Scaling GenAI Successfully

Adopting Direct-to-Chip Liquid Cooling to Help Address Exponentially Increasing Power Consumption


Andrew Buss
IDC IT infrastructure EMEA


Luis Fernandes
IDC IT infrastructure EMEA

Table of Contents

CLICK BELOW TO NAVIGATE TO EACH SECTION IN THIS DOCUMENT.

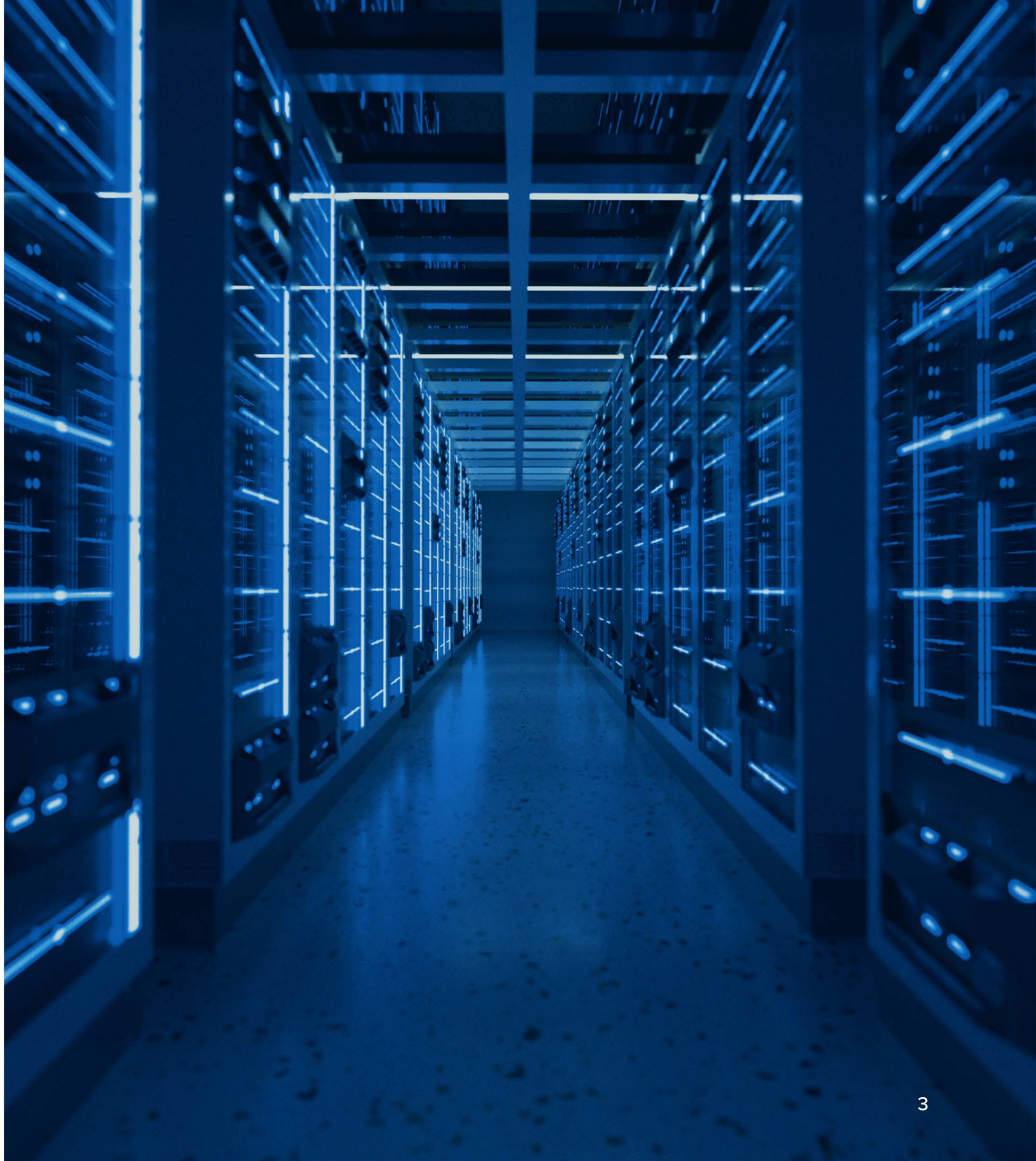
In this InfoBrief

This InfoBrief is designed to offer a comprehensive perspective on the status and future strategies related to European organizations' implementation of new solutions to cope with the power and cooling needs of a dynamic ecosystem. It delves into current adoption rates, prevailing attitudes, and challenges associated with new workloads and managing a state-of-the-art IT infrastructure.

Throughout the pages of this InfoBrief, we analyze the adoption of generative AI (GenAI) and the impact on power consumption and rack power density and examine attitudes towards air and liquid cooling technologies as datacenter architectures evolve at a faster rate than ever before.

The document concludes by presenting crucial recommendations tailored specifically for CIOs to facilitate the seamless integration of new cooling solutions within their respective organizations.

This thorough overview is based on survey data collected from a diverse group of 300 European CIOs, representing multiple verticals. The survey data was collected in the summer of 2025. Survey findings are further enriched by ongoing research conducted by IDC on the subject of IT infrastructures.



Executive summary

Among European enterprises, there is a growing realization of the urgent need for scalable, energy-efficient solutions to support new and more power-dense workloads, in the evolving landscape of GenAI infrastructure. GenAI is driving unprecedented increases in power density and cooling demands, and traditional air-cooling methods are proving obsolete.

This document highlights the strategic importance of digital infrastructure modernization, particularly the adoption of liquid cooling technologies, to support next-generation AI deployments.

Bigger companies, as well as the most digitally advanced, are leading the charge in GenAI adoption and infrastructure transformation; they are therefore more likely to also adopt new ways to cool their IT real estate.

Different approaches to liquid cooling are being tested and deployed, but is yet unclear if one will emerge as the clear winner, since widespread adoption is still a long way off.

Digital infrastructure: Driver of differentiation for successful businesses

The economy is increasingly **moving towards more digital engagement models, products, and services.**

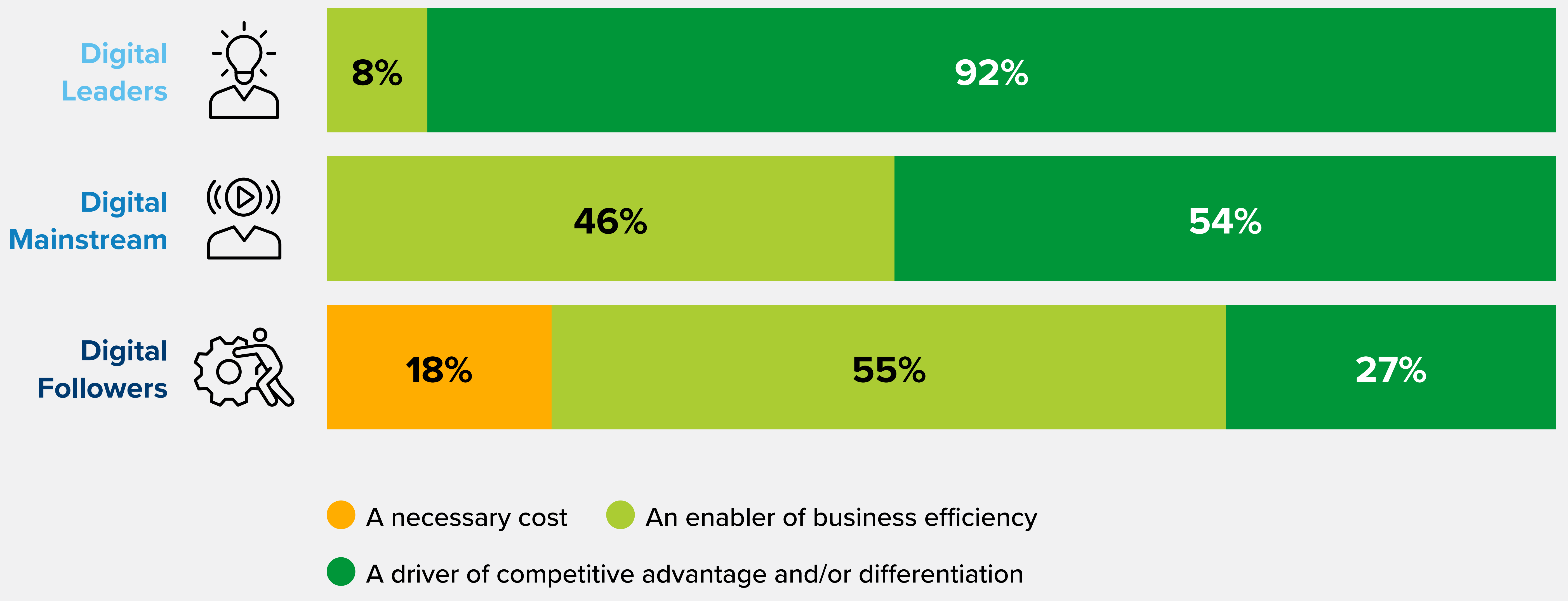
IT is moving out of the back office — where it has been responsible for driving **efficiency savings** — and is now responsible for user, employee, and partner **digital experiences,** increasingly becoming the **business model** itself.

Companies closest to this business model are applying multiple elements that make up a digital business and are aptly named **Digital Leaders** and make up about 25% of companies in EMEA.

The following 50% of companies have started to use some elements of a truly digital business but still have a long road ahead of them. These represent the **Digital Mainstream**.

On the other end are companies that are currently the least prepared to embark on their digital transformation journey. Also representing about 25% of all companies, these are classified as **Digital Followers.**

Senior management's view of the role of IT within the organization



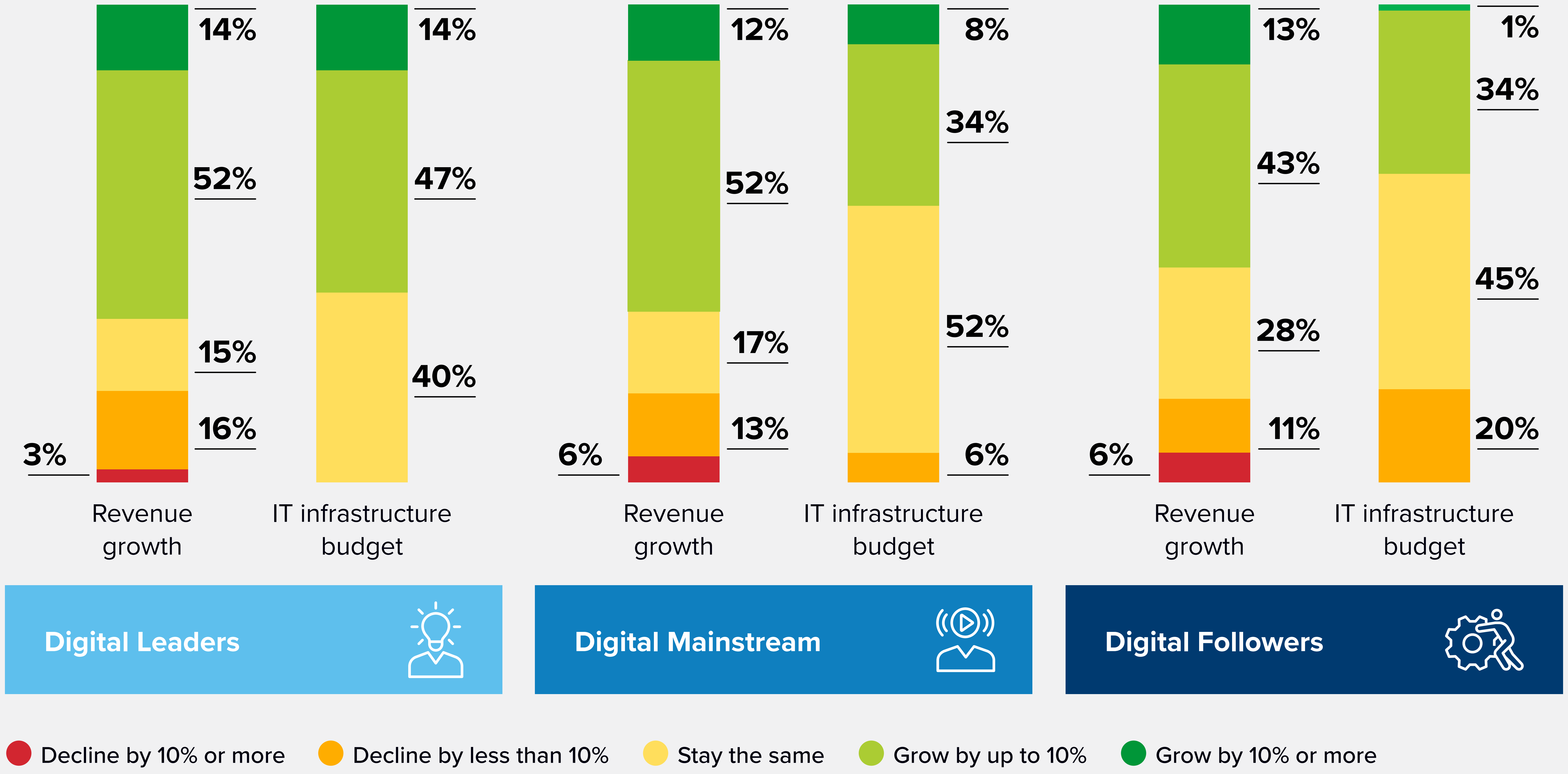
Source: IDC's *Survey Europe September 2025* (n = 300)

Investing in digital infrastructure for real dividends

In this new digital economy, **companies that invest more** heavily in their digital infrastructure are **more agile and flexible, without compromising on security and resilience.**

This gives them the competitive edge needed to capture more market share or pivot to new markets.

Company revenue and IT infrastructure budget growth from previous year to current year



Source: IDC's *Survey Europe September 2025* (n = 300)

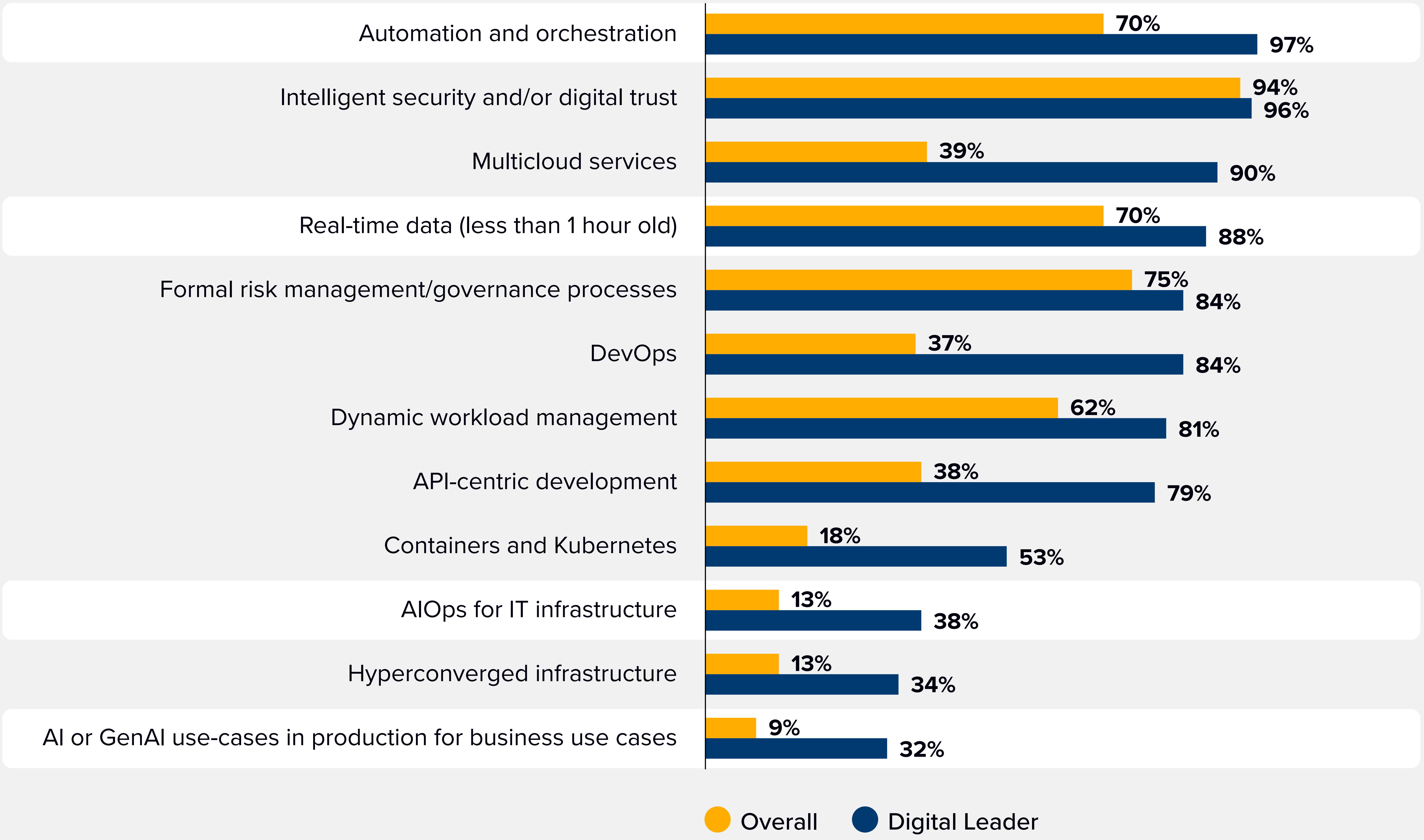
Driving AI success through digital leadership

The top quarter most digitally advanced organizations — the **Digital Leaders** — invest far more extensively across a set of advanced IT infrastructure capabilities than the overall market does.

This cloud-native infrastructure gives them the agility, flexibility, and scalability needed to adapt as the market evolves.

Crucially, adoption of **real-time data, automation and orchestration, and AIOps** provides Digital Leaders with the foundation to move beyond limited proofs-of-concept for GenAI and agentic AI and into full-scale production for these AI workloads.

Extensive use of the following advanced IT infrastructure elements



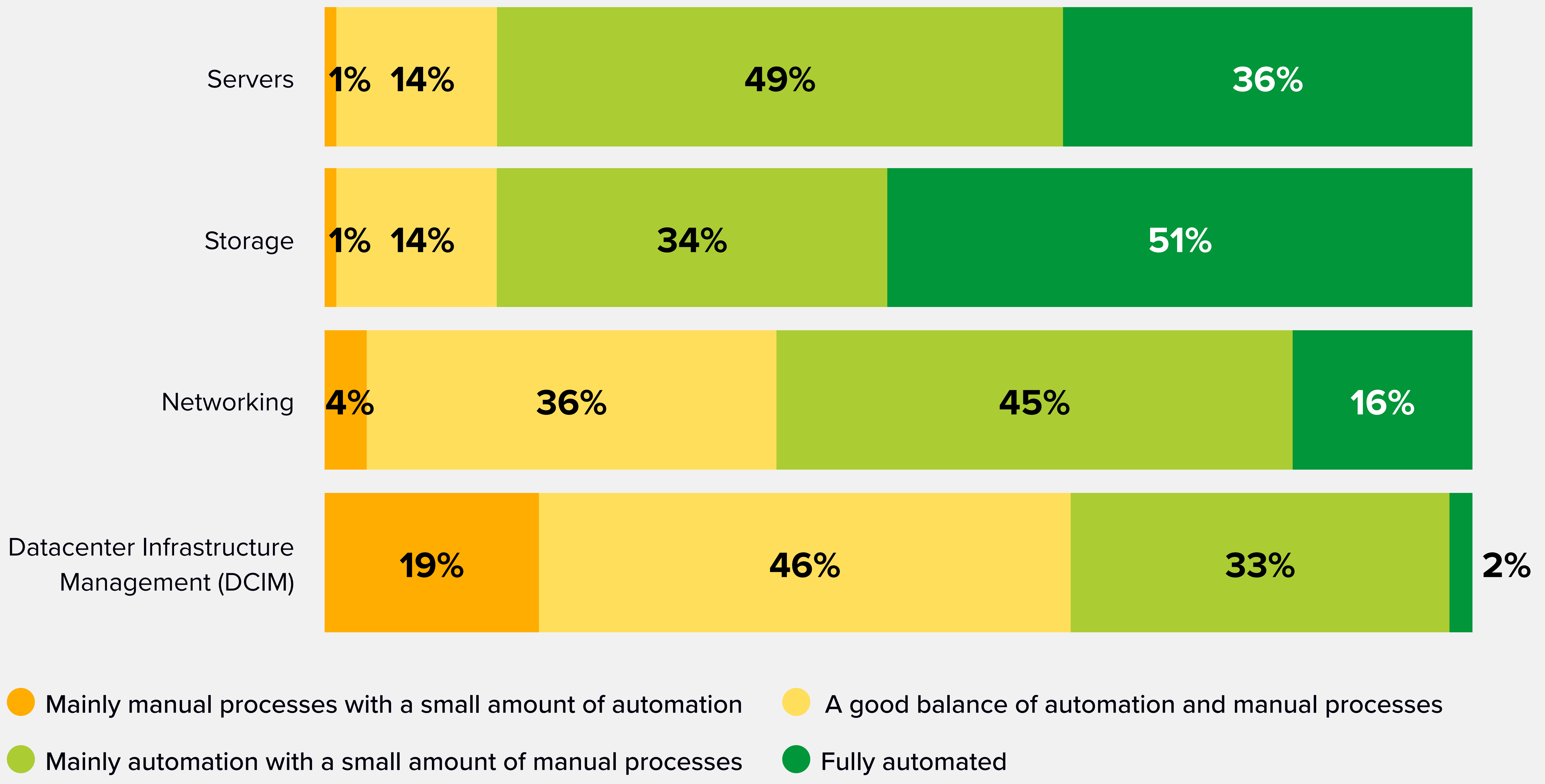
Source: IDC's *Survey Europe September 2025* (n = 300)

Moving beyond the single glass of pain

A modern digital business must be highly automated to succeed — otherwise, change is too slow and too unpredictable to deliver scalable and secure digital services. With GenAI and agentic AI workloads growing, this is more critical than ever.

While great progress has been made towards more integrated end-to-end automation in compute and storage, the networking fabric and, particularly, the datacenter infrastructure and facilities still have a long way to go.

Level of automation in infrastructure management



A modern approach to managing IT real estate is essential for digital transformation and relies heavily on automation and orchestration.

Source: IDC's *Survey Europe September 2025* (n = 300)

Rethinking scale for GenAI infrastructure buildout

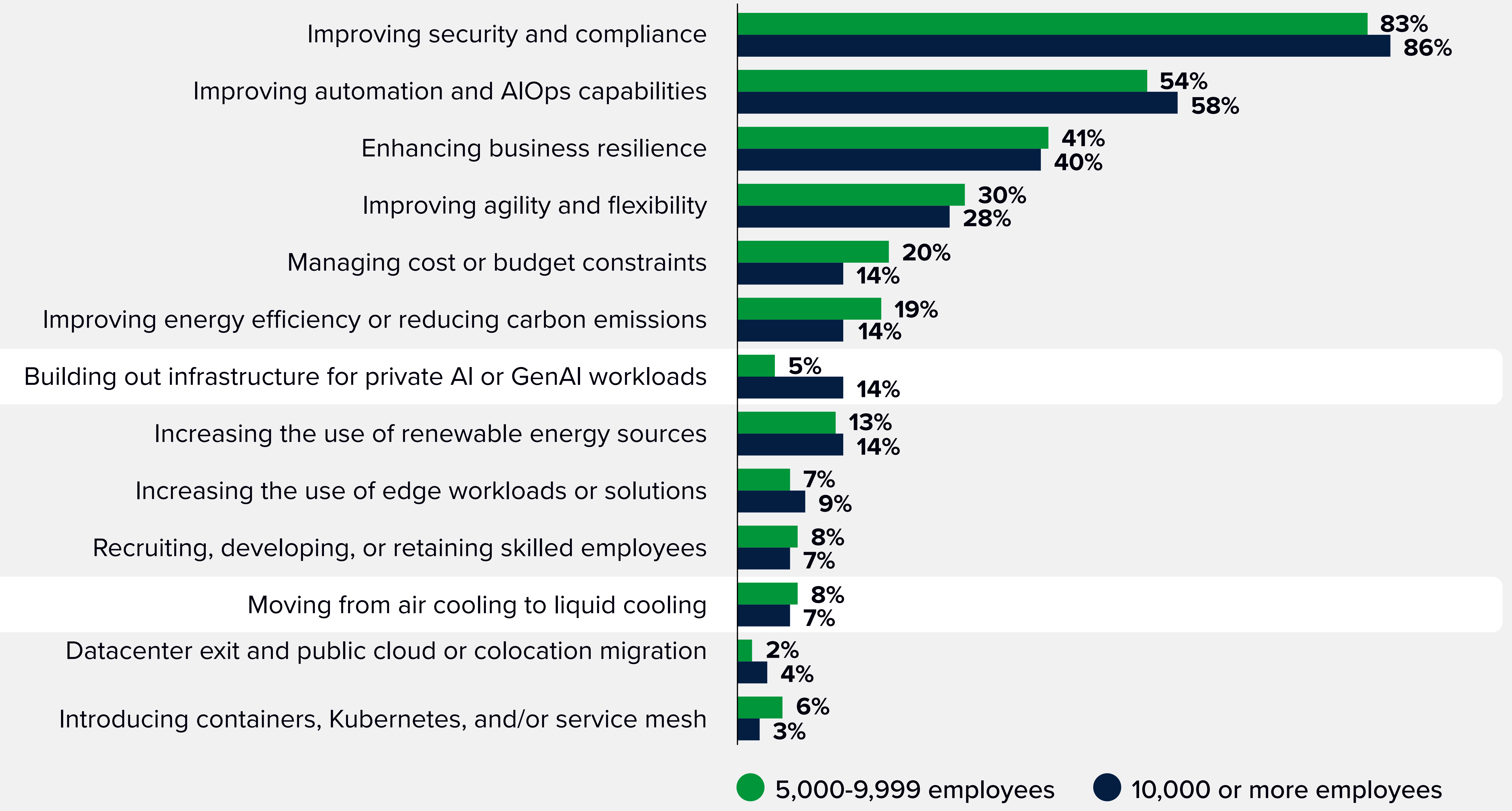
Companies are still at the early days of building out GenAI infrastructure.

Very large organizations adopted GenAI earlier and more extensively; they therefore also **give higher priority to building out infrastructure** to support these new, power-dense workloads.

Despite the fact that GenAI is becoming more prevalent, very few organizations today are focused on moving from air to liquid cooling — which is both an economic and ecological issue.

GenAI is **rapidly increasing rack power density.** From under 15kW per rack in 2020, new GenAI racks are drawing upwards of 100kW per rack in 2025, and this will increase to over 500kW before 2030.

Key priorities relating to IT infrastructure operations over the next 18 months



Liquid cooling will be an essential component of the AI-enabled datacenter.

Source: IDC's *Survey Europe September 2025* (n = 300)

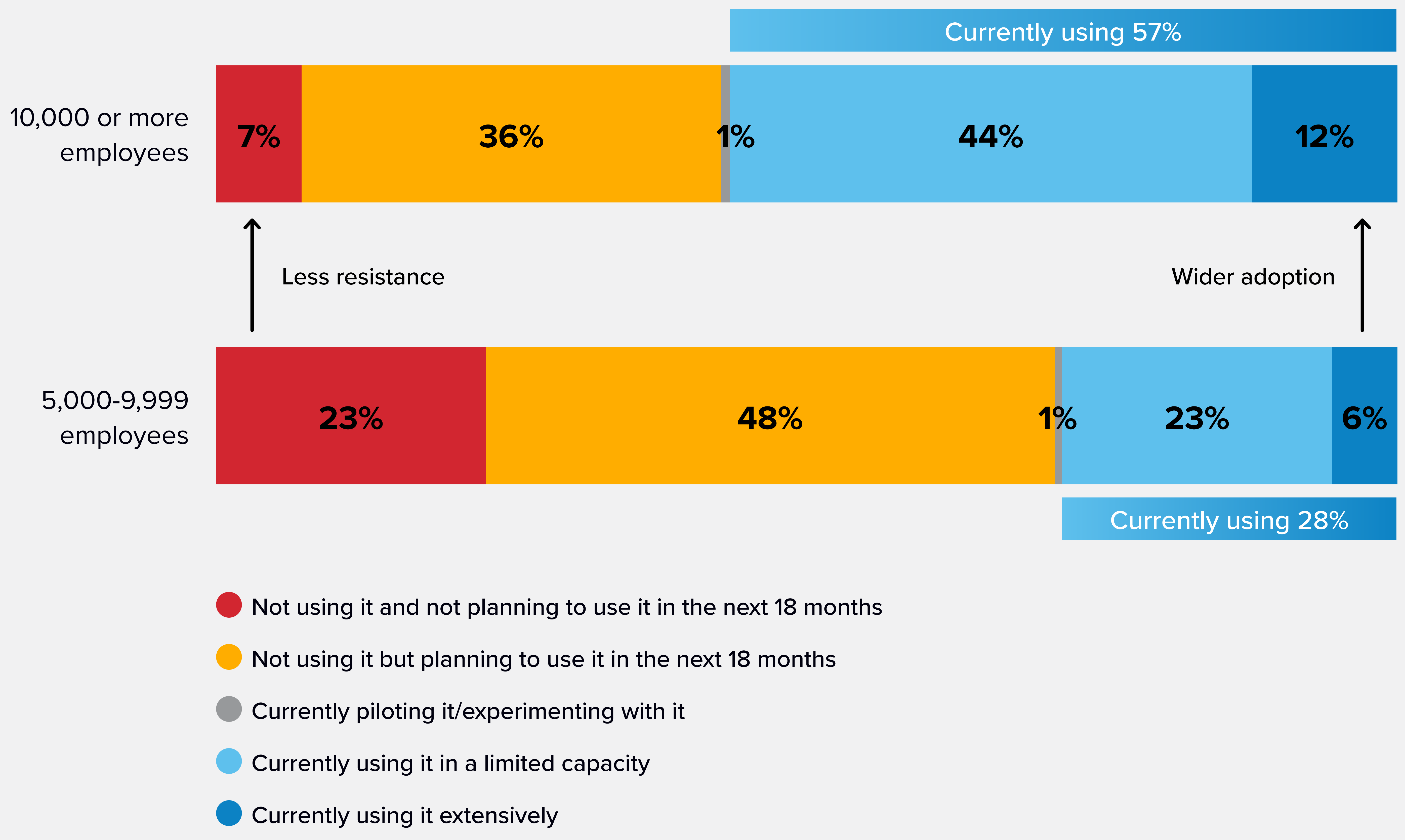
Large digital enterprises leading the GenAI frontier

GenAI is being pioneered by larger enterprises, particularly those above 10,000 employees where proofs of concept (PoCs) have moved into limited production more ostentatiously than in the case of companies with 5,000 to 9,999 employees.

Crucially, the largest enterprises have also accelerated into extensive use of GenAI in supporting the business at double the rate of the 5,000 to 9,999 employees' companies.

What these companies do today is a leading indicator of coming trends: **Companies of all sizes should be working now to be ready for the GenAI impact** on their digital infrastructure over the coming years.

Extent of AI or GenAI in production for business use cases



Source: IDC's *Survey Europe September 2025* (n = 300)

Shift in server buying patterns towards accelerated systems

The need for GenAI infrastructure will drive demand for specialized AI-accelerated servers.

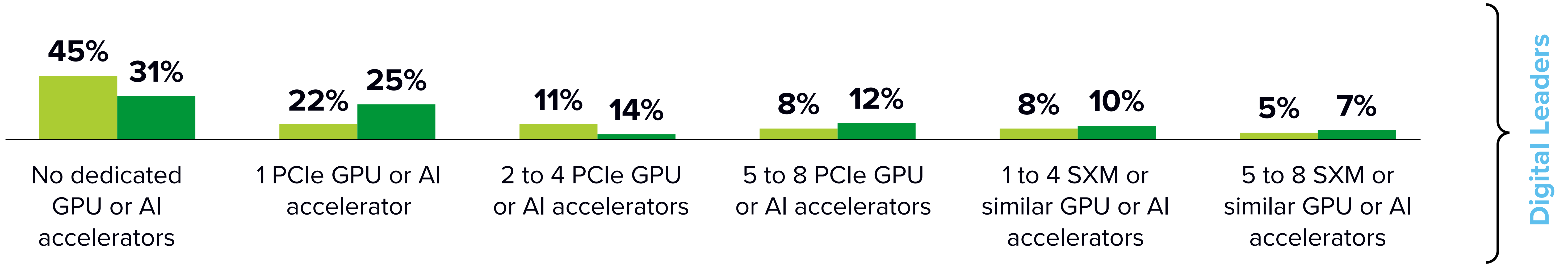
Servers with GPU-type or dedicated AI accelerators are becoming a bigger part of the IT infrastructure.

Already, respondents are saying that almost half of all their deployed servers are accelerated, with this number expected to rise significantly in the next two years.

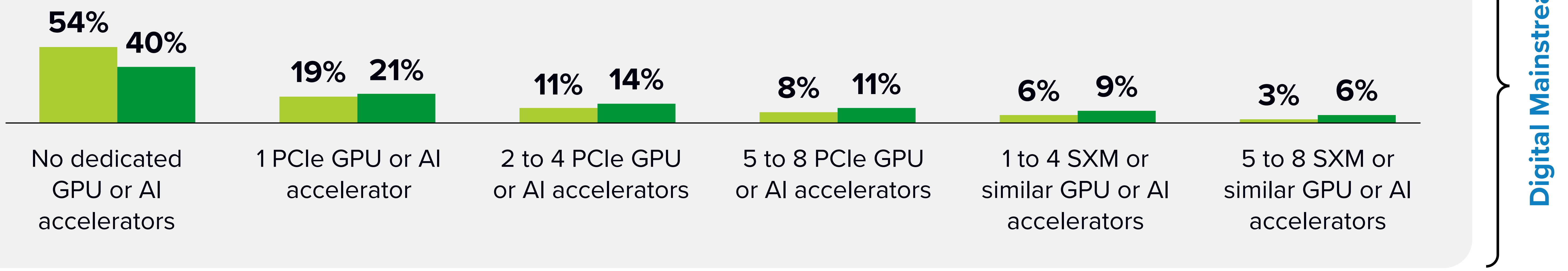
The adoption of AI accelerators generates significantly more heat, which will need to be carried away and dissipated — challenging the effectiveness of existing air-cooling technologies.

Servers deployed by type, 2025 vs 2027

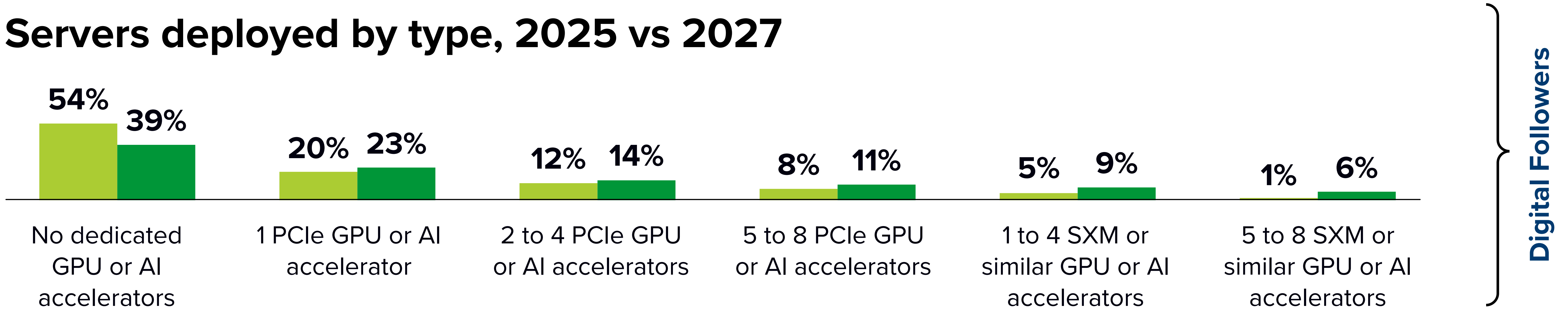
2025 2027



Servers deployed by type, 2025 vs 2027



Servers deployed by type, 2025 vs 2027



Source: IDC's *Survey Europe September 2025* (n = 300)

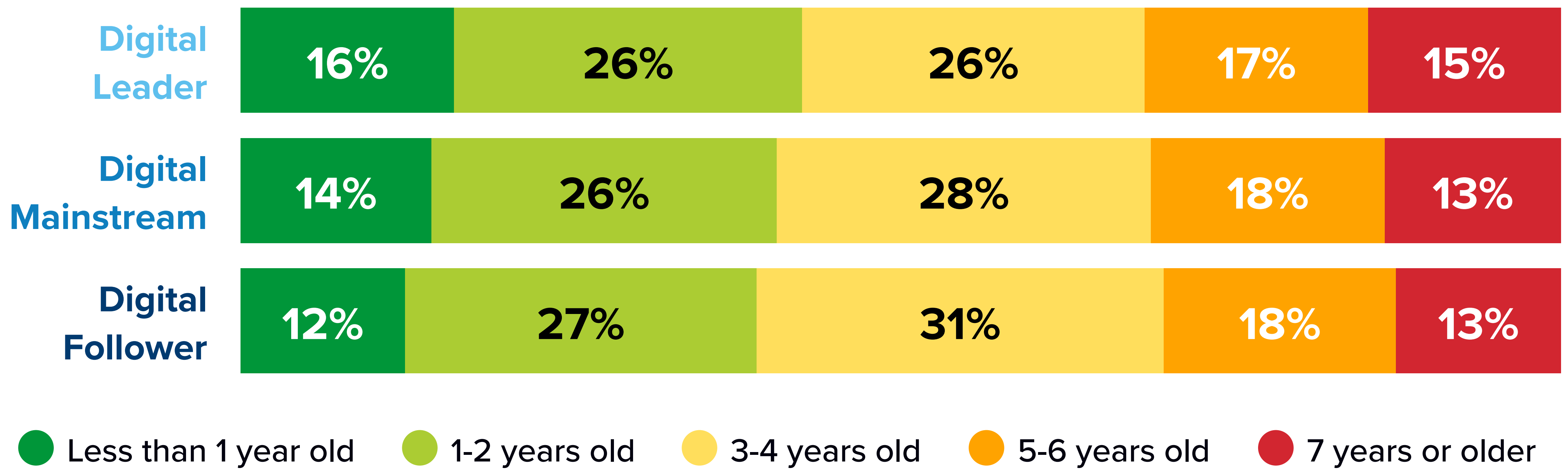
Value creation through infrastructure modernization

Many installed servers are three or more years old, and are often **under-utilized**, with many organizations running their servers — whatever their age — at 30% utilization or less.

They also have older generations of CPUs and GPUs that deliver lower performance per watt, meaning they consume a greater amount of power for minimal performance.

Proactively **modernizing IT infrastructure** with the latest generations of efficient **and** scalable servers, running at high levels of utilization using **automation and AIOps,** can result in a meaningful decrease in overall server **power consumption**.

Share of IT infrastructure by age band



Leading criteria for refreshing IT infrastructure hardware

1. Increased security
2. Better reliability
3. Better automation and orchestration
4. Direct-to-chip liquid cooling integration

The time to think about liquid cooling is now

Support for **liquid cooling** has yet to become a top selection criteria, but with coming **increases in rack power density,** this will become an increasingly critical factor when contemplating hardware refresh.

Source: IDC's *Survey Europe September 2025* (n = 300)

AI workloads driving demand for alternative cooling methods

Most datacenter IT racks have a power consumption of less than 15kW per rack, which has been the level of typical power consumption for some time.

With the move to CPUs with a high core count in the late 2010s, **power consumption per server started to increase noticeably** and the power density of IT racks increased.

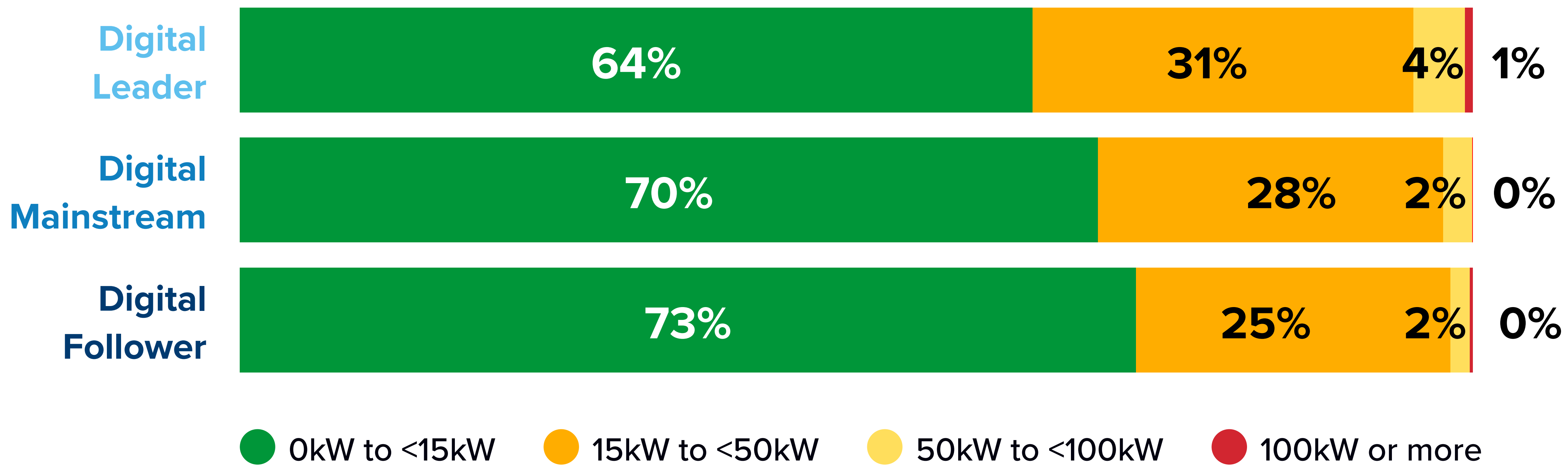
By 2022, the advent of **GenAI** and the rapid adoption of datacenter GPUs began to accelerate AI workloads.

The **high performance** of AI acceleration comes at the cost of **higher power consumption and density**, with rack power densities of over 100kW per rack starting to be deployed by Digital Leaders. The IT industry roadmaps show this will rise above 500kW per rack by 2030.

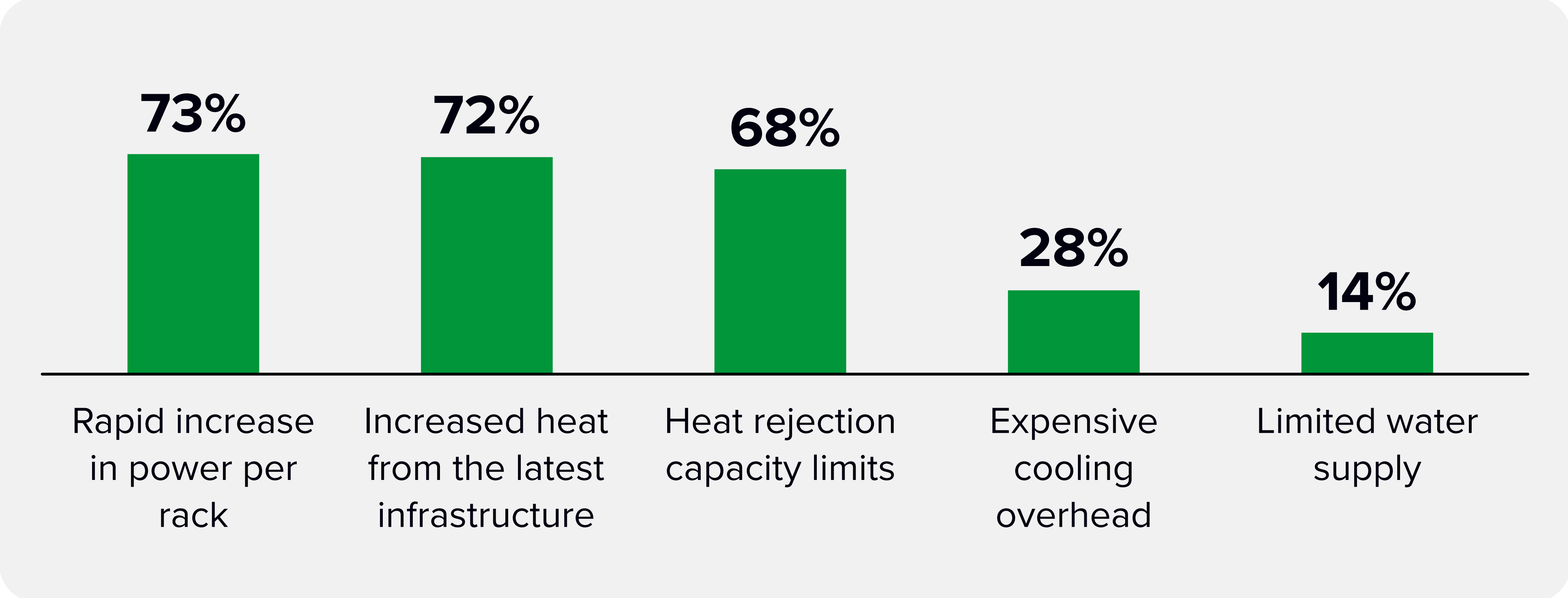
Key issues to consider:

- The rise in power consumption and density per rack means new and more effective datacenter cooling technologies will be required.
- Air cooling overheads in the datacenter that could be tolerated at under 15kW per rack become very noticeable and costly at 100kW+ per rack.

Proportion of datacenter racks per power band



Top datacenter power and cooling challenges



Source: IDC's *Survey Europe September 2025* (n = 300)

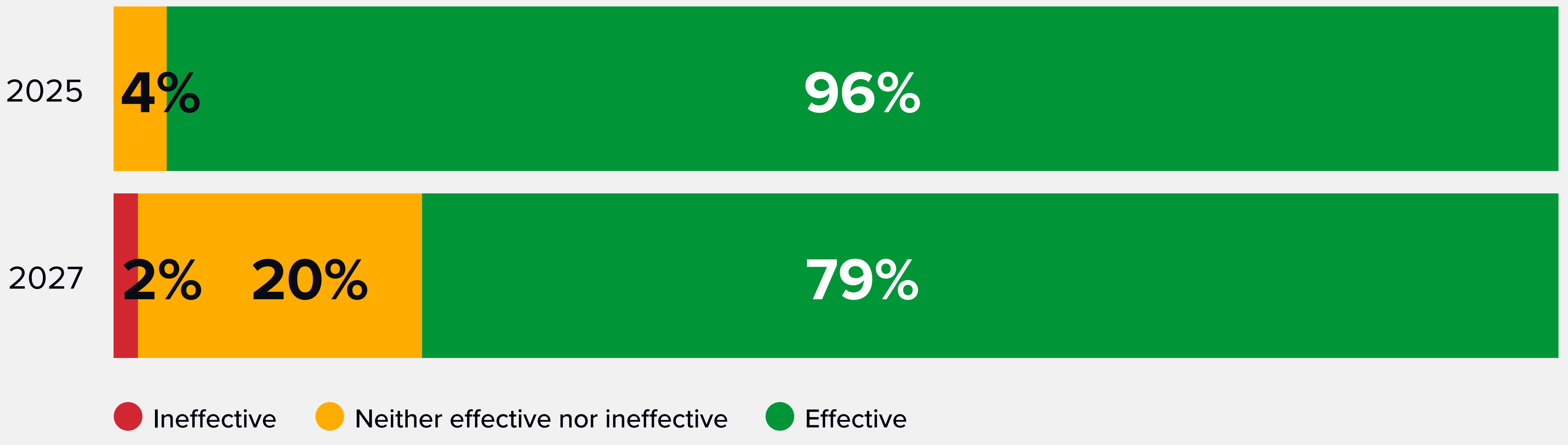
Overcoming resistance to liquid cooling

In 2025, most companies view **air cooling** as very or **adequately effective for their** datacenter cooling **needs**. This is understandable, as **most IT racks** still **have a power consumption of less than 15kW per rack.**

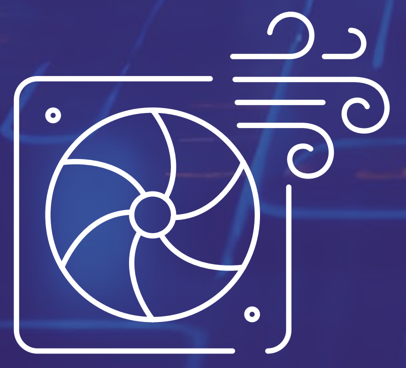
With the increase in rack power densities above 100kW per rack, this perception is starting to change. In 2025, only 4% of respondents cited that air cooling is no longer effective, while it is expected to increase to 20% in 2027. Critically, by 2027 companies are expected to start seeing air cooling as being ineffective for their needs.

New GenAI and agentic AI workloads will inevitably require increased amounts of power. **IT leaders and datacenter architects need to begin thinking today about their cooling solutions for tomorrow.**

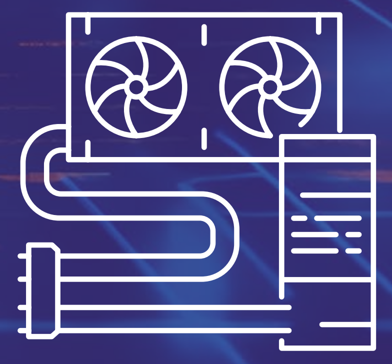
Level of effectiveness of air cooling approaches for organization's IT infrastructure in 2025 and in 2027



Key considerations:



Air cooling can cope with some increases in power consumption per rack, but will become obsolete as power density continues to climb.



New datacenters targeted at large-scale IT consolidation or AI use cases will require liquid rather than air cooling approaches.

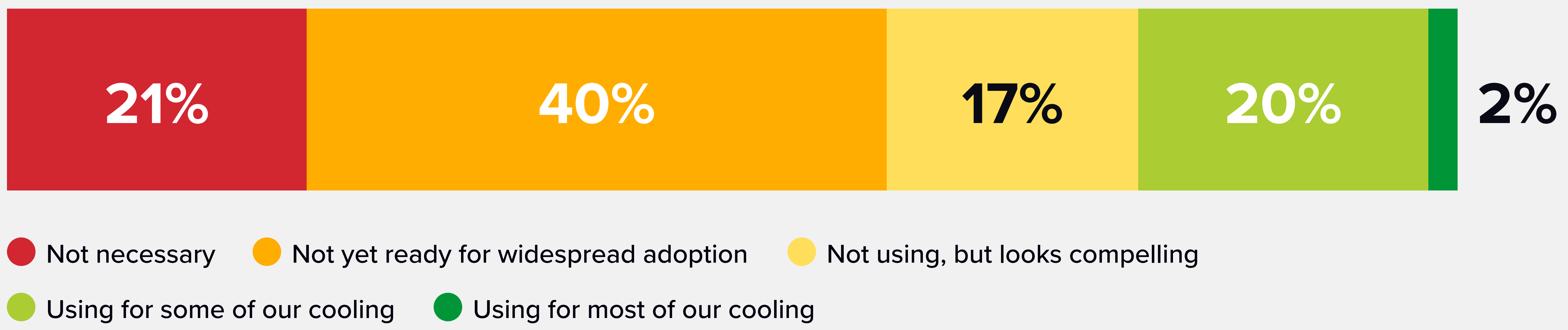
Source: IDC's *Survey Europe September 2025* (n = 300)

Making an ROI case for liquid cooling

Despite the many advantages of liquid cooling, companies often hesitate to invest due to limited hands-on experience and understanding.

Many organizations still struggle to see a clear ROI in switching from air to liquid cooling. However, as power densities rise, air cooling becomes increasingly costly. IT teams should proactively collect data and collaborate with vendors and partners to build a compelling business case for liquid cooling adoption.

Stance on direct-to-chip liquid cooling



Top 3 concerns regarding liquid cooling

- Lack of clear ROI
- Risk of damage in case of leak
- Environmental implications of liquids used

Many IT organizations worry about liquid cooling leaks and their impact on datacenter operations. However, modern direct-to-chip liquid cooling systems have matured significantly, featuring fewer hoses, more reliable quick connectors, and advanced leak detection sensors — making leaks less likely than fan failures in traditional air-cooled setups.

Source: IDC's *Survey Europe September 2025* (n = 300)

Direct-to-chip liquid cooling for power-dense infrastructure

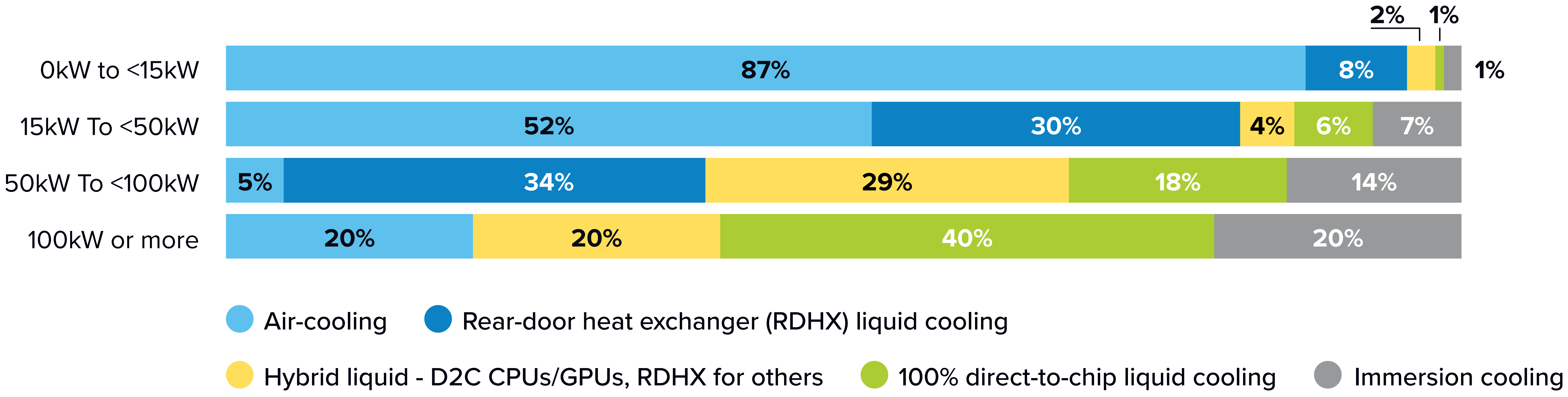
As rack power densities grow from under 15kW per rack to over 100kW, preferences are clearly shifting from air cooling towards various liquid cooling approaches.

Rear-door heat exchanger (RDHX) liquid cooling solutions are a good first step for many — but as power density rises even more, direct-to-chip becomes more popular.

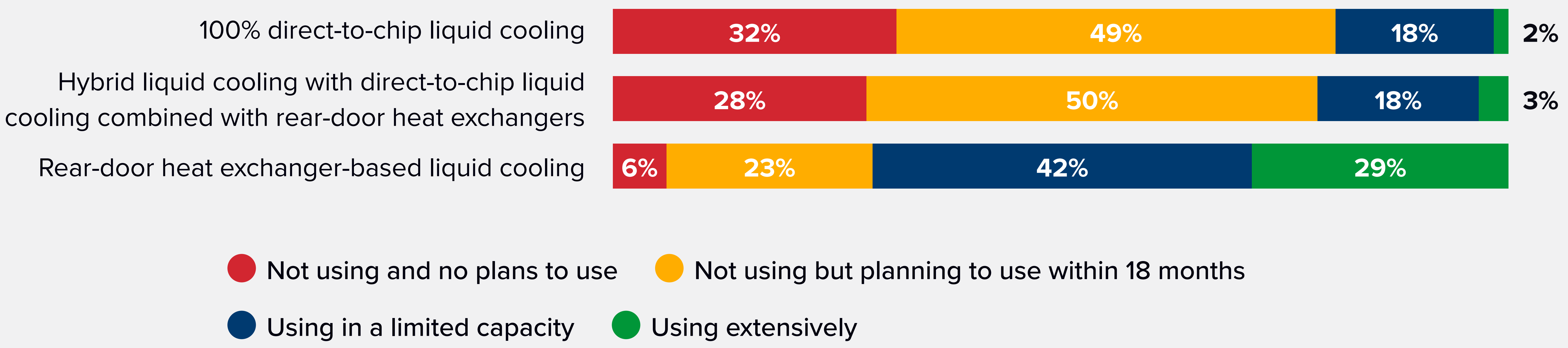
Initially, rear-door heat exchangers are deployed in a hybrid manner, but at 100kW or more there is typically a big jump in the use of 100% direct-to-chip liquid cooling architectures.

Immersion cooling will be important for certain specific but niche high-power workloads.

Most relevant cooling approaches by rack power densities



Use of different liquid cooling technologies in datacenters



Source: IDC's *Survey Europe September 2025* (n = 300)

IDC guidance

Building on research findings, IDC suggests that customers should follow these recommendations to enhance power efficiency and fulfil requests from relevant stakeholders for the IT department's technology requirements.

1

The time to invest in GenAI is now. Large enterprises are driving the first wave of GenAI adoption and are reaping the benefits that extend far beyond productivity increases to include AI-services-related revenue growth, revenue protection, and less costly manufacturing and warranty support.

2

Datacenter design and cooling strategies need to be re-engineered. Power demands will keep growing with GenAI, with rack power consumption exceeding 100kW in 2025 and in line to surpass 500kW by 2030. Organizations must rethink their designs to ensure they are fit for future needs.

3

Prioritize direct-to-chip liquid cooling. If your datacenter plans involve power densities of over 50kW per rack, you should focus on incorporating direct-to-chip liquid cooling where possible.

4

Liquid cooling investment should be a no-brainer. The cost of air cooling is increasing rapidly, and many companies do not track or report these overheads, making building an ROI case much more challenging. Organizations should start tracking early and using the data to help build a business case for investment in liquid cooling.

Message from the sponsors

AI factories from Supermicro and NVIDIA are complete, turnkey solutions that make deploying AI at any scale easier, leading to quicker time-to-online and time-to-revenue. These full-stack solutions include the latest compute, software, networking, and storage, and are first-to-market based on NVIDIA Enterprise Reference Architecture Designs. They also feature proven Supermicro rack-level validation and deployment services for complete AI confidence.

For more information

About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications, and consumer technology markets.

With more than 1,300 analysts worldwide, IDC offers global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries. IDC's analysis and insight help IT professionals, business executives, and the investment community to make fact-based technology decisions and to achieve their key business objectives.

Founded in 1964, IDC is a wholly-owned subsidiary of International Data Group (IDG, Inc.), the world's leading tech media, data, and marketing services company.

IDC Custom Solutions

This publication was produced by IDC Custom Solutions. As a premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications, and consumer technology markets, IDC's Custom Solutions group helps clients plan, market, sell, and succeed in the global marketplace. We create actionable market intelligence and influential content marketing programs that yield measurable results.

IDC UK
1st floor, Whitfield Street, London, W1T 2RE, United Kingdom
T 44.208.987.7100

@idc | @idc | idc.com