

Executive summary

As generative AI (GenAI) adoption accelerates, moving into production and implementation, enterprises face challenges posed by outdated legacy infrastructure. While many enterprises have modernized by migrating to public cloud platforms, these deployments remain concentrated in centralized datacenters, often far from remote business locations and customers.

To address this, enterprises are anticipating **an infrastructure evolution at the edge** aimed at overcoming technical debt and enabling true AI readiness. This shift requires tackling two priorities: modernizing legacy infrastructure and extending public cloud capabilities to edge locations. GenAI's data-intensive requirements for proximity to customers call for investments in a modern infrastructure that extends from core to edge, addressing scalability, latency, cost, and performance challenges.

This transformation, or what is emerging as **“The Edge Evolution,”** necessitates a new connected digital backbone to handle the exponential growth of edge data with enhanced performance, robust security, and scalability. By leveraging public cloud-based infrastructure across the core to edge, enterprises can rightsize workloads and run applications efficiently, wherever they are needed. This decentralized approach to IT services improves performance, optimizes IT costs, and enhances customer experiences by bringing services closer to users.

IDC forecasts a significant shift toward public cloud-based provisioned services at the edge over the next three years. Spending on public cloud services at the edge is projected to reach US\$29 billion by 2028, representing a 17% compound annual growth rate (CAGR) from 2024 to 2028 in the Asia/Pacific (excluding Japan) region.¹

This IDC InfoBrief explores the evolving IT landscape, highlighting how enterprises are building their digital infrastructure around connected public cloud models. It also provides best practices and actionable recommendations for designing a modern digital infrastructure that meets both current and future needs.



By 2028, in response to the growth of GenAI inferencing workloads, **30%** of A2000 in Asia/Pacific will supersize edge IT by doubling spend on remote compute, storage, and networking resources.²

¹ IDC Worldwide Edge Spending Guide — Forecast, 2025, ² IDC FutureScape: Worldwide Digital Infrastructure 2025 Predictions — Asia/Pacific (Excluding Japan Implications)