Center Research Practice aimed at providing a brief, quick, and informal update on the key trends disrupting the market.
- Omdia will distribute this report via e-mail first, to maximize the speed of delivery, and publish it on omdia.com afterward.
- **You can subscribe to this newsletter by using this link.**
- This content is for personal use and should not be distributed to third parties outside your organization or posted online.
- Distribution within your organization is permitted.
- **If you would like to add colleagues to our e-mail distribution list, please forward the sign-up link to them.**

Content

- Industry outlook is increasingly positive
- We've raised out data center investment forecast
- …across all segments
- More than 2X growth in data center power capacity
- What factors impacted this forecast?
- General purpose server consolidation is ongoing
- The power capacity of a single AI chip is inversely proportional the number of AI chips per server
- AI adoption is the primary driver for data center capacity increase
- A steady growth in new IT load rollout

Industry outlook is increasingly positive

Our data center investment forecast is tied to data center operator capex plans, vendor guidance and technology roadmaps. To eliminate bias, we have also anchored it in net new IT load additions – an independent analysis anchored in semiconductor design, server engineering and useful life.

Growth boosting factors

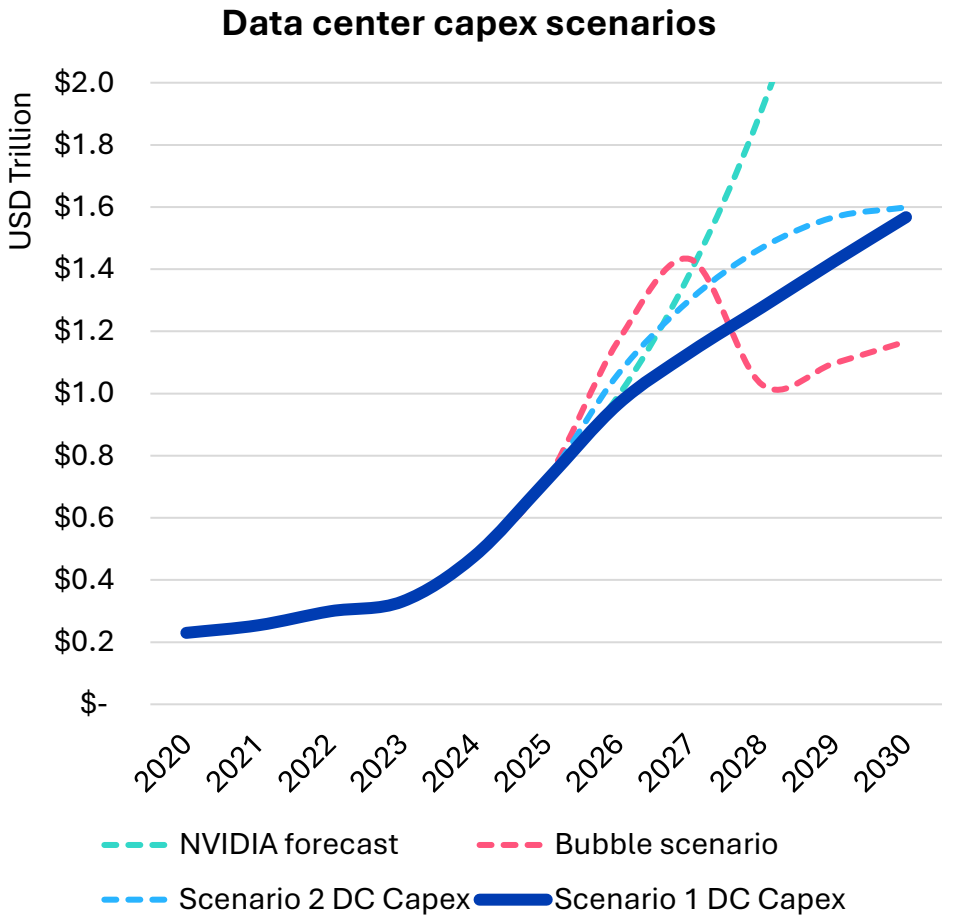
- Demand for AI computing is very strong. Most developers complain of capacity constraints.
  - LLMs are getting larger, increasing AI training compute demand
  - Reasoning models are consuming more compute during generation (AI inference)
  - New model features result in end-user demand spikes
- This is translating in higher performing and higher power AI chips, rising server, rack and data center power density.
- AI adoption is still low, more users and higher usage per user is coming.
- Within 2025 and 2026 a general-purpose server refresh cycle is boosting server shipments.
- Supply chain constraints is resulting in higher cost for commodity components.

Growth stifling factors

- A continuous pursuit of efficiency is driving a simplification in data center design, and a reduction of $/Watt, e.g. a simpler power distribution model utilising 800VDC.
- New higher capacity thermal management equipment is driving a decrease in $/Watt for cooling.
- Competition in AI computing is heating up with AMD, Google, AWS and Huawei challenging NVIDIA's dominance. This could erode AI chip margins slowing down the growth in the average selling prices of AI servers.
- Consolidation of general-purpose computing servers is occurring during the refresh cycle, limiting net power additions.
- AI server useful life is being extended over time.

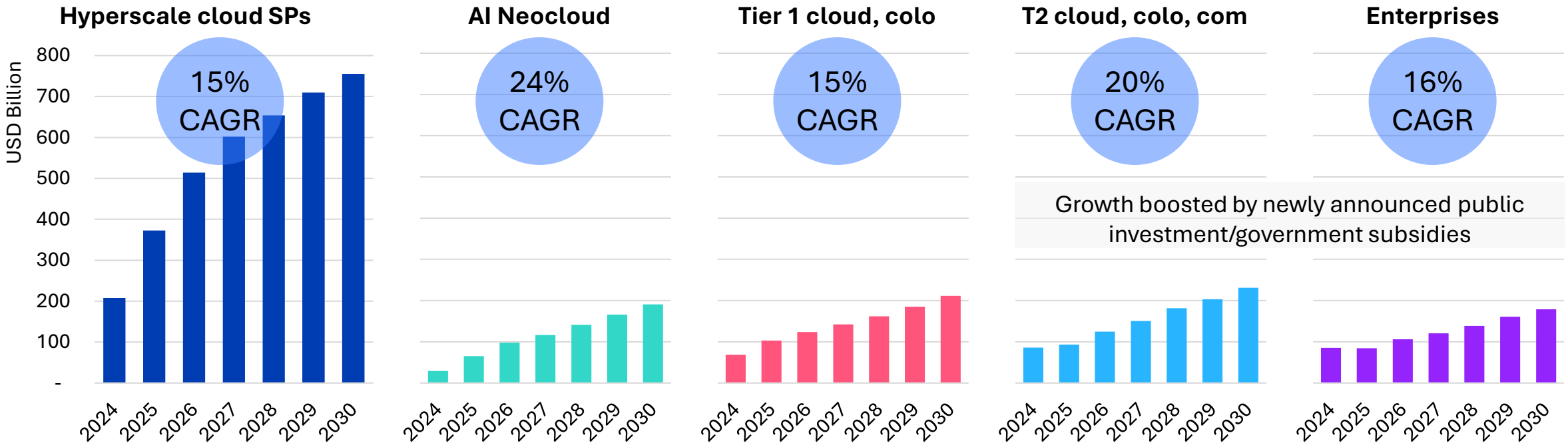
We've raised out data center investment forecast

- We expect data center capex in 2030 to be just shy of $1.6T, growing at a 17% CAGR from 2025 till 2030. We considered four scenarios before reaching our consensus:
  - **Scenario 1** (70% likelihood) considers actual order pipeline and demand which are both strong. This is balanced against numerous constraints (power availability, manufacturing capacity, supply chain hiccups).
    - We have aligned this scenario to the 2025-26 NVIDIA order backlog.
    - **Actual data center builds are occurring slower than deal announcements, which reduces the risk of a quick overbuild.**
    - **In the short term the only way to build AI computing capacity is by buying more AI chips.**
    - **In the long term, we expect companies to build AI computing capacity through more avenues: (1) acquire a failing competitor to use their GPUs, (2) extend AI server useful life.**
  - **Scenario 2** (20% likelihood) considers constraints are having a lesser impact in the short term, e.g. electrical equipment manufacturers can quickly increase capacity. An accelerated pace of development results in accelerated time to failure for some developers, with consolidation occurring at a faster rate.
  - **Scenario 3** the "bubble scenario" (5% likelihood), reflected a failure to realize productivity gains through AI use quickly enough, and after 5 years of accelerated investment, investors get spooked. We consider 2027 to be a key year because key AI developers have made revenue commitments for that year.
  - **Scenario 4** the "NVIDIA forecast" (5% likelihood) is what NVIDIA shared. We assume this does not consider all the constraints (e.g. power and manufacturing capacity) and how slowly enterprises adopt new technology.



Omdia by informa techtarget

...across all segments



Amazon
Google
Meta
Microsoft

CoreWeave
Crusoe
Nebius
xAI
...

Alibaba
Apple
Oracle
ByteDance
...

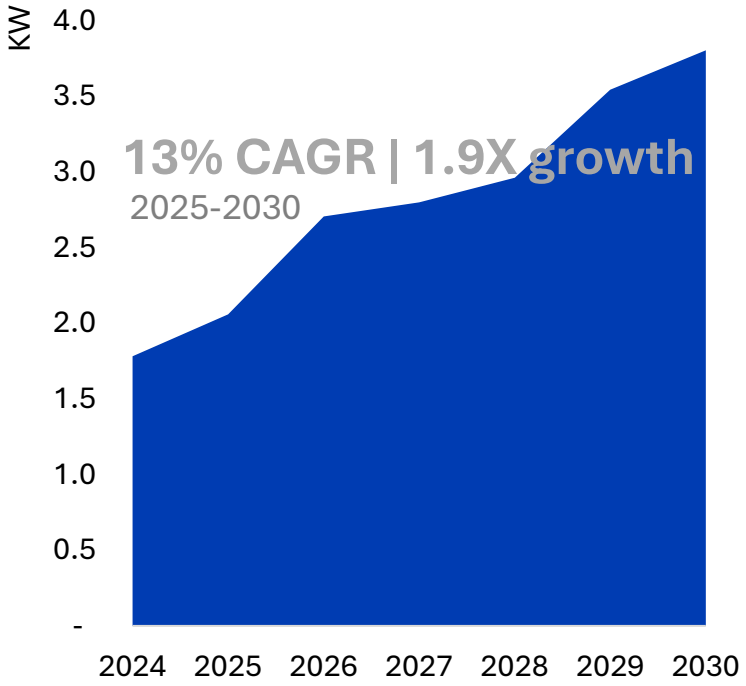
OVH Cloud
Rakuten
Reliance Jio
Yandex
...

US DOE
Goldman Sachs
UK NHS
TSMC
...

Omdia by informa techtarget

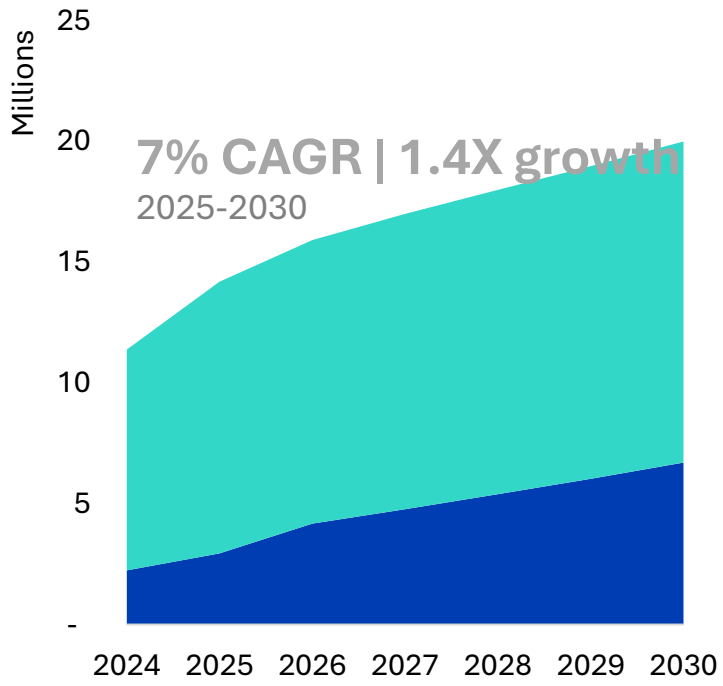
More than 2X growth in data center power capacity

Average server power capacity*

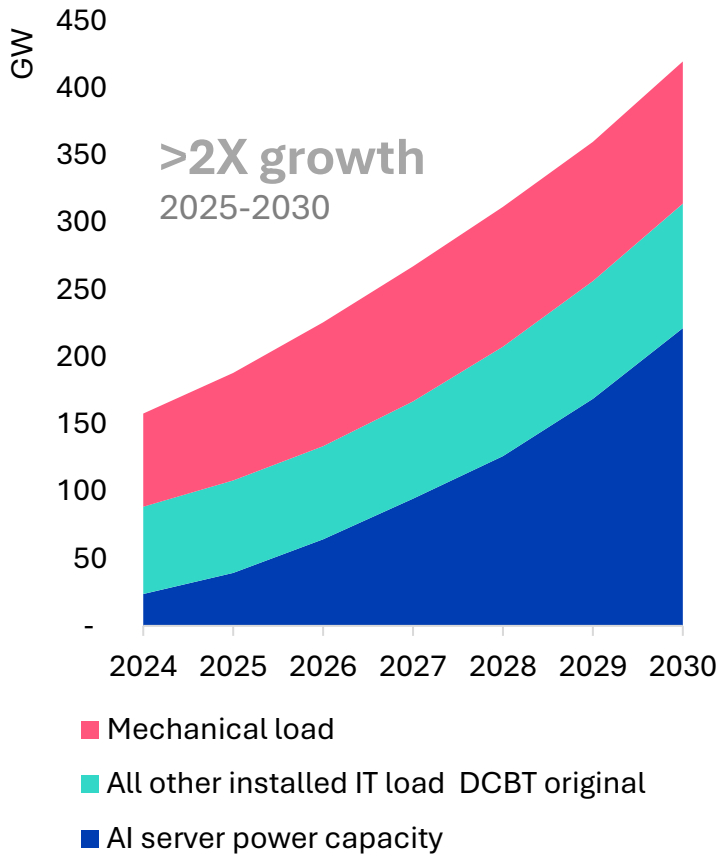


*Considers server architecture changes, move from 8-GPU to 4-GPU to 2-GPU AI server between 2023 and 2027

Server shipments

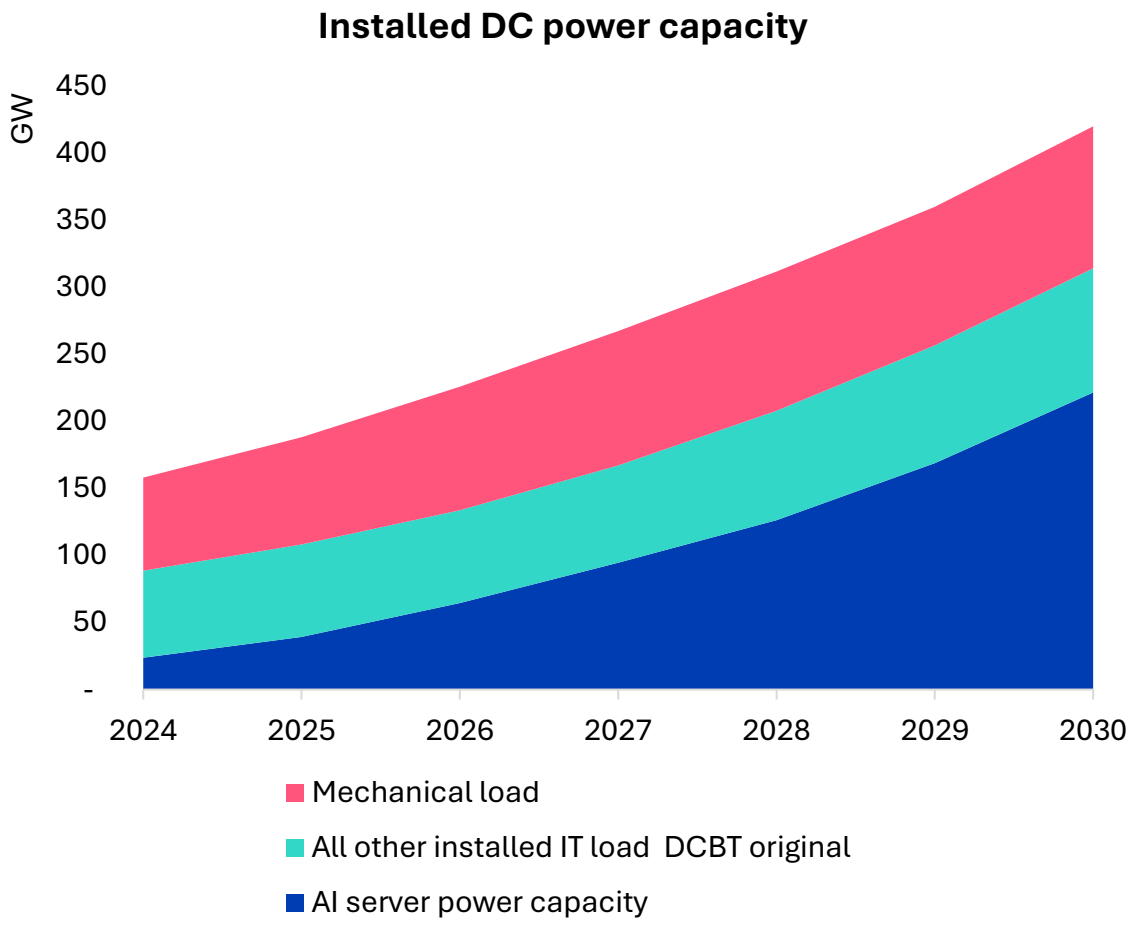


Installed DC power capacity



Omdia by informa techtarget

What factors influence this curve?

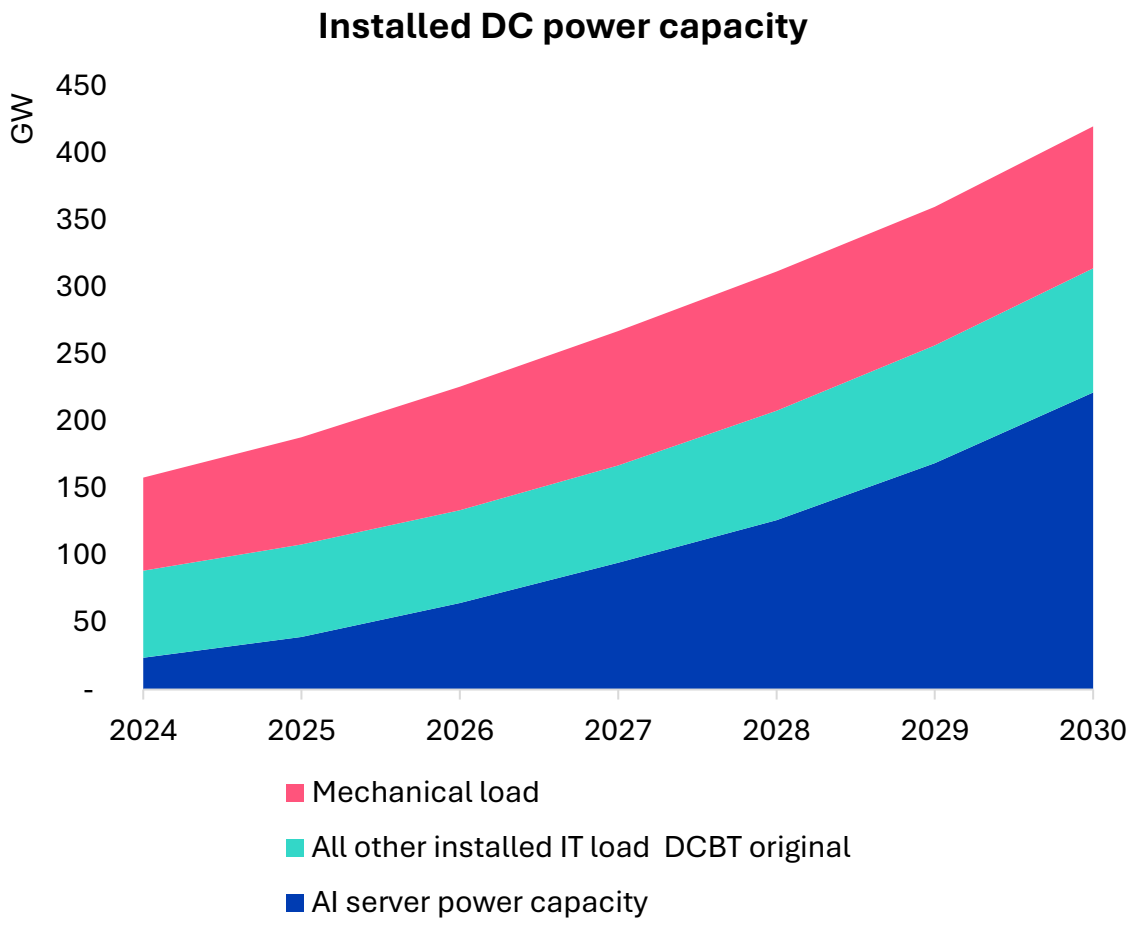


- Increase in processor performance (directly correlated with an increase in power consumption)
- Increase in the number of processors per server
- Increase in server/storage/switch shipments
- Increase in server/storage/switch useful life
- Underutilizing servers/storage/switches
- Lower data center efficiency (PUE increase)

- Change in processor design priorities, use Moore's Law to reduce power not increase performance
- Reduce the number of processors per server
- Slow down server/storage/switch installation
- Reduce server/storage/switch useful life
- Improve the utilization of servers/storage/switches
- Better data center efficiency (PUE decrease)

Omdia by informa techtarget

What factors impacted this forecast?



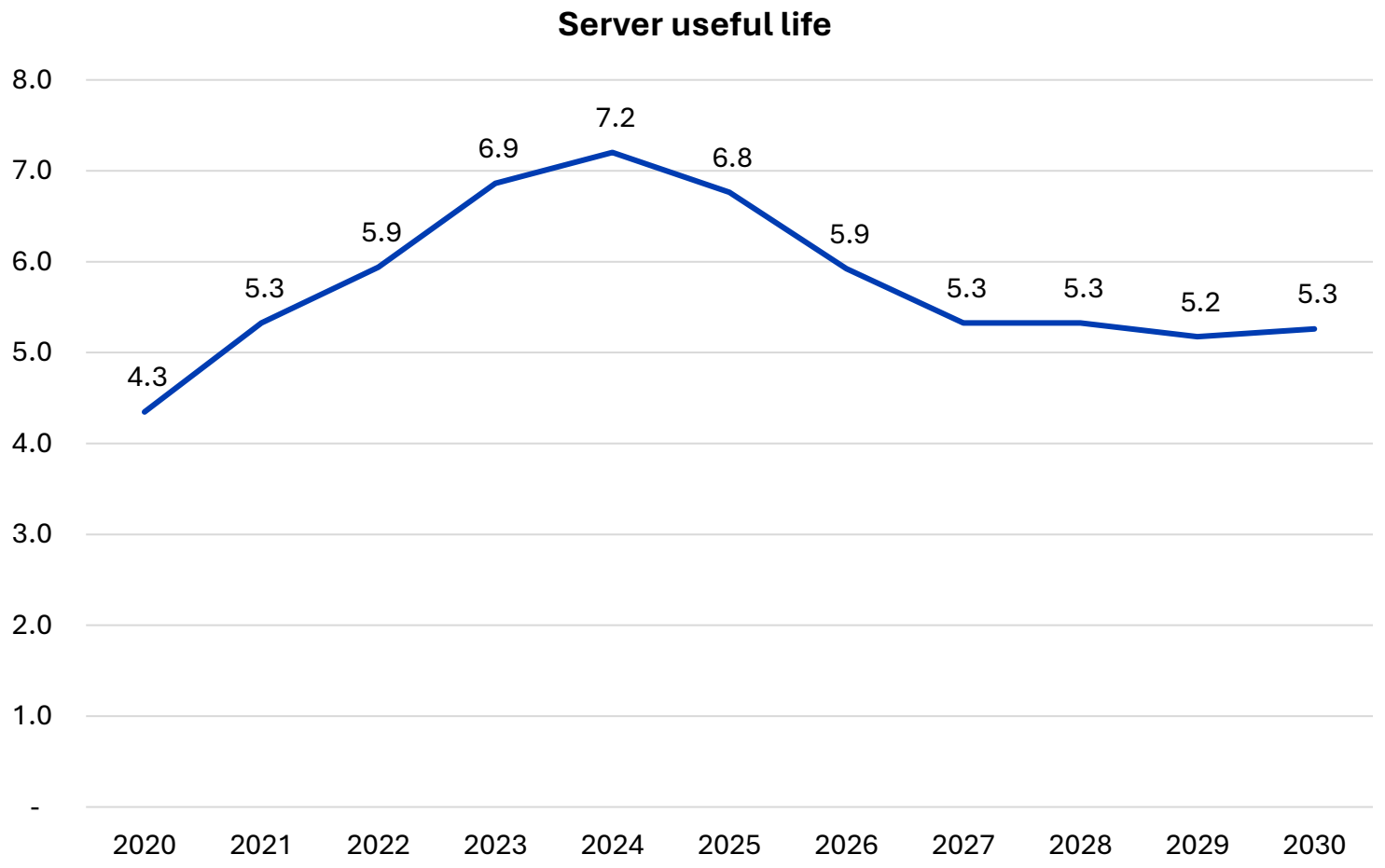
- Increase in processor performance (directly correlated with an increase in power consumption)
- Increase in the number of processors per server
- Increase in server/storage/switch shipments
- Increase in server/storage/switch useful life
- Underutilizing servers/storage/switches
- Lower data center efficiency (PUE increase)

- Change in processor design priorities, use Moore's Law to reduce power not increase performance
- Reduce the number of processors per server
- Slow down server/storage/switch installation
- Reduce server/storage/switch useful life
- Improve the utilization of servers/storage/switches
- Improve data center efficiency (PUE decrease)

Omdia by informa techtarget

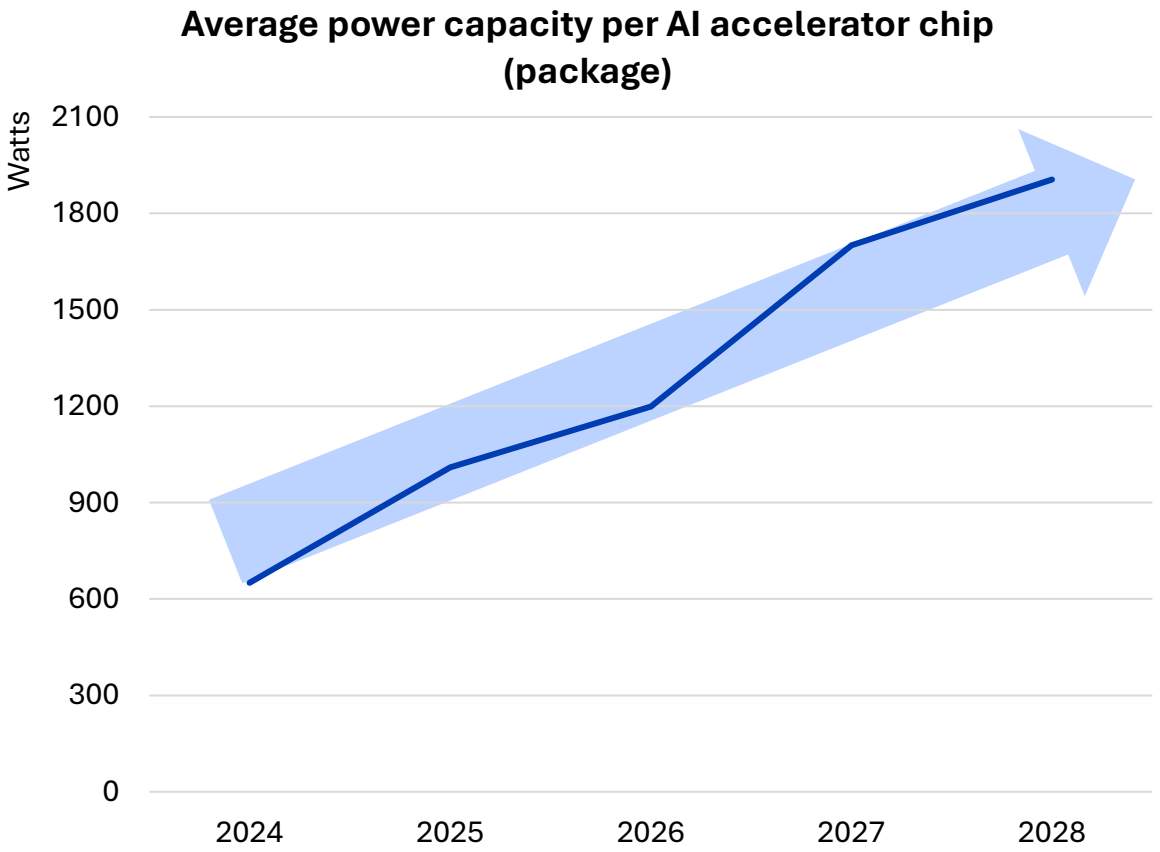
General purpose server consolidation is ongoing

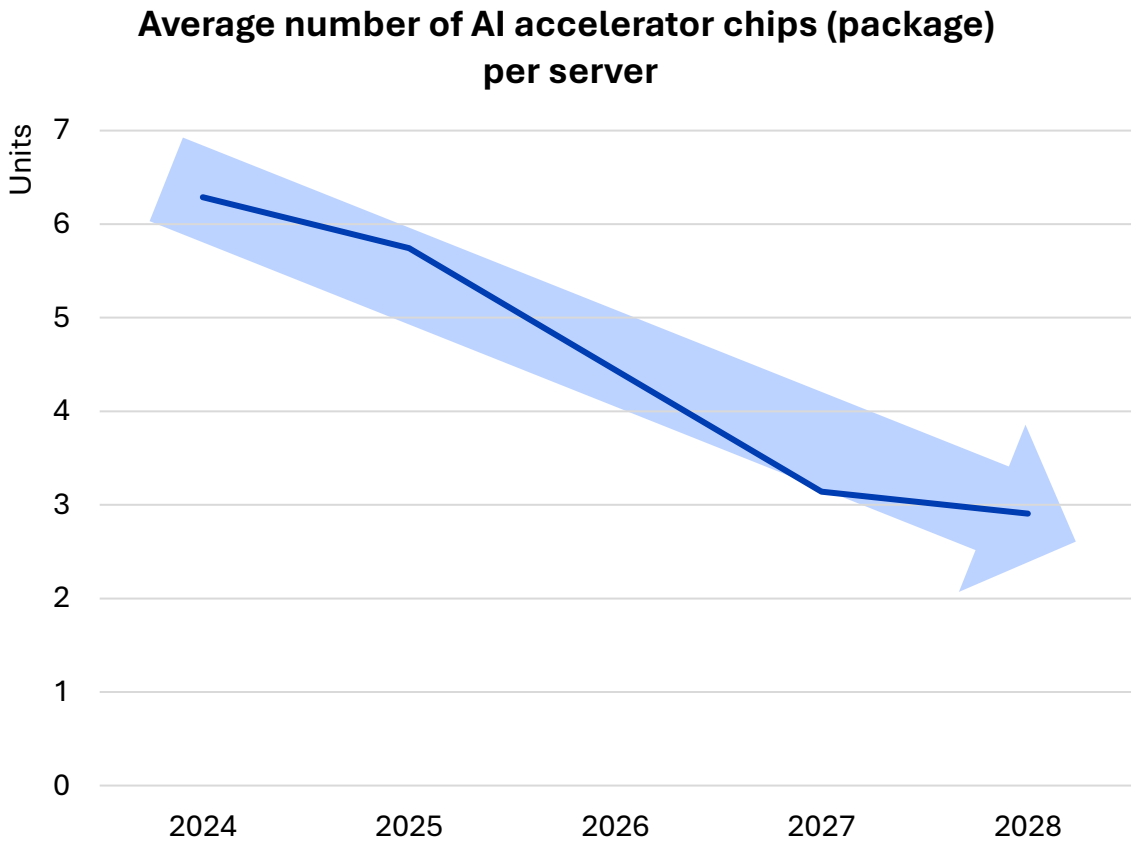
- A server refresh cycle kicked off in 2025 and will continue for 6-8 quarters reducing server useful life from a record reached in 2024.
- Servers used for database & analytics, media, graphics & gaming, web services, networking, security & telecommunications, and other IT applications & business processes are all targets for the refresh.
- One higher performance new servers can replace several older generation servers. This consolidation results in efficiency gains.
- Approximately 26GW of net savings will be realised between 2025 and 2030 from the refresh as older inefficient servers are decommissioned.
- The only category where we expect server useful life to increase is AI where useful life is on the rise. CoreWeave, for example, indicated it intends to use its newly installed AI servers for 7 years.
- We expect most NVIDIA Blackwell servers to still be in use in 2030.



Omdia by informa techtarget

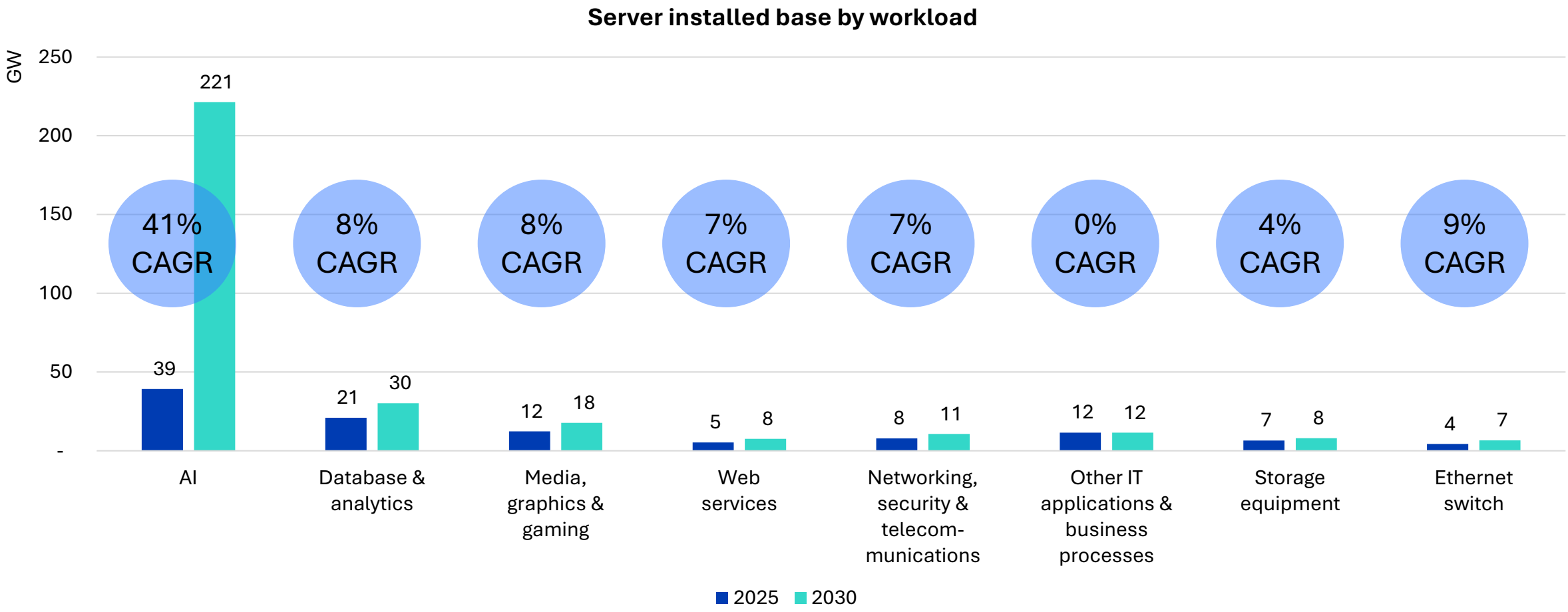
The power capacity of a single AI chip is inversely proportional the number of AI chips per server





Omdia by informa techtarget

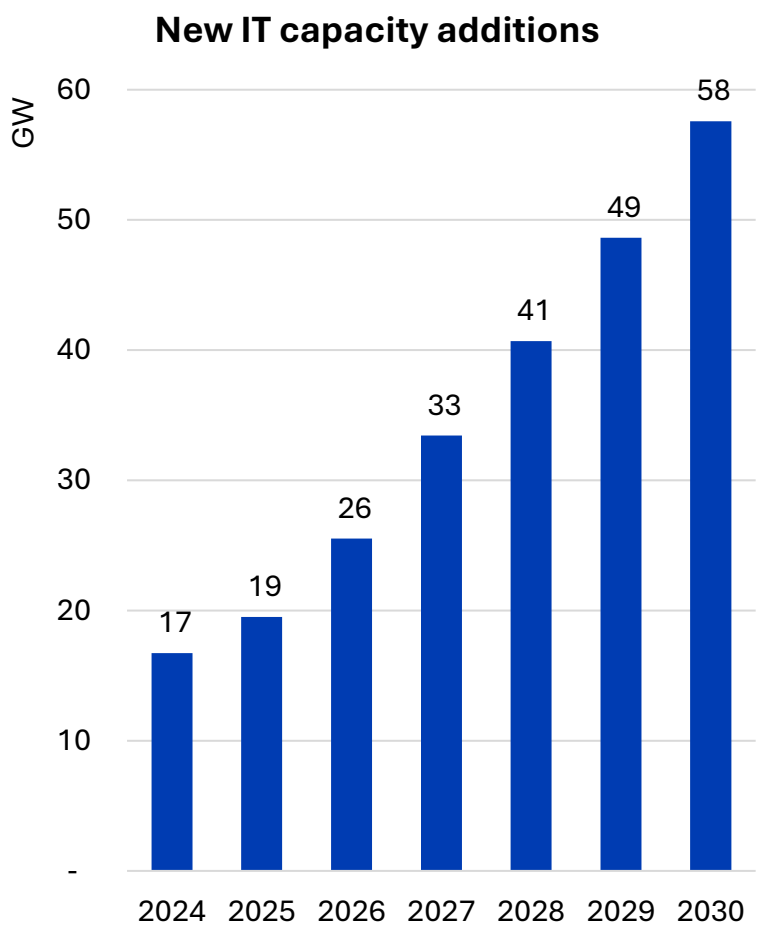
AI adoption is the primary driver for data center capacity increase



Omdia by informa techtarget

A steady growth in new IT load rollout

- We expect each year to add much more IT capacity across the global data center market, driven primarily by the rollout of new AI computing clusters.
- Net new IT capacity to be installed globally is 205GW (between 2026 and 2030).
- Just 20 companies or megaprojects account for nearly half of this number.
- It is likely that many of the companies on this list will make further project announcements.
- Hundreds of billions of additional government funding across Europe, Asia and North America will play a significant role in tens of GW on top of the announced projects in this table.
- We expect 20-30 GW to be added in China alone between 2026 and 2030.



Announced projects	GW
Microsoft	16.1
Google	12.8
Amazon	10.9
Meta	12.7
OpenAI, Oracle, Crusoe	10.0
CoreWeave	3.5
Reliance Jio	3.0
Adani ConneX	1.0
Tata	1.0
STT Global Data Center	1.0
UAE/Dubai megaproject	5.4
Saudi Arabia (cumulative projects)	3.5
MistralAI (Paris project)	1.4
Start Campus (Portugal)	1.5
IREN (US projects)	3.2
Quantum Loophole	2.0
Port Washington/Prince William	1.0
Data4 (France)	1.0
Korea's largest announced project	3.0
Japan's largest announced project	2.0
	96.0

Omdia by informa techtarget

Bottom line and recommendations

Data-based insights	Recommendations for vendors	Recommendations for users
Despite speculation of an AI bubble, both adoption and investment continue at a fast pace. AI developers remain capacity constrained.	Ignore the noise and speculation and focus on optimizing your roadmaps to remain competitive in the AI economy.	Early adopters of AI will reap the biggest benefits. The pace of model improvement is very fast. Early challenges should not discourage new projects.
The pace of investment across tier 2 cloud and AI Neocloud is accelerating as multiple governments in Europe and Asia launch incentives for local capacity build.	Pursue these from two side – (1) direct through wholesale tenants (e.g. Microsoft), (2) indirect through SIs and the other channel partners, in addition to forming relationships with the AI Neoclouds.	Utilise new public funding and the expanding ecosystem to accelerate time to AI adoption, time to productivity gains.
New data centers will be engineered differently. The design of AI chips, or AI-optimized servers, racks, data center thermal management, power transformation, distribution and backup is all changing very quickly.	It is not safe to bet only on one architecture. A diverse roadmap will be a competitive advantage. Staying close to the user community has never been more important. Form a continuous feedback loop with the leading cloud SPs.	Data center and IT teams need to avoid complacency at all costs. Deep understanding of new technologies should be pursued as a matter of urgency. Keep an open mind about doing things differently.

Disclaimer

The Omdia research, data and information referenced herein (the “Omdia Materials”) are the copyrighted property of TechTarget, Inc. and its subsidiaries or affiliates (together “Informa TechTarget”) or its third party data providers and represent data, research, opinions, or viewpoints published by Informa TechTarget, and are not representations of fact.

The Omdia Materials reflect information and opinions from the original publication date and not from the date of this document. The information and opinions expressed in the Omdia Materials are subject to change without notice and Informa TechTarget does not have any duty or responsibility to update the Omdia Materials or this publication as a result.

Omdia Materials are delivered on an “as-is” and “as-available” basis. No representation or warranty, express or implied, is made as to the fairness, accuracy, completeness, or correctness of the information, opinions, and conclusions contained in Omdia Materials.

To the maximum extent permitted by law, Informa TechTarget and its affiliates, officers, directors, employees, agents, and third party data providers disclaim any liability (including, without limitation, any liability arising from fault or negligence) as to the accuracy or completeness or use of the Omdia Materials. Informa TechTarget will not, under any circumstance whatsoever, be liable for any trading, investment, commercial, or other decisions based on or made in reliance of the Omdia Materials.

Get in touch

Americas
customersuccess@omdia.com
08:00 – 18:00 GMT -5

Europe, Middle East & Africa
customersuccess@omdia.com
8:00 – 18:00 GMT

Asia Pacific
customersuccess@omdia.com
08:00 – 18:00 GMT + 8