



AI-ready Hybrid Cloud

REPORT

The State of AI Infrastructure Report 2025





Table of Contents

Methodology and Research Goals	3
Summary	4
Key Findings.....	5
Conclusion	30
Recommendations.....	31
About A10 Networks	32

Methodology and Research Goals

Enterprise adoption of artificial intelligence (AI) is increasing, but infrastructure is not currently set up to handle the workload. Obstacles to greater use of AI include security, performance, scalability, and visibility. This report is based on a survey of IT professionals across industries, including technology, healthcare, financial services, manufacturing, and education. Respondents span organizational levels from staff to C-level executives, with roles in IT operations, cybersecurity, infrastructure, and architecture.

The survey explored:

- ✓ Current AI adoption and workload types
- ✓ Emerging AI priorities and strategic approaches
- ✓ Infrastructure hosting environments
- ✓ Performance and scaling capabilities
- ✓ Infrastructure pain points and modernization plans
- ✓ Security, visibility, and data governance challenges

Summary



GenAI is Becoming Mainstream

76 percent of organizations are already using GenAI applications such as LLMs, chatbots, and coding assistants

GenAI is no longer experimental, it's rapidly becoming a core part of enterprise traffic



Hybrid Cloud is the Preferred Model

42 percent use hybrid cloud for AI workloads, balancing scalability with control

Public cloud is the top choice for hyper-scaling AI workloads



Infrastructure is Lagging Behind AI Adoption

79 percent of organizations plan to modernize infrastructure within 18 months

Only 19 percent have automated scaling for AI workloads

53 percent are only "somewhat confident" in their infrastructure's ability to support future AI needs



Security, Performance, and Visibility are top Challenges

49 percent cite security as the biggest pain point

Visibility gaps and compute limitations are widespread

Monitoring and governance for AI workloads remain underdeveloped

Key Findings

AI Adoption is Strong, Diverse, and Growing

Organizations are actively embracing AI, with one-third of survey respondents reporting that they have already deployed AI and machine learning (ML) in production for select applications. Forty-five percent are currently “experimenting or running pilot projects with AI/ML,” followed closely by 33 percent who are running AI/ML in production. The 13 percent who are not currently using the technology, are planning to do so within the next 18 months.

Current Use of AI

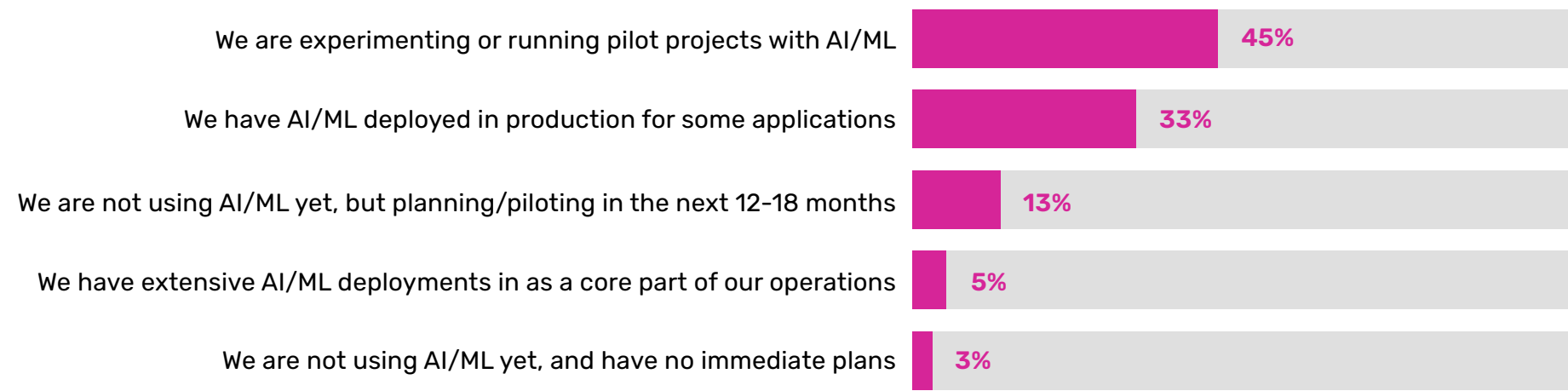


Figure 1 – Responses to the question, “Which statement best describes your organization’s current use of AI/ML?”

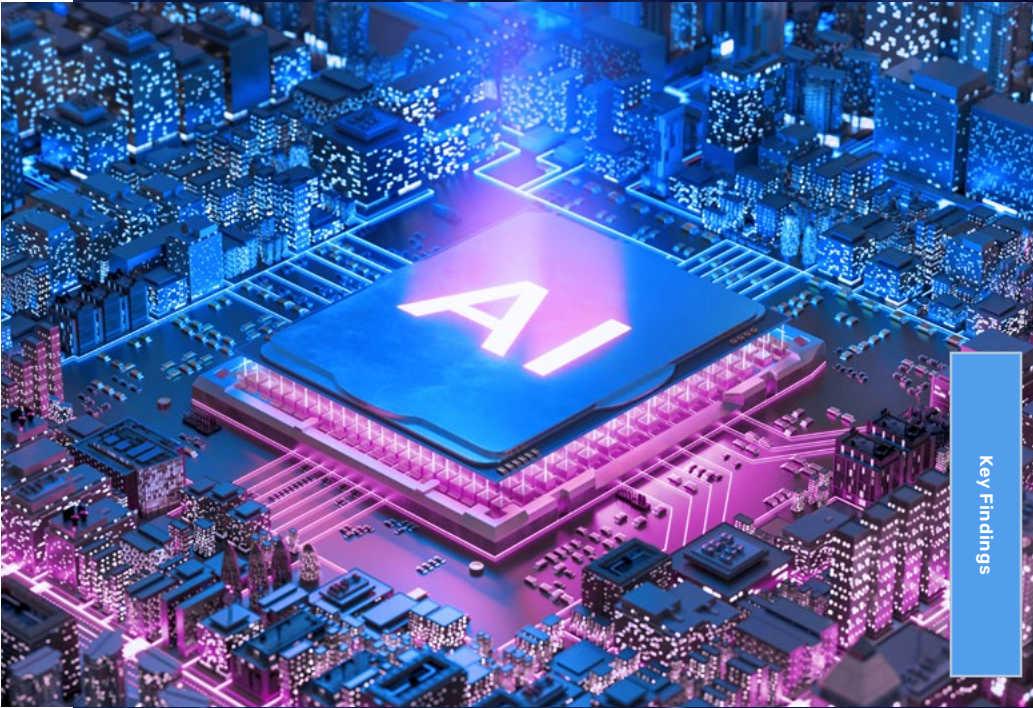
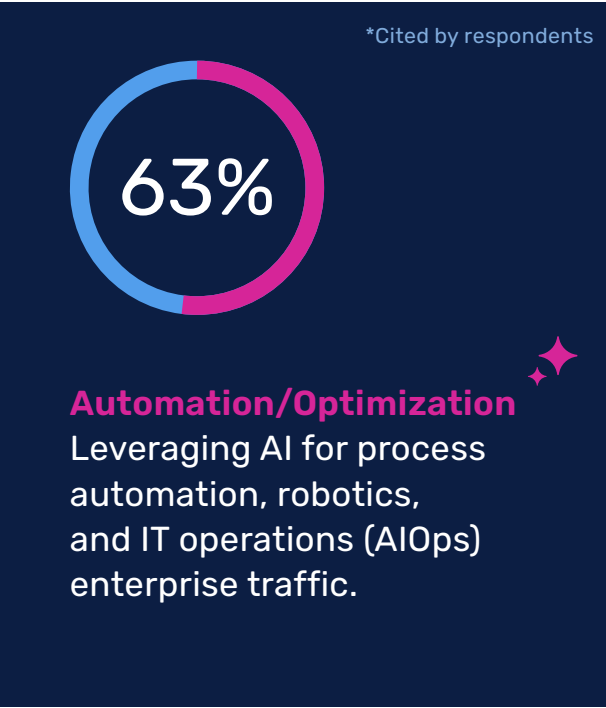
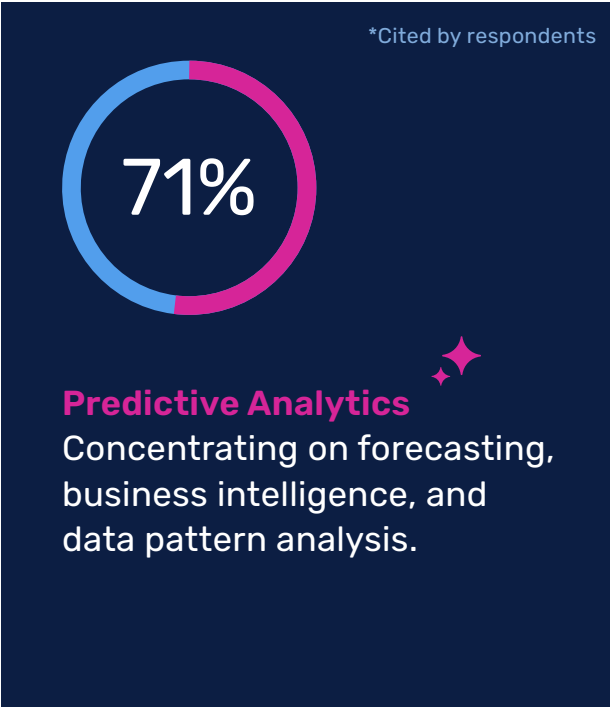
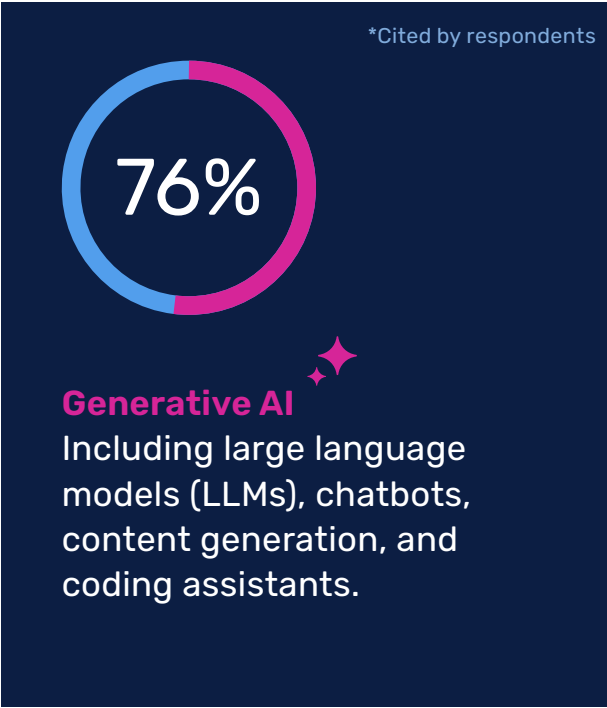
33%

of survey respondents reported that they have already deployed AI and machine learning (ML) in production for select applications.



Key Findings

Three Workloads Dominate



Computer vision lagged, with just 28 percent adoption, suggesting that this workload remains a more specialized or industry-specific use case.

Adoption varied by industry. Technology/software was the biggest adopter of GenAI, at 80 percent. This makes sense given the strategic importance of AI for companies in that field. Seventy-one percent of financial services firms surveyed are using AI for predictive analytics, which aligns with that sector's emphasis on using technology to gain competitive advantage. Adoption was also high in manufacturing, with 75 percent of firms using AI for predictive analytics, production optimization, and maintenance. Given tight margins and fast-paced production schedules in that industry, it's not surprising that manufacturers are leaning into AI to cut costs and avoid production outages.

42%

of respondents favor a balanced approach between on-premises and cloud infrastructure.



AI Hosting and Factors that Affect Infrastructure Choices

Edge facilities, which have been somewhat hyped for AI, represent just 4% of AI hosting.

Organizations are increasingly adopting hybrid cloud environments to host AI workloads, with 42 percent of respondents favoring a balanced approach between on-premises and cloud infrastructure. Drivers of hybrid cloud hosting include latency, availability, and performance. Hosting is divided between cloud and on-premises infrastructure. This is a logical choice, reflecting a split between workloads that require ultra-low latency, which favors on-premises infrastructure, and less performance-sensitive workloads, which can be handled adequately in the cloud. Public cloud hosting is the second most common choice, selected by 35 percent of respondents, largely due to its cost-effectiveness and flexible scalability.

Only 15 percent of organizations rely primarily on on-premises data centers. Edge facilities, despite being touted as ideal for AI, are used by just 4 percent. This limited adoption reflects the complexity and resource demands of deploying AI at the edge, underscoring the importance of robust infrastructure planning as GenAI workloads scale.

Primary AI Hosting Environments

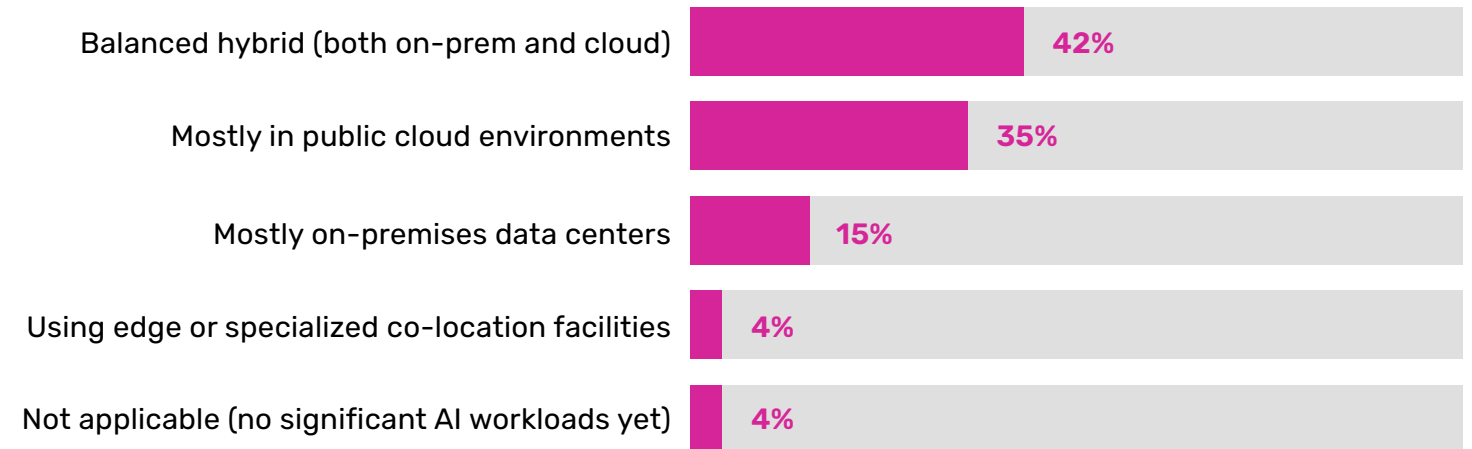


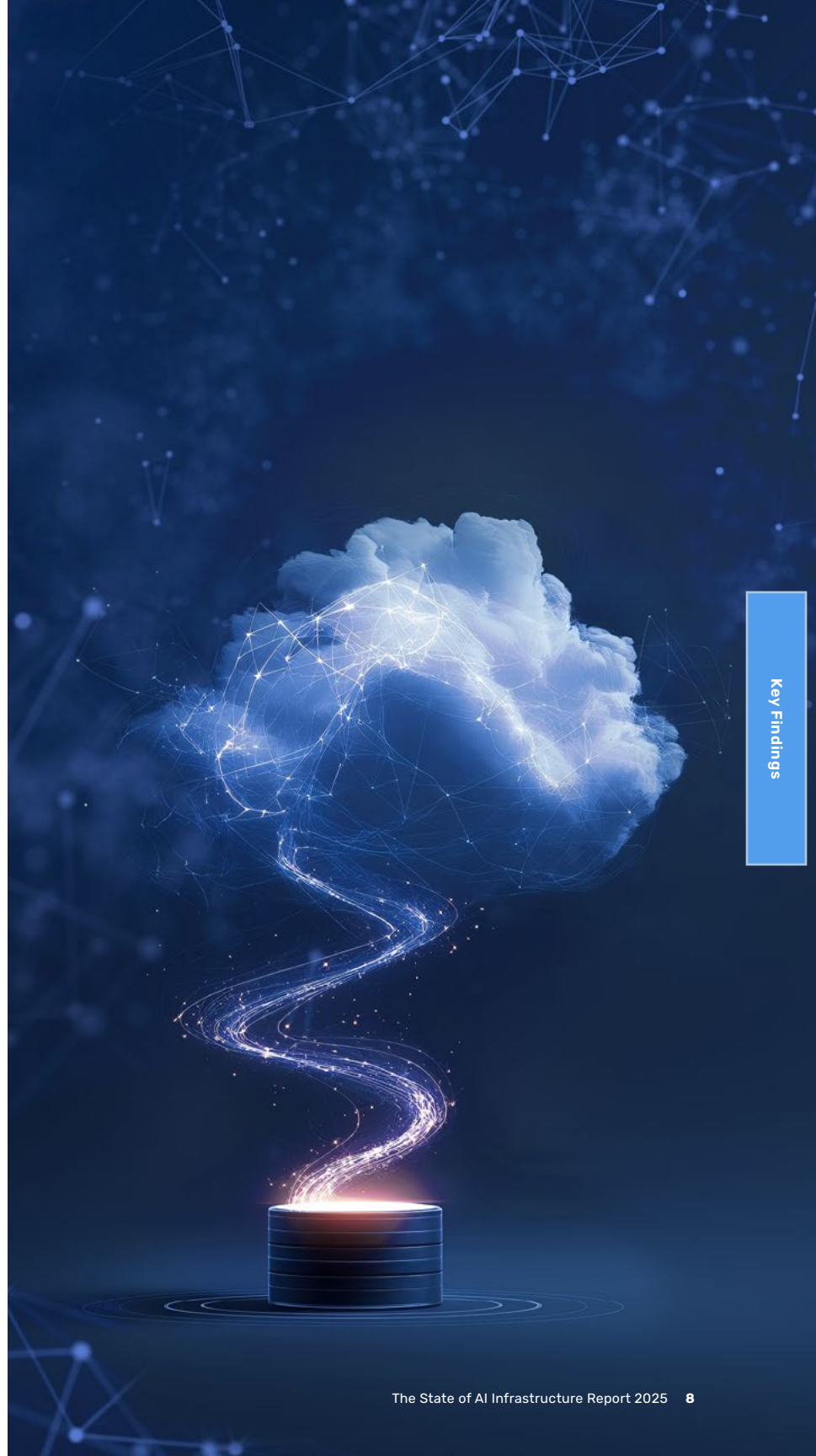
Figure 2 – Responses to the question, “Where are your AI workloads primarily hosted or run?”

Tech Companies Favor Hybrid Clouds

Forty-nine percent of technology companies host workloads in a hybrid cloud versus 38 percent in a public cloud. AI use cases in tech range from high-performing products in production, which may fit best with on-premises hosting, to development and early adopter offerings, which are suited to the cloud.

Thirty-one percent of manufacturing companies host AI on-premises. These businesses often require extremely fast response from AI, for example, for analysis of safety cameras in real time, so on-premises is the best option.

Fifty-eight percent of healthcare organizations favor hybrid approaches. Regulatory requirements, such as HIPAA, probably drive this behavior.



49%

of the respondents say
that AI application performance
is "important."



Low Latency and High Availability Matter

Nearly half of the respondents said that AI application performance is "important," while another 23 percent rate it as "extremely important."

This emphasis tracks with the strategic focus on AI. With AI integrated in business operations like customer support and manufacturing safety assurance, it is logical that users want close to zero downtime. That said, it appears that certain workloads are less critical. Twenty-two percent of respondents said that performance and latency were "moderately important," partly because their AI workloads involved batch processes.

Importance of Low Latency and High Availability

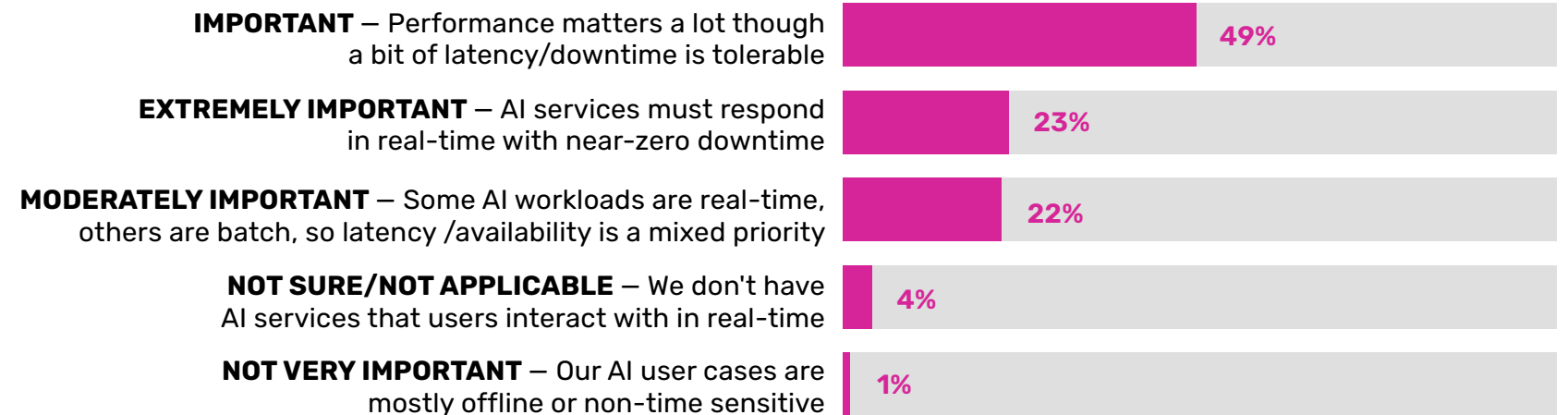


Figure 3 – Responses to the question, "How important are low latency and high availability for the AI-enabled applications in your organization?"

Application Delivery Infrastructure can Mostly Maintain AI Performance

When asked whether their current application delivery infrastructure such as load balancers and application delivery controllers can maintain the required performance and uptime for AI workloads, 50 percent of respondents said “mostly.” This suggests that while their systems generally meet performance expectations, they occasionally approach capacity limits or experience minor issues.

Another 22 percent reported that their infrastructure is “not entirely” sufficient, citing performance bottlenecks and availability challenges. Only 16 percent expressed full confidence, stating their infrastructure meets AI workload demands with capacity to spare. Meanwhile, 8 percent were unsure or found the question not applicable, and 3 percent acknowledged that their infrastructure is already struggling to keep up with AI-driven traffic and scaling needs.

These findings underscore the growing pressure on application delivery systems and the need for modern, scalable infrastructure to support the performance demands of GenAI workloads.

Ability of Current App Delivery Infrastructure to Handle AI Workloads

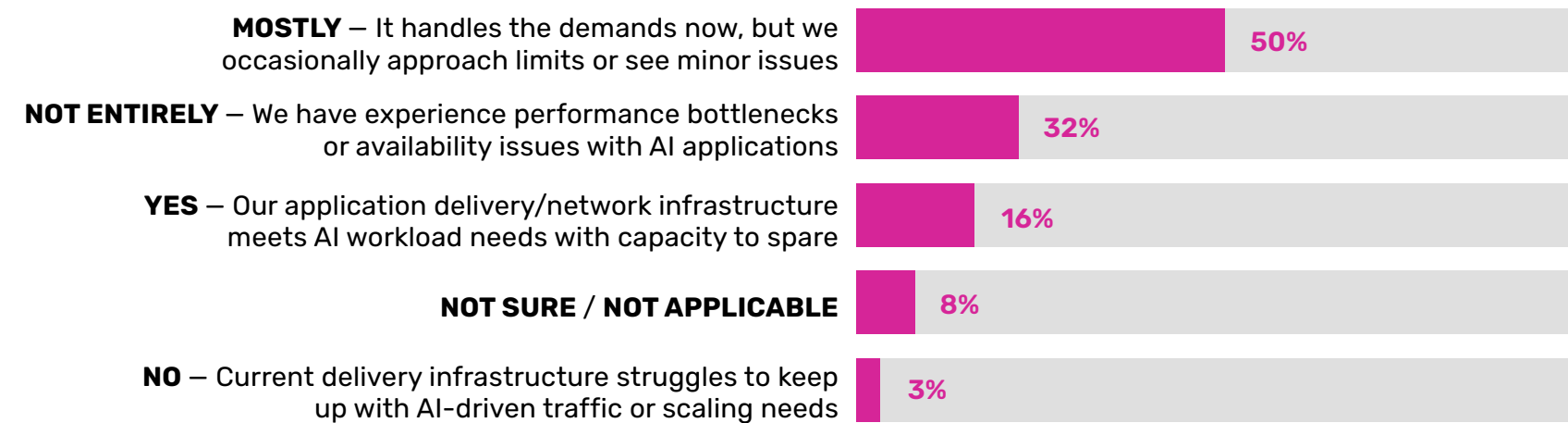


Figure 4 – Responses to the question, “Is your current application delivery infrastructure (e.g., load balancers, application delivery controllers) able to maintain the required performance and uptime for your AI workloads?”

50%

of survey respondents said their current application delivery infrastructure is “mostly” able to maintain the required performance and uptime for their AI workloads.



37%

of the respondents say
when faced with the need for sudden
and massive AI scale-out, many companies
are turning to the public cloud.

They are planning to use it as a
key strategy for rapid expansion.



The Public Cloud is the “Go-to” for a Massive AI Scale-out

When faced with the need for sudden and massive AI scale-out, many companies are turning to the public cloud, with 37 percent of respondents planning to use it as a key strategy for rapid expansion. A hybrid cloud approach, which maximizes on-premises capacity, comes in second, with 33 percent. These approaches make sense given the public cloud’s infinite, on-demand capacity. Scaling out on-premises infrastructure most often cannot be done rapidly and would require a significant financial commitment. In fact, just 18 percent said they will scale out AI on-premises.

Responding the Need to Hyperscale

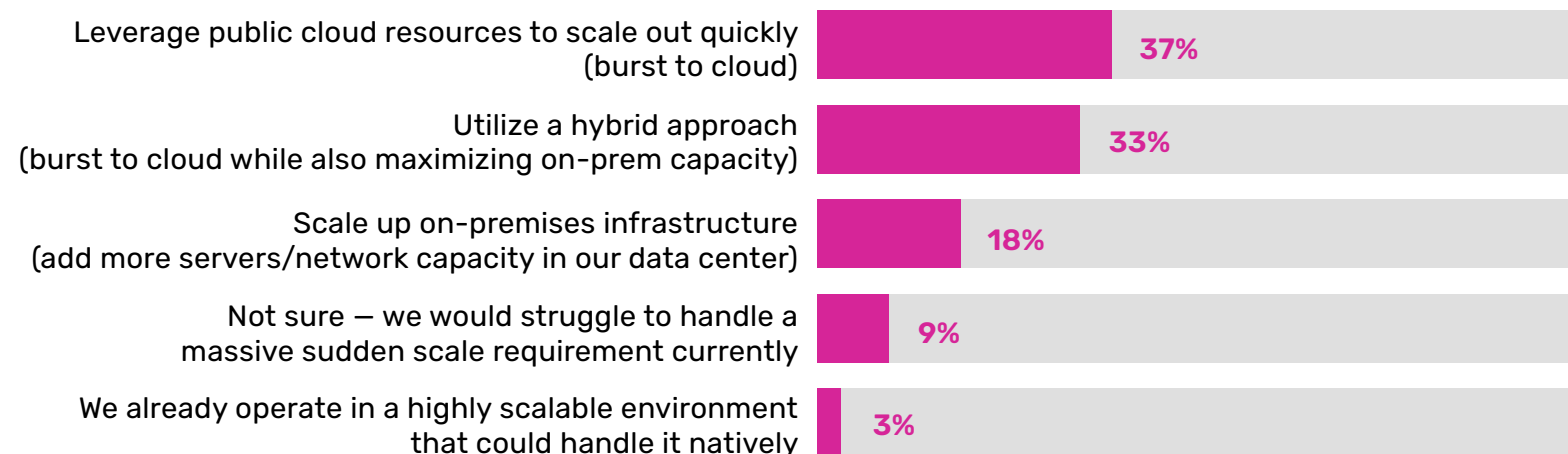


Figure 5 – Responses to the question, “If you suddenly needed to hyperscale (massively scale out) your AI workloads to meet a surge in demand, what would be your primary approach?”

Auto-scaling is Rare for AI Workloads

Continuing the topic of scaling, respondents were asked about their use of auto-scaling as a solution for situations where they needed more AI infrastructure capacity. More than six in ten were only using auto-scaling “partially” or “not yet.” The 35 percent that are “partial” users of auto-scaling rely on manual intervention at least some of the time. These findings point to the fact that automated scaling is challenging to implement.

The needs are real, however. While 71 percent of organizations are already using or experimenting with AI, only 19 percent have fully automated scaling for AI workloads, creating a potential bottleneck as AI workloads grow. For 10 percent, all scaling is manual or uses fixed capacity.

Use of Auto-scaling

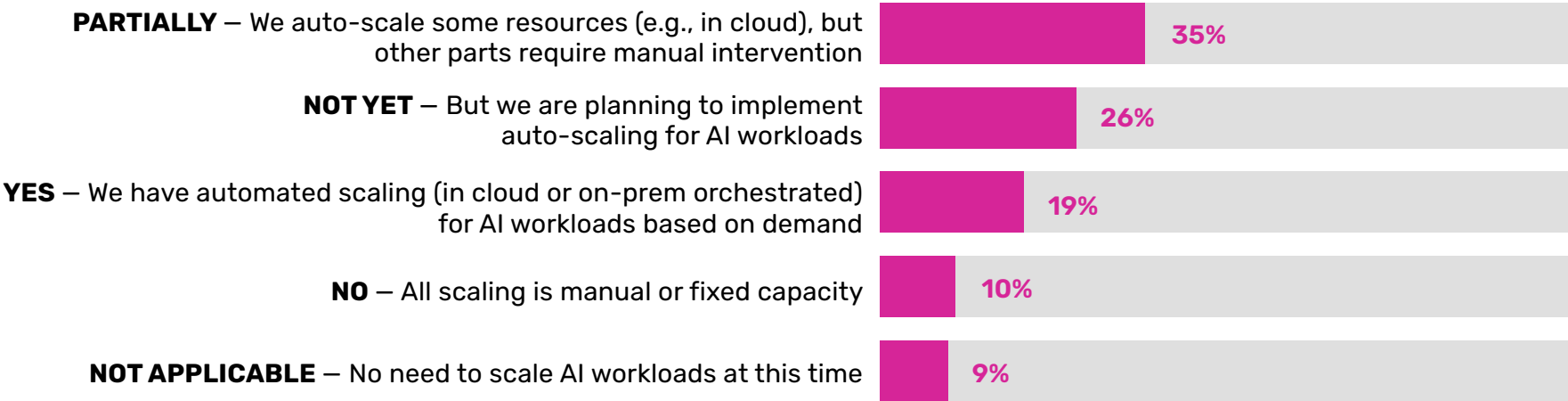
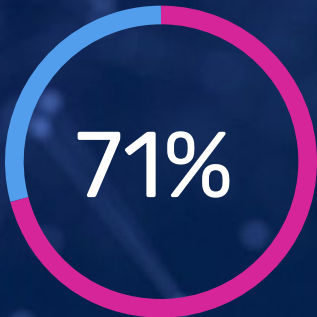
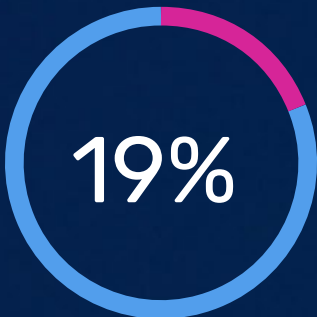


Figure 6 – Responses to the question, “Do you currently use any form of auto-scaling or on-demand scaling for your AI applications/infrastructure?”



of organizations are already using or experimenting with AI,



have fully automated scaling for AI workloads, creating a potential bottleneck as AI workloads grow,



all scaling is manual or uses fixed capacity.

79%

were either “somewhat confident” or “not very confident in the ability of their current AI infrastructure to meet the demands of AI-driven applications in the next two to three years.



Confidence in Current AI Infrastructure is Relatively Low

Respondents don’t seem to have a lot of confidence in the ability of their current AI infrastructure to meet the demands of AI-driven applications in the next two to three years. Seventy-nine percent said they were either “somewhat confident” or “not very confident.” Those who were “somewhat confident” believed they needed upgrades to handle the performance, latency, and availability demands of AI-driven applications in the future. Only 13 percent said they were “very confident.” These findings should not be surprising, given that AI-driven workloads are still new and evolving and infrastructure is catching up.

Level of Confidence in Network and Infrastructure

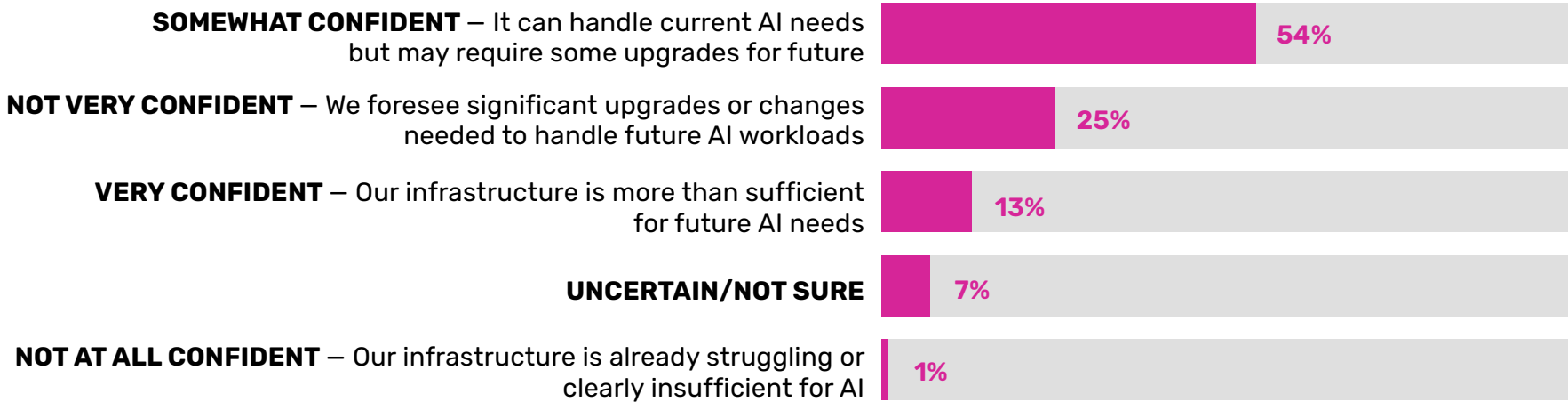


Figure 7 – Responses to the question, “How confident are you that your current network and application infrastructure can meet the performance, latency and availability demands of AI-driven applications over the next 2–3 years?”

“Security is our biggest concern, specifically data leakage.”



Enterprises Face Wide Ranging Infrastructure Challenges in Supporting AI Workloads

What challenges are organizations facing with AI infrastructure? Respondents report several concerns, though security constraints dominate all organizational levels and industries. Forty-nine percent said that security is their biggest limitation for current infrastructure to support AI workloads. They feel that current security appliances and solutions cannot effectively protect or inspect AI-related traffic. This finding presents something of a paradox. While 49 percent of organizations identify security constraints as their biggest infrastructure limitation, only 40 percent have deployed new AI-specific security solutions.

Biggest Pain Points

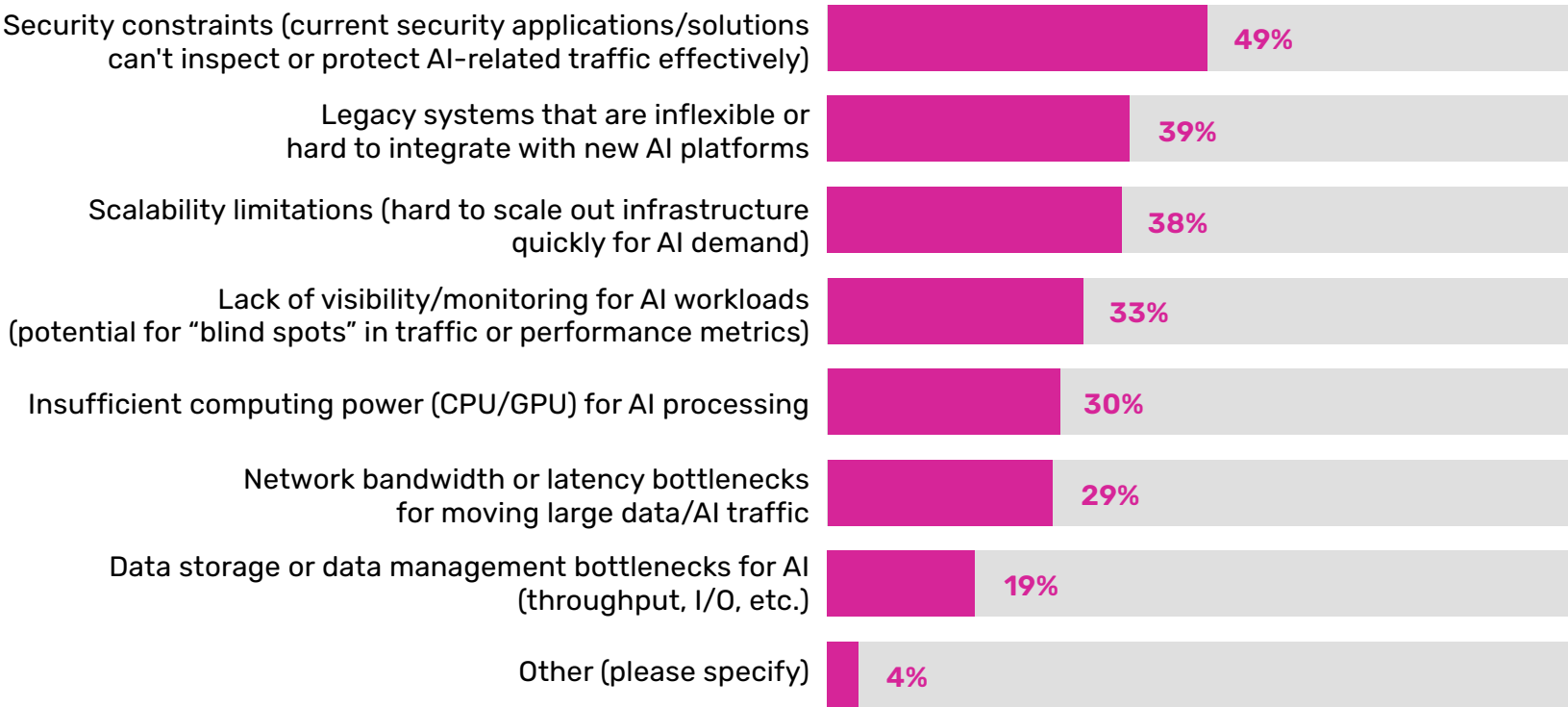


Figure 8 – Responses to the question, “What are the biggest limitations or pain points in your current infrastructure when it comes to supporting AI workloads?”

Security emerged as an area of concern for AI workloads in the analysis of open-ended text responses.



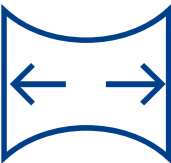
Notable issues included:

“Security is our biggest concern, specifically data leakage”

“Potential for AI to access sensitive systems or data”

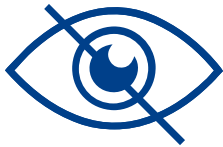
“Inability of current security tools to understand AI traffic patterns”

Infrastructure Challenges in Supporting AI Workloads



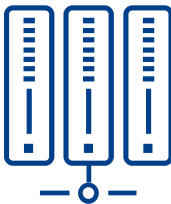
Scalability

Scalability is perceived as a limitation by 38 percent. It is viewed that it will be difficult to scale infrastructure quickly to meet fluctuating AI demands. The same number (38%) is concerned about legacy system integration, claiming that inflexible or incompatible legacy systems are impeding AI adoption. The data-intensive nature of AI partially explains this worry. Extracting data from legacy systems to train AI models can be a challenging and resource-intensive process.



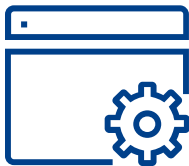
Visibility Gaps

Visibility gaps affect 33 percent of organizations, primarily due to insufficient monitoring and analytics for AI workloads. Although existing infrastructure may support AI, many organizations lack properly configured tools to track AI-specific activity. Key concerns include limited insight into how AI accesses data, poor visibility into model performance and drift, and difficulty monitoring resource usage.



Data Management

Data governance and management challenges are a top concern for AI workloads. Many organizations struggle with handling data volume, ensuring quality, and securing sensitive information. While infrastructure efforts often focus on compute and networking, the foundational task of managing the data that powers AI is frequently overlooked. Common concerns include the use of real versus synthetic data, scaling and capacity limitations, and risks of data leakage during training.



Resource Management

Resources also emerged as an area of concern, based on the comments below:

- “Ever growing demand of compute and network resources”
- “Unpredictable spikes in AI workload requirements”
- “Difficulty predicting capacity needs for scaling AI initiatives”

Infrastructure Challenges in Supporting AI Workloads

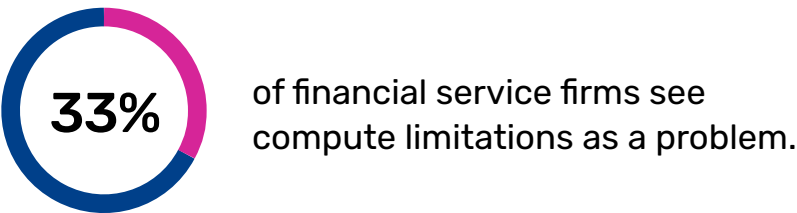
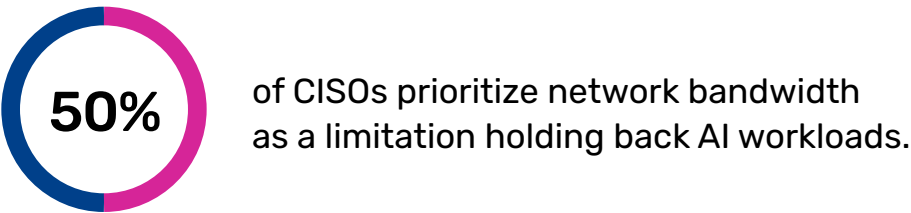


Compute, Network, and Storage

Compute, network bandwidth, and storage capacity emerged as limitations for 30 percent, 29 percent, and 19 percent, respectively, of respondents. They lack adequate CPU/GPU resources for AI processing, and legacy networks and storage solutions are also not up to the job. This may represent a matter of catching up, with organizations planning to add compute, network, and storage to support AI through their ongoing infrastructure upgrade cycles.

Roles affect points of view on these issues:

- “CISOs prioritize network bandwidth (50%) more highly than other roles”
- “VPs/SVPs emphasize visibility gaps (62%) more than other roles”
- “CIOs/CTOs place greater emphasis on compute limitations (67%)”



Performance Bottlenecks

Perspectives about performance bottlenecks reveal how AI challenges current infrastructure, which was not set up to handle this performance- and data-intensive workload. Compute limitations like CPU/GPU processing power again emerge as a factor that holds back AI performance. A third of respondents cited computational limitations as a performance bottleneck, followed by 23 percent who said inefficient application architecture and software constraints get in the way of AI performance. Network latency is an issue for 15 percent. Memory and storage inputs/outputs per second (IOPs) do not appear to be serious problems. Just nine percent of respondents called it a bottleneck. However, storage will soon be a future concern (Figure 8).

Biggest Performance Bottlenecks

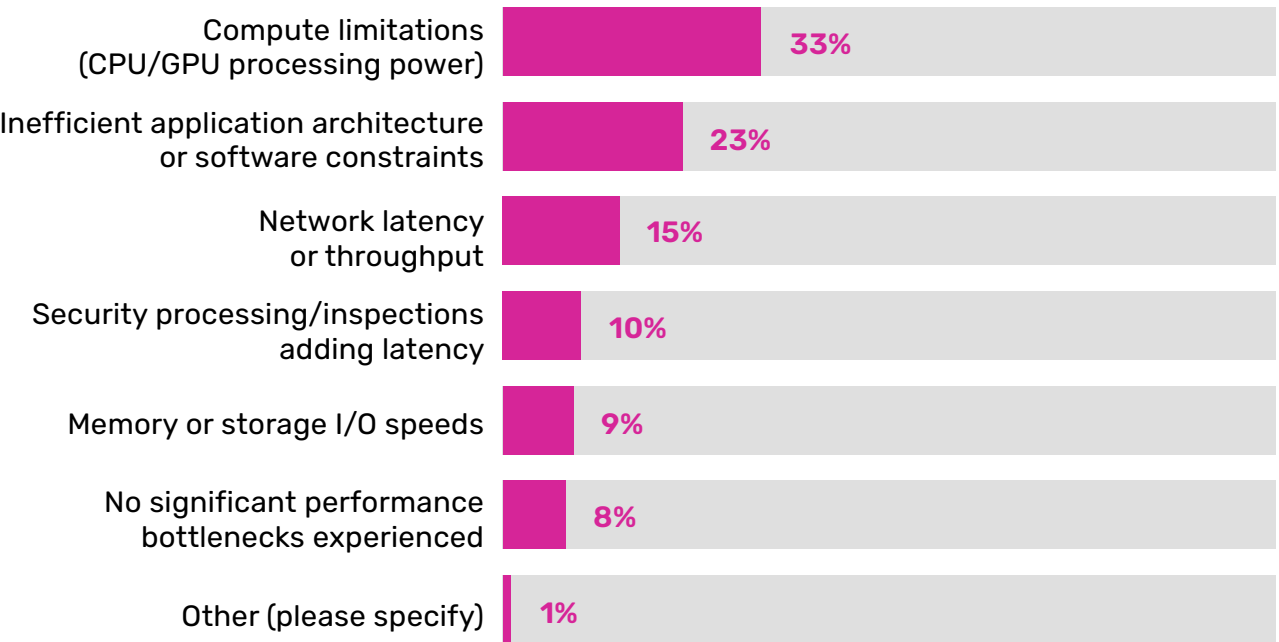
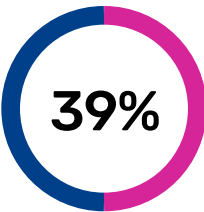


Figure 9 – Responses to the question, “What do you consider the biggest performance bottleneck for AI workloads in your environment today?”

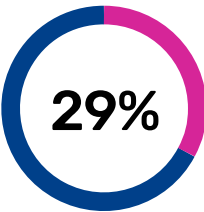
Infrastructure Challenges in Supporting AI Workloads

While compute is the most visible bottleneck, underlying causes of performance issues in AI workloads span the entire infrastructure stack. The best solution, as a result, will be one that takes a holistic approach to infrastructure modernization rather than focusing solely on adding more compute resources.

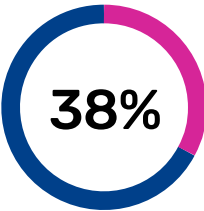
Bottleneck Perceptions Vary by Industry:



of technology companies and 34 percent of financial services firms see compute limitations as the problem. This is likely to be a temporary problem as these industries modernize their AI infrastructures.



in healthcare perceive inefficient application architecture to be the bottleneck. Regulatory restrictions and the need to integrate different system types, such as electronic health records (EHR) with imaging solutions, etc., explain this point of view.



of educational institutions blame security processing overhead. This sector is typically under-resourced, so it makes sense that security processing would inhibit AI workload performance.



Compute is the most visible bottleneck, but underlying causes of performance issues in AI workloads span the entire infrastructure stack.



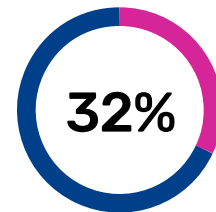
42%

of enterprises are using high-performance networking, including dedicated high-bandwidth links and a low-latency network fabric to optimize their AI application delivery and performance.

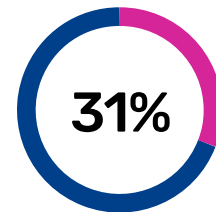


How Enterprises are Currently Addressing their AI Infrastructure Pain Points

Organizations that took the survey are implementing various technologies to address their AI infrastructure pain points. For example, to optimize AI application delivery and performance, 42 percent are using high-performance networking, including dedicated high-bandwidth links and a low-latency network fabric. Nearly 40 percent of the respondents are using advanced load balancing, comprising application delivery controllers tuned for AI traffic patterns.



of organizations are using content caching to reduce latency,



of organizations use edge computing to put AI computing and storage infrastructure closer to users.

This last finding would seem to contradict that only 4 percent said they host AI workloads at the edge (Figure 2). The discrepancy may be due to varying definitions of the “edge” or the broader answer choice in this question, “Edge computing or deploying AI services closer to users/data sources to cut down latency.”

33%

of organizations
cite compute as the biggest
performance bottleneck.



Interestingly, while 33 percent cite compute as the biggest performance bottleneck (Figure 9), 27 percent are using specialized hardware like a GPU or accelerator offline to improve throughput. This divergence may be related to timing, with 50 percent of organizations planning to deploy a compute performance solution in the future (Figure 13). On a related note, 26 percent report using standard infrastructure without special optimizations for AI. This is either a gap in AI infrastructure optimization strategies or a problem waiting to be solved through future investments.

Delivery Optimization for AI Applications

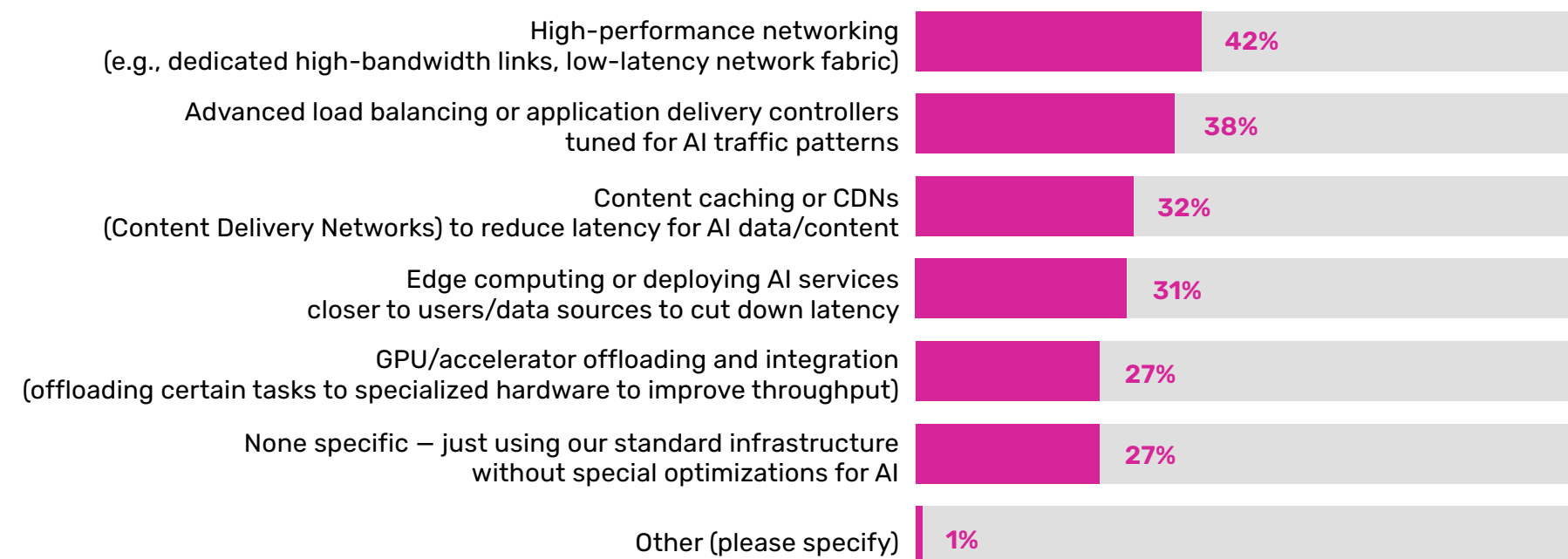


Figure 10 – Responses to the question, “What technologies or strategies are you using to optimize the delivery and performance of AI applications?”

56%

of organizations
plan to address security and privacy
concerns related to AI by training
staff and users on safe AI usage
and potential risks.



When asked about how they plan to address security and privacy concerns related to AI, 56 percent said “training staff and users on safe AI usage and potential risks.” Normally, training is considered an ineffective solution for security and privacy risks, given that people are prone to make security mistakes. However, modern AI creates new types of security risk, such as data leaking into the AI training model, so it’s reasonable that organizations feel the need to teach their employees how to handle AI from a security perspective.

Fifty-five percent said that they were implementing stricter data governance and access control for AI datasets and models. Like employee training, governance is a non-technical solution for AI security. The novelty of AI and its potential to defy traditional security controls likely accounts for this level of support for governance. More technical solutions were cited less often, with 45 percent saying they would use encryption and Zero Trust strategies and 40 percent would use new advanced firewalls and AI monitoring tools. This suggests that respondents have less faith in technology to solve the AI security and privacy problem than they do in human-centric knowledge sharing and rules.

AI-related Security and Privacy Concerns

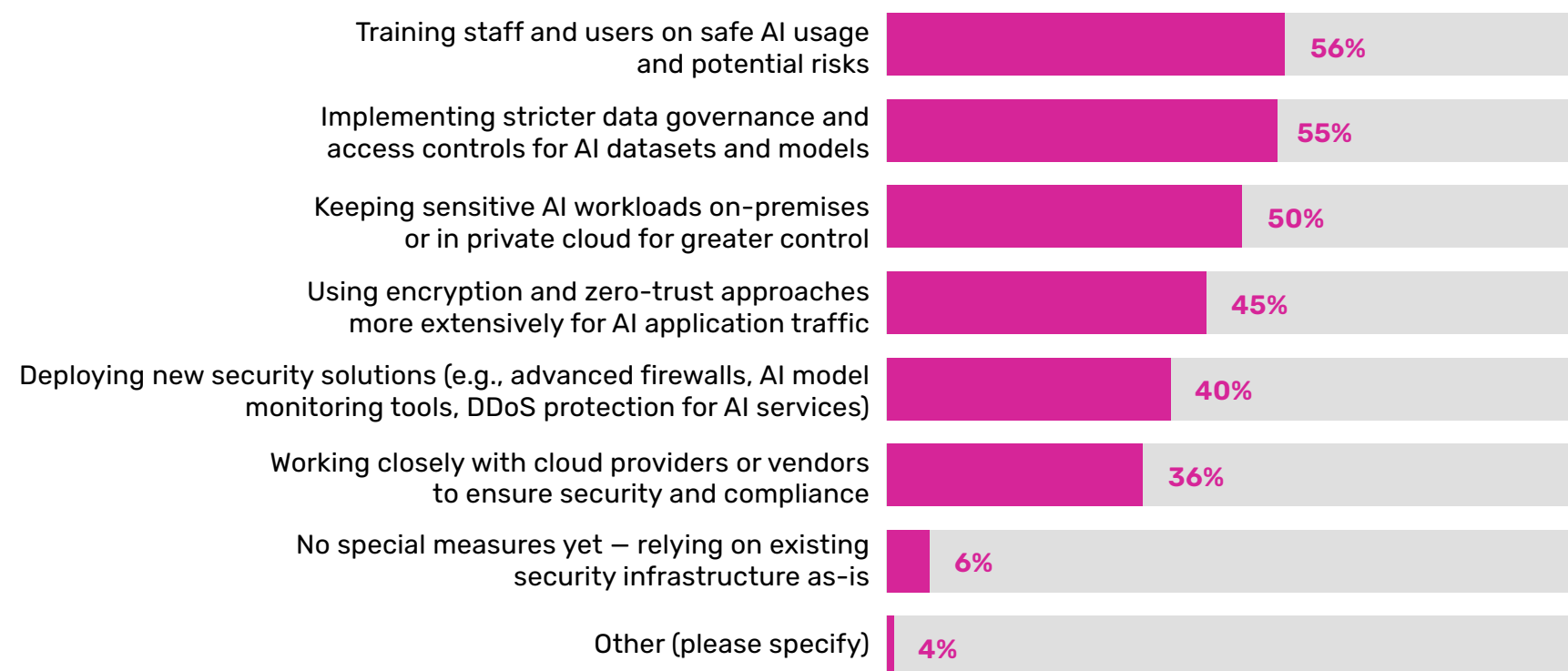


Figure 11 – Responses to the question, “How is your organization addressing these AI-related security and privacy concerns?”

Extensive Infrastructure Modernization Steps Planned for AI Workloads

Nearly 80 percent of organizations responding to the survey plan to modernize or upgrade parts of their infrastructure within the next three years. Twenty-one percent are actively modernizing now or plan to in the next six months, with projects in progress; 37 percent plan to within the next 12 to 18 months; and 21 percent will within two to three years. Only 17 percent are taking a “wait and see” approach. These findings show that most organizations take AI seriously, see it as an important workload that requires satisfactory infrastructure in a near-term, and plan to take action to address the necessary infrastructure changes.

Plans to Modernize AI Infrastructure

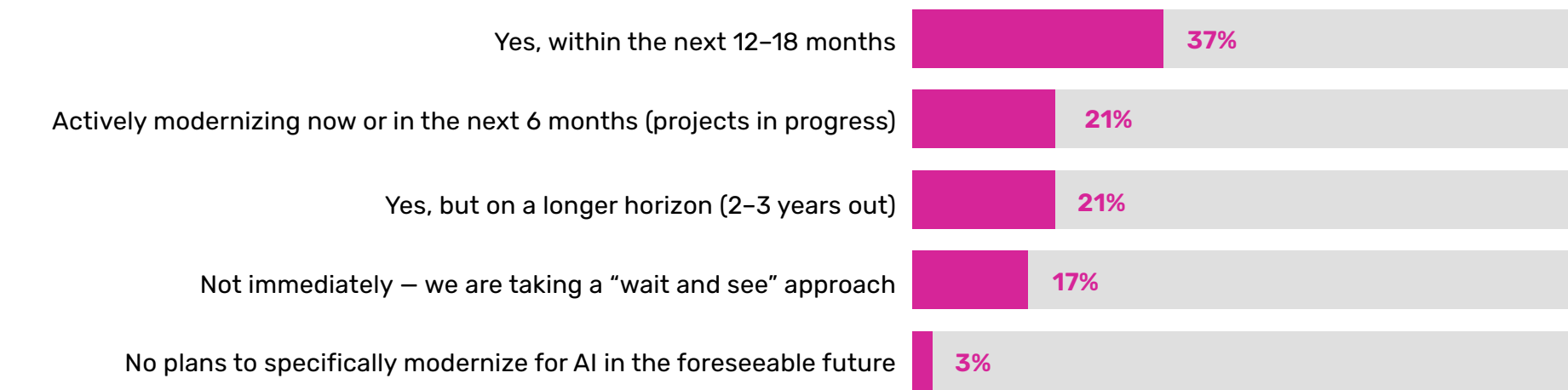


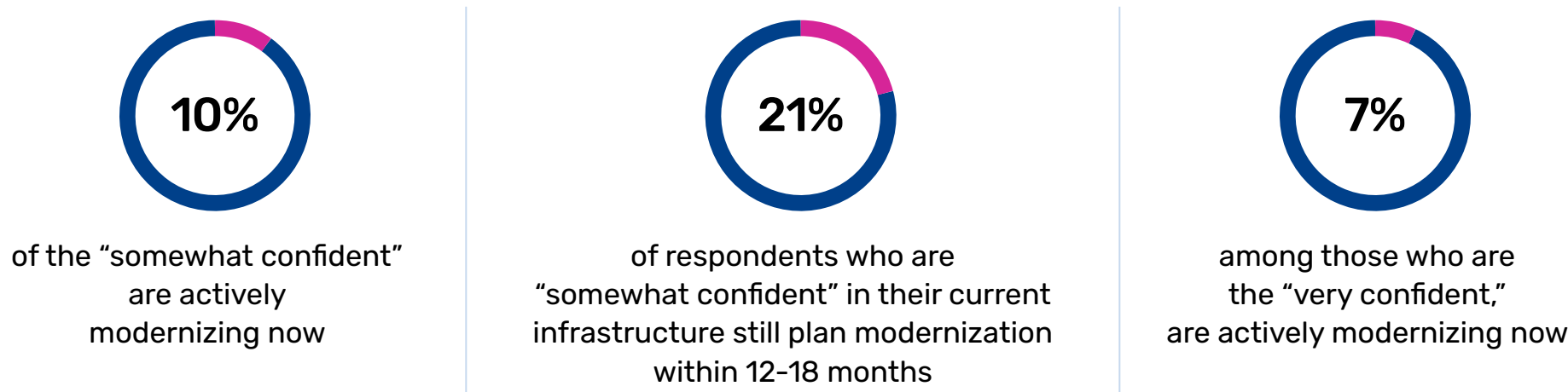
Figure 12 – Responses to the question, “Are you planning to modernize or upgrade parts of your application infrastructure to better support AI workloads?”

A significant majority of respondents plan to modernize or upgrade parts of their application infrastructure to provide better support for AI workloads.



Extensive Infrastructure Modernization Steps Planned for AI Workloads

Cross-referencing infrastructure confidence (Figure 7) with upgrade planning reveals that even organizations who express confidence in their current infrastructure are planning to modernize.

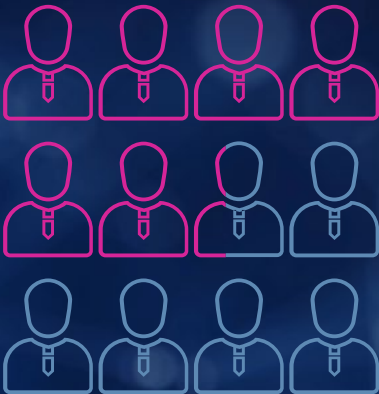


While 53 percent of respondents are “somewhat confident” their current infrastructure can handle AI needs, 79 percent are already planning infrastructure modernization within the next 18 months. The difference indicates a recognition of looming capacity limitations. Indeed, it seems that early adopters of AI and leaders in AI strategy recognize that they will need more modern infrastructure to keep up with expected growth in AI workloads and future demands on infrastructure.

At the other end of the scale, 11 percent of respondents who lack confidence plan on modernization within 12-18 months. They are later adopters who may not see the urgency of infrastructure modernization for the sake of AI. Or their internal processes are comparatively slow.

53%

of survey respondents are “somewhat confident” their current infrastructure can handle AI needs.



Extensive Infrastructure Modernization Steps Planned for AI Workloads

Highest Priority Areas for Modernization

Modernization plans are multi-faceted. The highest priority appears to be for security infrastructure modernization, such as AI firewalls and application programming interface (API) security, which were cited by 60 percent of respondents. Compute infrastructure, such as GPUs, is at 50 percent.

The Highest Upgrade Priorities Outpace Current Concerns

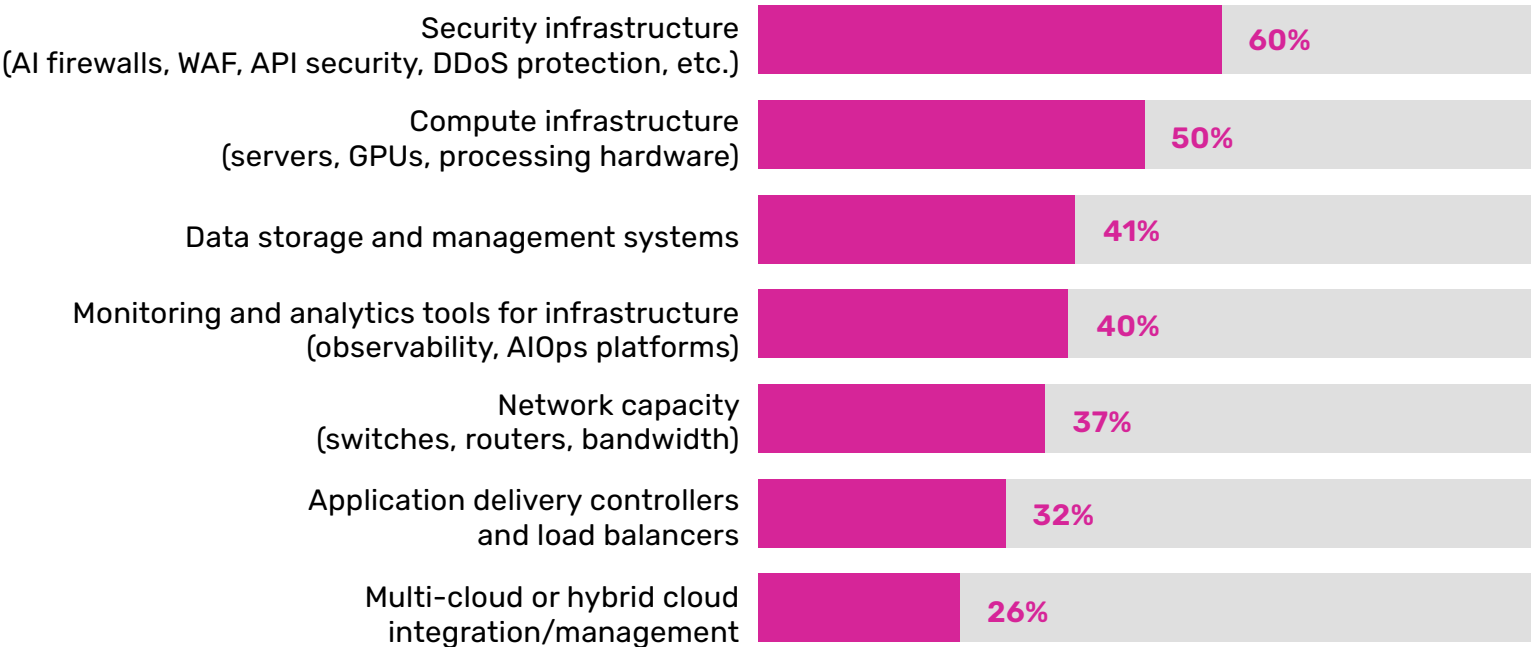


Figure 13 – Responses to the question, “Which infrastructure areas are the highest priority for upgrades or modernization to handle AI demands?”

60%

of survey respondents say highest priority appears is security infrastructure modernization.



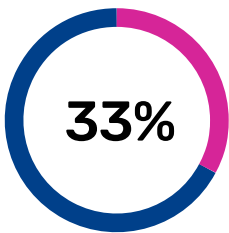
Key Findings

Extensive Infrastructure Modernization Steps Planned for AI Workloads

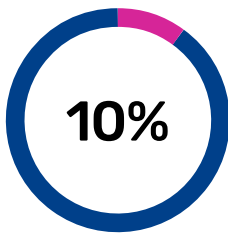
The highest upgrade priorities outpace current concerns.

Functional area	Perceived as a current limitation or pain point (Figure 8)	Highest priority for upgrades or modernization
Security	49%	60%
Computing Power	30%	50%
Network Bandwidth	29%	37%
Data Storage	19%	41%

These differing levels of emphasis may reflect a belief that issues like a lack of computing power or data storage will be greater in the future than current concerns would suggest. A comparison of future priorities with current bottlenecks (Figure 9) supports this hypothesis:



see compute limitations as a bottleneck for AI versus **50 percent who prioritize it for modernization**

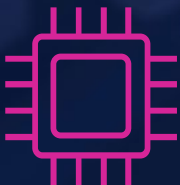


see security processing as a bottleneck versus **60 percent who prioritize it for modernization**

Highest Upgrade Priorities



Security



Computing Power



Network Bandwidth



Data Storage

Extensive Infrastructure Modernization Steps Planned for AI Workloads

Most Important Solutions to Implement

Security is the most important factor when it comes to selecting new solutions to implement for AI infrastructure. Forty percent of respondents said that API security solutions were the most important, with another 28 percent citing AI firewalls.

Ability of Current App Delivery Infrastructure to Handle AI Workloads

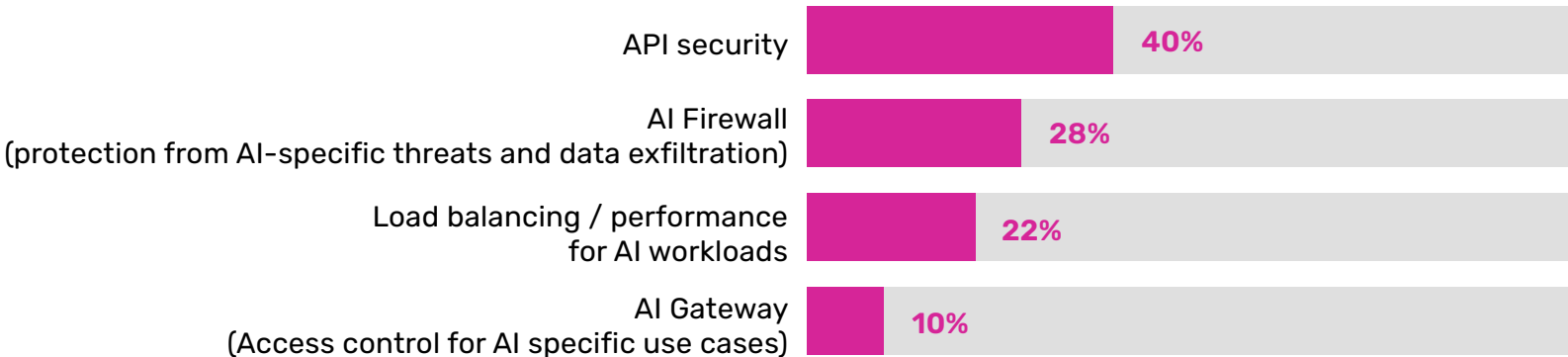


Figure 14 – Responses to the question, “What are the most important solutions you are considering specific to AI?”

The emphasis on API security is logical given the role that APIs play in AI model training and AI-based application functionality. AI models need to ingest data from a wide variety of sources, typically accessing that data through APIs. Without strong API security, negative outcomes from this process include access to data that has been designated “off limits” to the AI model as well as excessive API requests creating performance problems for data repositories. AI firewalls offer a comparable countermeasure.

40%

of survey respondents said that API security solutions were the most important.



Key Findings

Extensive Infrastructure Modernization Steps Planned for AI Workloads

A look at open-ended questions about the most important potential AI infrastructure solutions also surfaced concerns about data security.



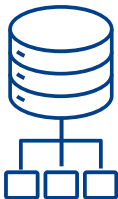
Eighty-six respondents mentioned a need for AI-specific security.

- “AI Firewall” (protection from AI-specific threats and data exfiltration)
- “API security”
- “Access control gateway”



Other respondents mentioned data protection in their comments.

- “Data loss prevention for AI systems”
- “Encryption for AI data in transit and at rest”
- “Model security and governance”



Load balancing and performance for AI workloads, cited by 22 percent of respondents as the most important solution under consideration, also came up in open-ended responses.

Twenty-one percent of respondents mentioned the following issues in their open-ended responses:

- “Load balancing / performance for AI workloads”
- “Application acceleration technologies”
- “Traffic management for AI applications”

62%

of respondents preferred vendors with platform strategies over standalone products.



This desire is likely based on the challenges of implementing too many point solutions in an already complex environment. A platform approach is preferable because it reduces the number of vendors and specialized skills required for AI infrastructure modernization.

30%

of respondents said budget constraints and costs are the leading obstacle holding organizations back from modernizing their infrastructures to support AI workloads.



Variety of Factors May Constrain Modernization of AI Infrastructure

What is holding organizations back from modernizing their infrastructures to support AI workloads? Budget constraints and costs are the leading obstacle cited by 30 percent of respondents. This problem does not appear to be related to a lack of support from senior management, which only 3 percent say is an obstacle. The issue could simply be a lack of resources, though the perceived difficulty in getting budget may be related to the divergence between enthusiasm for AI as a strategy versus the existence of a proven business case for the technology.

Constraints on Modernization

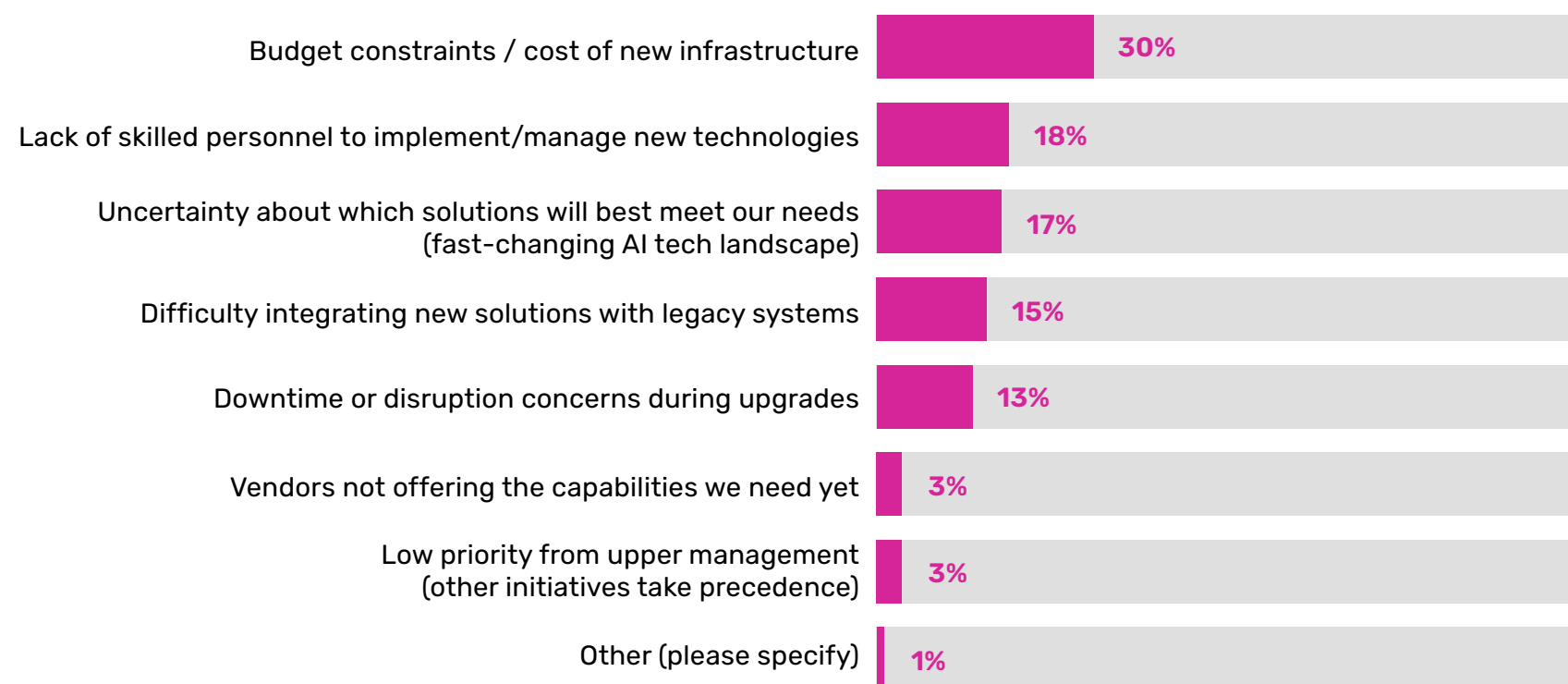


Figure 15 – Responses to the question, “What is the biggest obstacle or challenge your organization faces in modernizing your infrastructure for AI?”

Obstacles To Modernization

18%

lack of skilled personnel.

17%

uncertainty about which solutions will be best to meet the needs of fast-changing AI workloads.

15%

expected difficulty in integrating new solutions with legacy systems.

Technical and personnel constraints also emerged as obstacles to modernization, with 18 percent saying that a lack of skilled personnel is a challenge and 17 percent citing uncertainty about which solutions will be best to meet the needs of fast-changing AI workloads. Another 15 percent said they expected difficulty in integrating new solutions with legacy systems.

Viewpoints on these issues varied by organizational level.

C-Level (CIO/CTO)

Budget (50%), skills gap (17%). Senior managers must make the business case for modernization, and this level of concern over budget suggests that they are having trouble making such a case.

Director

Budget (27%), skills gap (26%). Directors are on the line for implementation, so it makes sense that they worry about the skills gap. Without the right skills, modernization will stall, and directors will feel the pressure to solve the problem.

Manager

Budget (28%), solution uncertainty (18%). Like directors, managers are expected to drive modernization in the real world. If it isn't clear that available solutions can do the job, that is cause for concern.

CISO

Legacy integration (50%), downtime concerns (50%). Security executives seemed worried about disruptions to information availability, a pillar of information security, due to AI modernization.



Taken together, these concerns suggest a relatively high level of uncertainty about the practical challenges involved in modernization. In theory, modernization is a high priority. In practice, issues like skills, integration, and solution features emerge as obstacles, apparently influencing budget considerations.

Conclusion

Generative AI is rapidly moving from niche use to the core of enterprise operations, reshaping how organizations deliver services and build applications across hybrid cloud environments. This shift brings significant opportunities but also introduces new challenges, as AI workloads demand greater flexibility, performance, security, and governance. Many existing infrastructures are not yet equipped to meet these demands. To keep pace, organizations must modernize their environments to support real-time, data-intensive, and scalable AI workloads.

The State of AI Infrastructure Survey 2025 from A10 Networks highlights an urgent and widespread trend. Seventy-six percent of organizations have already adopted GenAI technologies, and 71 percent are actively using or piloting AI solutions. The shift from experimental use to enterprise-grade workloads is well underway.

Despite this momentum, many organizations are unprepared to support the unique demands of AI. Infrastructure modernization is lagging, with gaps in security, visibility, and automation. Only a small fraction of respondents reported having fully automated scaling capabilities, and most expressed only moderate confidence in their infrastructure's readiness to support future AI growth.

Report Highlights

67%

of organizations have already adopted GenAI technologies.

71%

of organizations are actively using or piloting AI solutions.





Recommendations

To meet this moment, organizations should take the following steps.

- **Modernize infrastructure** across all layers, including compute, storage, networking, observability, and data governance
- **Embrace hybrid cloud strategies** that offer both scalability and control, especially for workloads with performance sensitivities or regulatory needs
- **Invest in performance optimization and automated scaling technologies** to ensure systems can adapt in real time to changing AI workload demands
- **Create long-term infrastructure roadmaps** that align with projected AI adoption and growth within the organization
- **Customize infrastructure strategies** to reflect industry-specific needs, acknowledging that sectors such as finance, technology, healthcare and education face distinct regulatory and operational challenges

AI adoption is accelerating with pilot projects rapidly moving into production. As workloads expand, legacy infrastructure will struggle to keep up, making modernization essential. IT and security professionals understand what they must do to modernize their infrastructures to support AI workloads. Budget limitations and skills gaps are obstacles, but these can be overcome. A strong business case for AI with clear alignment to innovation, efficiency, and competitive advantage can help unlock the resources and talent needed to move forward.

By acting decisively now, organizations can build resilient, scalable, and secure environments that harness the full potential of AI. This transformation will not only support current workloads but also lay the foundation for future innovation and sustained growth.

About A10 Networks

A10 Networks provides security and infrastructure solutions for on-premises, hybrid cloud, and edge-cloud environments. Our 7000+ customers span global large enterprises and communications, cloud and web service providers who must ensure business-critical applications and networks are secure, available, and efficient. Founded in 2004, A10 Networks is based in San Jose, Calif. and serves customers globally.

For more information, visit [A10Networks.com](https://www.a10networks.com) and follow us [@A10Networks](https://twitter.com/A10Networks).

Learn More

[About A10 Networks](#)

Contact Us

[A10Networks.com/contact](https://www.a10networks.com/contact)

©2025 A10 Networks, Inc. All rights reserved. A10 Networks, the A10 Networks logo, ACOS, Thunder, Harmony and SSL Insight are trademarks or registered trademarks of A10 Networks, Inc. in the United States and other countries. All other trademarks are property of their respective owners. A10 Networks assumes no responsibility for any inaccuracies in this document. A10 Networks reserves the right to change, modify, transfer, or otherwise revise this publication without notice. For the full list of trademarks, visit: [A10Networks.com/A10-trademarks](https://www.a10networks.com/A10-trademarks).

Part Number: A10-EB-14176-EN-01 SEP 2025