Deloitte. Insights

Four futures for technology infrastructure: Which one are you building toward?

Infrastructure choices made today could shape tomorrow's enterprise flexibility. Four possible futures show why leaders should design for resilience, trust, and optionality now.

Article • 8-min read • 15 July 2026 • Deloitte Center for Integrated Research

By 2032, the technology stack will likely look far different from the one many organizations are modernizing today. AI agents could be managing workloads across cloud, edge, and on-premises environments. Interfaces could become more intent-driven. Trust, identity, orchestration, and resilience might be needed not just in applications and infrastructure but also across increasingly autonomous systems.

For C-suite leaders, the challenge is not predicting one future with precision but making infrastructure decisions today that maximize flexibility for several plausible futures while managing existing constraints.

To explore these possible futures, Deloitte conducted interviews with more than 30 technology leaders to better understand their perspectives on architecture and stack decisions, security and trust innovations, and potential future states for tech infrastructure. We combined those insights with data from Deloitte's published research and an analysis of market signals. Many leaders envisioned a 2032 landscape with ambient experiences, multi-environment architectures, continuous identity verification, and adaptive systems, all while contending with constraints including limited computing resources, concentrated chip production, energy challenges, rising costs, network issues, technical debt, and skill gaps.

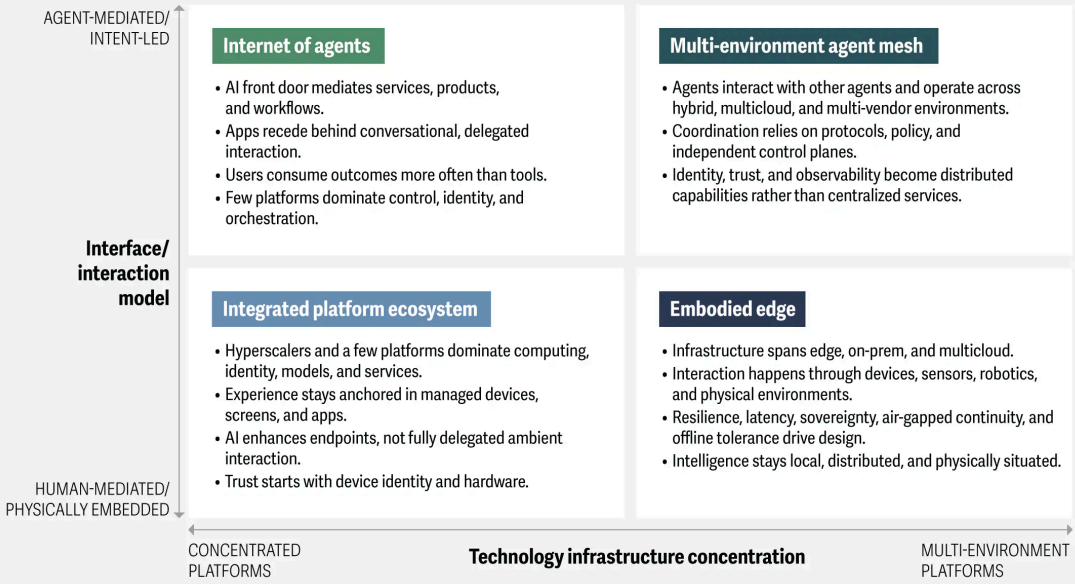
But the shape of that future remains uncertain, given two especially important technology infrastructure questions: Where will infrastructure control concentrate? And how will people, devices, applications, and agents interact?

- Technology infrastructure concentration: Infrastructure includes computing, data, platforms, orchestration, governance, and trust controls. Some forces point toward greater concentration, with a few hyperscalers and major platforms expected to centralize control over resources for efficiency and standardization. Others point toward multi-environment architectures, as enterprises distribute workloads across cloud, edge, on-premises, and sovereign environments to meet needs for latency, resilience, continuity, and regulatory control.
- Interface and interaction model: Interactions could remain largely human-mediated and device-led, focused on physically embedded experiences with screens and graphic user interfaces. Or they could shift toward intent-led, agent-mediated interactions. Physical AI and wearables like smart glasses highlight the shift toward more contextual, less app-centric experiences, blurring the line between device and interaction.

Combining these dimensions may produce four possible futures for technology infrastructure: an integrated platform ecosystem, an internet of agents, embodied edge, or a multi-environment agent mesh. The future may be one of these, more than one, or a blended combination. Each presents distinct enterprise risks and opportunities in architecture, orchestration, and security (figure 1).

Figure 1

4 scenarios for the future technology stack



Source: Deloitte Center for Integrated Research analysis based on interviews conducted with more than 30 technology leaders between April and June 2026.

Deloitte. Insights | deloitteinsights.com

Scenario 1: Integrated platform ecosystem (concentrated platforms, human-mediated or physically embedded interfaces)

In this future, infrastructure remains concentrated, but interaction could be human-mediated and device-led. A few hyperscalers and platform providers continue to dominate the underlying computing, model, identity, and service layers, while the user experience remains anchored in devices, screens, and applications. For devices and robotics, a small number of vertically integrated manufacturers may also gain outsized influence over physical AI stacks.[1]

This is the most familiar of the four futures because it extends many of the patterns enterprises already know. It offers scale, performance, standardized tooling, and relatively clear operating boundaries. For executives, the appeal can be straightforward: Platform concentration can simplify vendor management, accelerate deployment, and make it easier to scale AI without rewriting the entire enterprise interaction model.

As the vice president of technology at a tier-one bank says, “In the next five to seven years … the comfort level of going full cloud is going to happen … the security and the reliability is getting better and better to the point where it doesn’t make sense for us to pay for the redundancy of on-prem and to try to maintain it on-prem. It’s a slow migration, it’s not an overnight trust, but … eventually you go full cloud.”

Opportunity: The opportunity in this future is not novelty but leverage. Enterprises can use AI to make devices, apps, and workflows smarter without forcing wholesale behavior change across the workforce or customer base. The app, the dashboard, and endpoint still exist, but an intelligence experience layer above them might help users plan, navigate, and act across a more integrated platform ecosystem.

Risk: The risk in this scenario is dependency. A concentrated platform landscape can create lock-in not just at the infrastructure layer, but also across identity, service orchestration, data access, and governance. Emerging vendor friction, including tollgating around systems-of-record data,[2] may intensify this future. Resilience may look strong at the platform level, while strategic flexibility weakens at the enterprise level.

Scenario 2: Internet of agents (concentrated platforms, agent-mediated interfaces)

Here, infrastructure remains concentrated, but the interface changes radically. In this scenario, applications don't disappear, but they recede behind an AI "front door" that mediates services, products, and workflows. Users increasingly consume outcomes instead of tools. App-to-app or agent-to-agent interaction becomes common, while humans are mainly consulted for key decisions.

Beneath that experience, however, the underlying system still depends on a relatively small set of cloud and AI platform providers. The former head of systems integrator alliances and delivery strategy at one hyperscaler explains this scenario: "Everything, all services, products, information is accessed through an AI front door, but beneath that, everything is dependent on few large cloud players and providers of apps, identities, and networks."

This future is especially consequential because it shifts where enterprise value is created and captured. If interaction moves from app navigation to delegated intent, then the strategic battleground moves away from the application surface and toward orchestration, policy, context, and trust. Models themselves might become less differentiated, while value migrates toward the systems that orchestrate quality, route tasks, manage context, and integrate workloads.

Opportunity: For enterprises, the upside can be substantial. Work can be reorganized around outcomes rather than screens. Customers may no longer need to navigate fragmented digital journeys across multiple interfaces; agents could instead manage product discovery and transactions. Employees might spend less time coordinating manually across systems that should already work together. In principle, this could reduce friction dramatically and change the economics of service delivery.

Risk: This scenario introduces a more opaque form of dependency. If the control layer is concentrated, enterprises may lose visibility into how actions are brokered, how decisions are routed, which systems were touched, and where accountability sits. Trust becomes harder to inspect precisely when it becomes more important. As interfaces become more invisible and outcome-led, manipulated outputs can distort judgment at scale. Identity then becomes an agent problem, not just a user problem, requiring the least privilege, clear audit trails, and stronger controls over what agents can see and do.

There is also a hard economic edge to this scenario. Token metering, expensive frontier models, cloud costs, and long-running agents could be major constraints. Elegant agentic architectures may create significant value, but they may also become costly if enterprises do not engineer for efficiency and control from the start. As AI agents begin to operate across enterprise systems, software vendors may also move to meter, restrict, and monetize agent access to data and actions. That shifts value capture away from seats and interfaces and toward execution itself.[3]

Scenario 3: Embodied edge (multi-environment platforms, human-mediated or physically embedded

interfaces)

Infrastructure becomes even more distributed in this scenario, while interaction remains grounded in the physical world. Manufacturing, logistics, and regulated environments are already leaning into local inference and sovereign deployment. By 2032, that shift could accelerate as physical AI expands across wearables, biotech, vehicles, public infrastructure, robots, drones, space tech, and other connected devices.

The architectural bet is that intelligence needs to be close to where work happens: near machines, facilities, people, and regulated processes. AI earns its place by being embedded in physical operations, not simply layered on top of them. “In our manufacturing facilities, in our sites, in the hospitals, in the distribution centers, we have more and more of those AI applications, latency-sensitive use cases, the robotics, some autonomous AI systems, some remote monitoring systems,” says the head of digital and data science at a Fortune 500 life sciences company. “The cybersecurity is going to be embedded everywhere.”

In this scenario, the interface becomes more ubiquitous. Cameras, sensors, spatial systems, industrial controls, and physical environments all become ways for people to interact with technology. These interfaces are shaped by a need for real-time inference, local continuity, and jurisdictional control.

Opportunity: The opportunity here is resilience by design. Local processing, sovereign deployment, graceful degradation, offline tolerance, and air-gapped operations can reduce dependence on a concentrated set of vendors while preserving continuity. Hyperlocal processing can also give enterprises tighter control over inference and latency performance that centralized architectures may struggle to support.

Risk: The trade-off is cost and complexity. Edge infrastructure can require considerable investment, and distributed environments are harder to observe, patch, govern, and support. Orchestration will be important, especially across interfaces, middleware, data, and infrastructure. For many organizations, this scenario will likely present a near-term execution challenge rather than a structural ceiling as they build the operating discipline needed to manage it at scale.

Scenario 4: Multi-environment agent mesh (multi-environment platforms, agent-mediated interfaces)

This scenario is the most distributed, and arguably the most transformative, of the four futures. In this future, interfaces become invisible and agent-mediated, while infrastructure is spread across multiple vendors, clouds, edge environments, and geographies. Agents become the primary interaction and execution layer, but they operate across a hybrid mesh that no single provider fully controls.

In this scenario, the trust architecture is the architecture. Coordination depends on shared protocols, policy-based routing, and independent control points so that identity, orchestration, and resilience do not rely on one dominant platform. One director of IT and chief of technology operations at an oil and gas company says, “I’m seeing a great trend towards a distributed AI native and what I would say is an autonomous infrastructure stack … You have your edge systems, hyperscale cloud and AI agents, and secure digital networks operating as one coordinated system.”

Opportunity: The opportunity is strategic optionality. Enterprises could place workloads where they create the most value, preserve sovereignty where it matters, and reduce dependency on any single platform core. They could also participate in broader ecosystems of enterprise agents, partner agents, and service-provider agents without ceding full control to one dominant intermediary.

Risk: The risk is that this future might be the hardest to make operational. There is still no dominant standard for agent registries, inter-agent protocols, or agent provenance. Without those foundations,

interoperability risks may remain more aspirational than operational.

Trust becomes the defining challenge. Enterprises may need continuous validation, behavioral controls, stronger cryptography, and better ways to manage permissions, traceability, and observability in multi-agent environments. This is not a peripheral security issue. It's what would make this model work.

Enterprises might also need an agent observability layer that can monitor performance, enforce guardrails, and surface risk in real time. Immediate oversight could catch policy violations, data leakage, or harmful outputs, while longer-term analysis could detect reasoning drift, repeated tool failures, or coordination breakdowns across agents. Without that visibility, a multi-environment agent mesh could become too opaque to govern at enterprise scale.[4]

Design for multiple possible futures

These four futures are not predictions. They are ways to pressure-test the decisions technology leaders are making now about architecture, orchestration, and trust. The common thread across all four is the need for flexibility. Instead of building around one assumed future, leaders should design for optionality: modular designs, decoupled controls, reusable patterns, and trust models that can scale across concentrated and distributed environments. As AI agents take on more work, enterprises will also need clearer ways to govern what those agents can access, what actions they can take, and how their decisions are monitored.

But none of this works without a modern data stack supporting orchestration, observability, and trust across distributed environments. Data should be accessible, governed, observable, and secure enough to support orchestration in increasingly complex environments. Rules-based controls are a start, but they are likely not enough on their own. Enterprises may also need more adaptive approaches to identity, behavior monitoring, and cryptography, as well as agent oversight.

The future of technology infrastructure will not depend on computing power alone, but on how well enterprises connect experience, data management, orchestration, workload placement, and trust before market pressure forces the issue. In the next article in this series, we'll examine the 18- to 24-month playbook leaders can use to begin building for that future now.

Continue the conversation

Meet the industry leaders



Chris Thomas
Principal | Hybrid cloud infrastructure offering leader | Deloitte Consulting LLP



Parth Patwari
Principal, US AI & Engineering Leader | Deloitte Consulting LLP



Gopal Srinivasan
Principal | Global AI and data lead | Deloitte Consulting LLP



Oniel Cross
Principal | Government and public sector hybrid cloud infrastructure offering leader | Deloitte Consulting LLP



Diana Kearns-Manolatos
Senior manager, subject matter specialist | Deloitte Services LP

BY

Chris Thomas
United States

Parth Patwari
United States

Gopal Srinivasan
United States

Oniel Cross
United States

Diana Kearns-Manolatos
United States

Iram Parveen
India

ENDNOTES

1. George Chowdhury, "Humanoid robots market: Separating hype from reality," ABI Research, 2026.

2. Daniel Levy, "The emergence of software agent toll gates," Tech Startups, May 25, 2026.

3. Ibid.

4. Prakul Sharma, Parth Patwari, and Brijraj Limbad, "AI agent observability: Measuring what matters in multiagent systems," *The Wall Street Journal*, Feb. 19, 2026.

ACKNOWLEDGMENTS

We would like to thank Deloitte's subject matter experts—**Adnan Amjad**, **Arpan Tiwari**, **Colin Soutar**, **Duncan Stewart**, **Faruk Muratovic**, **Prakul Sharma**, **Scott Buccholz**, **Tim Davis**, and **Vinit Shah**—for taking the time to speak with us and for sharing their valuable insights, which helped shape this research.

We are also grateful to the marketing and public relations team for their guidance and leadership in extending the impact of these insights: **Andrew Ashenfelter**, **Anushka Bose**, **Christian Parsons**, **Cindy Chang**, **Jen Reid**, **Kaneez Fizza**, **Matt Merill**, **Nicole Bostock**, **Rachel Freya Rosenberg**, **Rebecca Lalez**, **Saurabh Rijhwani**, **Tafline Laylin**, and **Winslow Sowards**.

Finally, we appreciate the support of our research partners at 10EQS for collating input from market leaders and providing the analysis and synthesis of data that was instrumental to this work.

Editorial (including production and copyediting): **Corrie Commisso**, **Shyamili M**, **Pubali Dey**, and **Anu Augustine**

Design: **Molly Piersol** and **Sonya Vasilieff**

Cover image by: **Jim Slatton** and **Sonya Vasilieff**

Knowledge services: **Rishitha Bichapogu**

COPYRIGHT

Copyright © 2026 Deloitte Development LLC. All rights reserved. Member of Deloitte Touche Tohmatsu Limited