

ORACLE

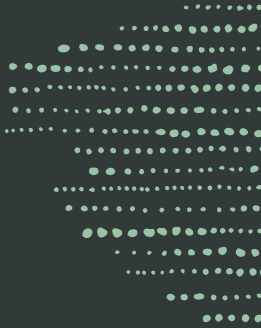
# 10 Cloud Trends CIOs Must Track in 2024

Promising ideas are becoming game-changing



# Table of contents

- Introduction..... 3
- Trend 1:** Cloud native AI services drive new applications ..... 4
- Trend 2:** Hybrid clouds go mainstream ..... 5
- Trend 3:** Multicloud ushers in a new era of cooperation ..... 6
- Trend 4:** CIOs look to rein in cloud costs ..... 7
- Trend 5:** More cloud-based applications are containerized ..... 8
- Trend 6:** No code and low code democratize development ..... 9
- Trend 7:** RAG and vector databases will improve generative AI results ..... 10
- Trend 8:** Cloud computing levels the security playing field ..... 11
- Trend 9:** The edge becomes the center of cloud innovation ..... 12
- Trend 10:** Organizations will evaluate data sovereignty and data privacy ... 13
- It's a cloud-first world ..... 14



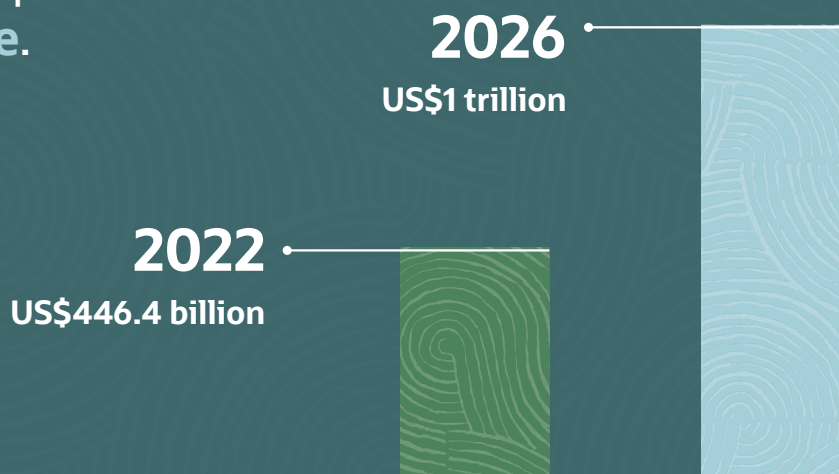
Last year saw some incredible technology advances. Generative artificial intelligence (AI) wowed us with its ability to quickly perform a wide variety of tasks, from recommending the best recipe to passing the bar exam. For most of 2023, CIOs imagined what this new type of AI could do for their businesses. Late in the year, we started to see providers build real, useful capabilities into their applications.

As generative AI piqued imaginations, something else remarkable happened. Most of us simply assumed that all this cool AI stuff would be delivered in the cloud. Someone, somewhere probably thought about building AI tools in an owned data center, but even if you did decide to run the technology on your own servers, most development would happen in the cloud because that's where the tools, infrastructure, and required services reside.

Using the cloud for development projects isn't new, and the consensus is that the cloud is where AI apps will reside. That's driving your IT architects to think more seriously about what else they can—and should—do in the cloud.

They aren't alone. Cloud spending is expected to more than double, from US\$446.4 billion in 2022 to more than US\$1 trillion by 2026, according to Forrester<sup>1</sup>. Your competitors are likely leveraging cloud's benefits. Here are ten trends that will make the cloud even more appealing for your business in 2024.

Cloud spending is expected to more **than double**.



<sup>1</sup> Forrester blog, *Announcing The Public Cloud Market Outlook, 2022 To 2026, November 2022*



## Cloud native AI services drive new applications

What innovations will we be unable to live without in 2025?

It's exceedingly rare that a new technology launches a revolution. Tech changes our lives and how we exist in the world, but it tends to do so by evolution. The smartphone seems revolutionary now, but in 2007 not much about Apple's original iPhone suggested that it would become as ubiquitous as it is today. You'll think of other examples. Our point is that it's impossible to pinpoint the ways generative AI will be used 15 or so years hence—though, of course, plenty of pundits are trying.

Right now, what we can see are many powerful AI services available in the cloud. This year, it will become easier to customize generative AI using your company's own data and then combine those customizations with other powerful technologies to create useful applications. And, of course, a range of vendors are doing the same. Your business will be the beneficiary.

Tired of poring over lengthy reports? Let generative AI summarize key findings. Want to improve the end user experience? Bring in an advanced AI chatbot. Need to find hints of changing customer sentiment in your transactions? Look for anomalies in real time. All of that's courtesy of AI services in the cloud.

**US\$151.1 billion**

Spending on generative AI solutions by 2027, a compound annual growth rate (CAGR) of 86.1% over the 2023 to 2027 forecast period.<sup>2</sup>

<sup>2</sup> IDC Press Release, IDC Forecasts Spending on GenAI Solutions Will Double in 2024 and Grow to \$151.1 Billion in 2027, December 2023



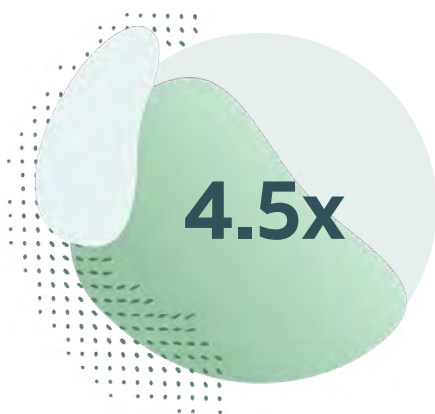
## Hybrid clouds go mainstream

The result? More value from your technology stack

Large organizations have huge investments in business applications, with many running in both the cloud and the data center. Making the most of the data these hybrid systems produce will get easier in 2024. Software vendors understand that even if the stated goal is to run your entire technology stack in the cloud, that will take years to achieve. With this in mind, you'll find new, more modular software architectures, faster APIs, higher-bandwidth networks, distributed cloud infrastructures, and even low-code/no-code tools that allow for rapid customization and the combining of apps that run on-premises with those that run in the cloud.

So, no matter where the hardware, software, or data resides, you can deploy complex applications that take the best advantage of all your resources. Visualize interchangeable building blocks, applications that can be built and rebuilt, and cloud infrastructure that can be deployed exactly where you need it. Benefits will include shorter development times, improved performance, and the ability to use the most cost-effective platform for your needs.

With such a modular, hybrid approach, you'll spend less time reinventing current wheels and more time creating new, better ones.



Expansion of the global hybrid cloud market, from US\$125.1 billion in 2023 to US\$558.6 billion by 2032, a growth rate of 17.5% annually.<sup>3</sup>

<sup>3</sup> [IMARC Group](#)



## Multicloud ushers in a new era of cooperation

Expanded interoperability will wring more mileage out of investments

The evolution of your technology stack likely came with unintended consequences and tradeoffs. For most enterprises, the decision about which cloud to use didn't involve whittling down the overall vendor list. For any number of reasons, you now have on-premises software from one set of suppliers and cloud-based systems from others. Getting everything to work together? Challenging at best.

In 2023, major cloud providers started actively supporting connections between their platforms, with tools, standards, and protocols that will let you run workloads where you want and then interconnect them to create new applications—with the full support of your suppliers. Some even went so far as deploying their cloud infrastructures in the same physical data centers, all to provide more choice for customers. Last September, for example, [Oracle's Larry Ellison traveled \(10:33\)](#) to Microsoft's headquarters and appeared with Satya Nadella to talk about their multicloud vision.

Expect more of this sort of cooperation in 2024, with CIOs the beneficiaries. Take a hard look at any vendor still looking to lock in your business by restricting interoperability.

# 98%

Enterprises using more than one cloud provider.<sup>4</sup>

<sup>4</sup> [Oracle, Press Release: 98% of Enterprise Using Public Cloud Have Adopted a Multicloud Infrastructure Provider Strategy](#)



## CIOs look to rein in cloud costs

Cutting egress charges, autoscaling, and other measures can help

Many companies that moved large workloads to the cloud and gave line-of-business leaders the ability to purchase seats and spin up instances are experiencing sticker shock. It doesn't help that it can be difficult to accurately estimate monthly costs and implement controls to avoid budget overruns.

A particular pain point is cloud data egress charges, where costs are incurred when data leaves a provider's network for another location, whether to an on-premises data center, a different provider, or your backup and disaster recovery sites. Another problem is that bills have gotten exceedingly complex. If you're challenged to decipher your cloud services invoice, you're not alone.

For 2024, look to innovative cloud providers that offer lower bandwidth pricing, few or no geography-based cost fluctuations, automated resource management and optimization, and flexible pricing to meet your specific requirements.

If you're sticking with a vendor that isn't helping you realize savings, cloud cost management tools and policy best practices can help. While most providers offer basic cost management utilities, specialized third-party systems that offer full visibility across multiple providers may be worth the investment. As to processes, keep an eye on autoscaling. Retailers may have seen traffic spikes over the holidays. If your provider automatically added instances, did anyone go in and shut down what you don't need? If your applications are containerized, are you running multiple containers in each VM? Now that your cloud use may be more predictable, have you considered less expensive reserved instances or prepaid fixed subscriptions versus pay as you go?

Educate anyone who contracts for cloud services on ways to minimize overprovisioning, and regularly review pricing offered by various providers. A workload or [cloud cost estimator](#) tool will help determine when it's worth shifting workloads.

**US\$1.34 trillion**

Worldwide public cloud services revenues by 2027, representing a five-year CAGR of 19.4%.<sup>5</sup>

<sup>5</sup> [IDC Press Release, Worldwide Public Cloud Services Revenues Grew 19.2% Year Over Year in the First Half of 2023, According to IDC Tracker, December 2023](#)



## More cloud-based applications are containerized

Benefits include cost savings and opportunities for innovation

Containers encapsulate all an app's dependencies and configurations into lightweight, self-contained packages that are portable and can be quickly scaled up and down on demand. Because containerization allows for easy deployment and right-sizing of applications, it helps contain costs. Need more capacity to handle end-of-month financials or a big holiday sale? Spin up more containers. Demand slowed a week later? Turn some of them off.

One valuable benefit of containerization is the ability to granularly, and often automatically, scale individual applications up and down based on CPU utilization, memory use, or other metrics. This flexibility ensures that your applications are always available and meet the needs of your users.

Containerization also promotes consistency and security when an application moves among cloud providers or runs partially on-premises by ensuring that it performs the same way regardless of the underlying infrastructure.

CIOs seeking the advantages of cloud native computing in 2024 should favor containers based on industry-standard, open-source technologies, such as Kubernetes and Kafka. Kubernetes is the platform for assembling and managing containers in cloud-based clusters, while Kafka gives those containers the ability to handle real-time data streams at scale. It's the win-win architecture of the future.

### 49%

Survey respondents saying Kubernetes has driven them to increase their cloud spending; 28% said Kubernetes represents half their cloud budgets.<sup>6</sup>

<sup>6</sup> *Cloud Native Computing Foundation*





## No code and low code democratize development

Small-scale and ad-hoc cloud software projects can pay off

Got a business project that doesn't require the heft of a cloud native or containerized application? A new generation of low-code and no-code tools provide organizations with the services and APIs to build new, end-to-end applications faster than they could with traditional programming languages, such as C++, or even with scripting languages, such as JavaScript, and then easily run them in the cloud.

Need a quick and simple app for a data analysis task? Low code and no code are the way to go.

Now, it's important to differentiate this newer breed of tool from traditional utilities aimed at line-of-business professionals. Executives might enjoy their spreadsheet macros, but whether it's "recording" a procedure or dragging and dropping visual icons, that popular rapid development methodology is worrisome because these applications often aren't tested, inventoried, documented, or secured. Today's cloud-centric low-code and no-code tools are different. They leverage cloud services using well-documented APIs and are secure, robust, and fault-tolerant while still giving your employees the freedom to try out new concepts quickly and inexpensively.

## RAG and vector search will improve generative AI results

Achieve more relevant and trusted responses to queries

If you don't know about [retrieval-augmented generation \(RAG\)](#), let us introduce you. Turns out, simply feeding generative AI your data won't necessarily help the system produce detailed and up-to-date answers to queries specific to your organization. Further, the process would be resource-intensive and could divulge information you don't want to share.

RAG and vectors can help. Here's how it works: RAG provides a mechanism to improve both the relevance and the trustworthiness of responses. It does that by pulling in your organization's most timely and applicable data while also enabling the model to cite source documents. To make it easy to determine what data is relevant, indexes—called vectors—are added to your data, making it highly searchable. Vectors are formed in such a way that it's possible to quickly determine what data RAG should supply to the AI for optimal results. And, you can keep certain sensitive information off limits. Now, the system uses only the data you select in producing its responses, making the output highly pertinent to queries within your expertise domain. Working as a team, RAG and vector search will simplify the way organizations deploy generative AI while dramatically improving performance, accuracy, and trust.



## Cloud computing levels the security playing field

Hyperscalers deliver top-tier detection services to all customers

Wider use of IoT and edge computing will demand speedier threat detection, and privacy regulations are likely to continue to change. Very large organizations have the resources to manage these issues in their own 24/7 security operations centers. But most businesses don't, unfortunately. That's a big problem because 2024 is likely to see threats evolve quickly; security talent remains rare and expensive; and bad actors will leverage new technologies, including AI, to find and exploit vulnerabilities almost instantaneously.

To be effective, security controls should be simple to use, simple to enforce, and simple to document for compliance. Security must also be highly automated, with around-the-cloud monitoring. Encryption must be the default, access controls must be stringent, and software updates must always be kept current.

A cloud provider's business viability depends on just that level of top-tier network, hardware, software, and data security. For hyperscale cloud providers, redundancy, data continuity, and security are designed into every layer, from hardware to control systems, from databases to communications. That effort, and the strict compliance and regulatory requirements that large providers implement to win government contracts, can pay off in a major way for ordinary businesses. Because every customer benefits from the same firewalls, key management, threat intelligence, network virtualization, patch management, and AI and security expertise, the cloud is the best way to stay level with competitors when it comes to securing sensitive information and delivering maximum uptime for employees and customers. These providers can also harness AI to help detect and prevent sophisticated cyberattacks.



## The edge becomes the center of cloud innovation

That demanding Internet of Things project? Now might just be the time to move forward

For 2024, the next generation of edge cloud architecture is moving more processing closer to end users and data sources. Not only is performance significantly improved through the judicious use of edge systems, but complex applications can be supported even where fast networking is unavailable.

Edge computing also enables real-time data processing and analysis, allowing for near-instantaneous decision-making and responses while still reporting activity and other data back to centralized systems. If your organization depends on applications that require real-time insights and actions, such as automated route optimization, emergency response systems, or mission-critical Internet of Things (IoT) devices, the edge is where you want to be.

As organizations embrace hybrid and multicloud, and the distinctions blur, edge computing can accelerate service delivery and enable new applications that require localized compute power.

**16.7 billion**

The number of connected Internet of Things devices as of year-end 2023.<sup>7</sup>

<sup>7</sup>IoT Analytics

## Organizations will evaluate data sovereignty and data privacy

Managing data access and storage to help an organization stay compliant

Increasingly, regulations around the world place strict limits on where and how information is stored, transmitted, and consumed. Business leaders need to understand how these laws affect their IT systems including cloud services and providing data access across borders. The financial and legal penalties for infractions can be substantial, and the trend is toward increasing regulation, oversight, and governance.

Across the board, technology decisions—including cloud use—will need to account for data sovereignty and data privacy just as they now consider cost, performance, and scalability.



# It's a cloud-first world

Whether you're looking to buy new applications or develop or improve your own, the cloud is increasingly the smartest place for those apps to be built and to live.

Application providers understand that most customers are comfortable using cloud-based apps with user interfaces running in browsers. Over time, that architecture delivers a simplified codebase and eliminates problems, such as supporting multiple generations of software.

Your developers get these same advantages, with less complexity. Specifically, they'll find rich sets of services delivered in the cloud to let them tap AI, databases, user identification systems, and more in a simple, standard way. With the browser as the universal user interface, a whole set of concerns about supporting a myriad of devices goes away.

As the cloud has evolved, so has the conversation about its use. A few years ago, you may have thought, "Maybe we should try this in the cloud." Then, "We should really do this in the cloud." Now, it's not a question. Cloud is the default, and not using it for new applications, storage, and analysis requires a very good argument.



Oracle understands the primacy of the cloud—and lives it. [Oracle Cloud Infrastructure](#) (OCI) was built from the ground up to support the complex needs of the world's largest organizations, and it works very well for that purpose. Oracle knows this because it's the only enterprise application provider to use its public cloud infrastructure to deliver its commercial SaaS applications: [Oracle Cloud Fusion Applications](#), NetSuite, and a lot more all run on OCI.

The services that Oracle uses to bring AI functionality to its applications are the same services you'll use to reap the benefits of [AI on OCI](#). Whatever your goals are for the cloud in 2024, Oracle can help.

[Learn more](#)

## Connect with us

Call +1.800.ORACLE1 or visit [oracle.com](#)

Outside North America, find your local office at [oracle.com/contact](#)

---

Copyright © 2024, Oracle and/or its affiliates. This document is provided for information purposes only, and the contents hereof are subject to change without notice. This document is not warranted to be error-free, nor subject to any other warranties or conditions, whether expressed orally or implied in law, including implied warranties and conditions of merchantability or fitness for a particular purpose. We specifically disclaim any liability with respect to this document, and no contractual obligations are formed either directly or indirectly by this document. This document may not be reproduced or transmitted in any form or by any means, electronic or mechanical, for any purpose, without our prior written permission. Oracle and Java are registered trademarks of Oracle and/or its affiliates. Other names may be trademarks of their respective owners.

