

What is the Cloud?

For decades, individuals, companies, and governments gained the benefits of increased computing power by accepting the expense and burden of maintaining and defending their own vulnerable IT systems. As IT spending has ballooned and threats—malign and incidental—to private data and systems abound, these users have increasingly turned to third-party companies to provide IT services. The most flexible and dominant arrangement is cloud computing, which enables the outsourcing of IT services that would otherwise occur on users’ own in-house computers (also known as “on-premises”).

Under this arrangement, users rely on cloud providers¹ —companies that operate vast networks of servers and associated data centers around the world—to offer optimized, secure, affordable, efficient, and powerful IT services over the internet. Amazon Web Services, Microsoft Azure, Google Cloud, Alibaba, and Tencent are leading examples. Cloud service customers can send and transfer data, access and store files, and use computer applications without having to shoulder the costs, responsibilities, and risks of managing their own IT systems. This outsourcing is why the cloud is often described as using “someone else’s computer.”

To date, the cloud has been efficient, comparatively affordable and accessible, and secure. It is a demonstrated accelerator of business growth and, consequently, promises to continue driving economic prosperity in the future.² Governments are turning to the cloud for activities from data storage to delivery of social-service benefits.³ Individuals have—even if they do not recognize it—been using the cloud for entertainment and education via social media, video games, and online streaming. Whether consciously or not, humanity is becoming ever more engaged with and dependent on the cloud for basic functioning.

But alongside these enormous benefits, the cloud raises vexing and interrelated governance challenges. If left unattended, these issues can undermine the economic and social gains of cloud adoption. This digital product by the Carnegie Endowment for International Peace organizes and analyzes these challenges, and explores how they may be managed or resolved.

Background on the Cloud

Fully appreciating the range of governance challenges requires some knowledge of the cloud’s users, benefits, underlying technology, commercial and technical deployment models, and evolution.

Who uses the Cloud?

Customers of the cloud fall into two categories: individuals and enterprises (businesses, government agencies, and other organizations). Individuals use the cloud via storage services like Apple's iCloud or Google Drive and via free⁴ email, social media, and digital games.

Enterprises rely on cloud computing solutions for back-office tasks such as managing large data sets and detecting fraud in real time, as well as for executing their missions, be that optimizing power generation or delivering customer-facing services like communications platforms and productivity tools. An illustrative market forecast projects that between 2021 and 2026, the global cloud market will grow from \$445 billion to \$947 billion.⁵

Benefits of Cloud Adoption

Enterprises and individuals use the cloud to reduce costs, operate at a larger scale, and build more flexibility into their IT usage. Transitioning to the cloud allows enterprises to offset the costs of procuring hardware for on-premises systems and the labor costs associated with maintaining and securing them. Instead, they contract with cloud providers on a more affordable pay-as-you-go operating model. The cloud providers harness economies of scale to profitably operate vast global networks of data centers, servers, and cables in order to provide computing solutions that can adapt and scale to users' needs.

Similarly, individuals can save money by using affordable "consumer cloud" services such as Google Drive, DropBox, and iCloud. By storing their files on Google Drive, for example, consumers can purchase phones and laptops with smaller local storage, saving money in the process. By lowering costs, increasing security and reliability, improving operational flexibility, and offering significantly greater computing power, the cloud allows billions of users to more efficiently operate in the digital world.

Types of Cloud Services and Deployment Models

At the most basic level, cloud service providers sell access to the services they develop, maintain, and offer in exchange for a fee. These services generally come in three forms—Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS), and Software-as-a-Service (SaaS). The differences lie in the functionality of these services, their costs, and the distribution of controls and responsibilities for the service between cloud providers and their customers.

- **IaaS** is the foundation of cloud-based services. Under an IaaS arrangement, a cloud service provider manages computing infrastructure; ensures that it is functioning, up-to-date, and secure; and sells access to customers. This allows customers to store data, host applications, and focus on their core operations without devoting time and resources to managing cumbersome IT systems. Paying as they go, customers can change the amount of storage or computing power they need frequently, avoiding the upfront costs of investing in their own on-premises infrastructure.

- **PaaS** provides customers a platform on which to build and run custom applications without the hassle of hosting them.⁶ Providers manage the tools needed to develop software, operating systems, databases, as well as all the infrastructure resources included in IaaS. For a fixed, pay-as-you-go fee, customers need only bring data and development skills to begin deploying new applications. Customers use PaaS to simplify the software development process. It allows them to keep the entire process in a single environment and avoid the challenge of managing licenses for the many tools they use.
- **SaaS**, sometimes described as on-demand software, allows customers to access applications over the internet. This model frees customers entirely from hardware and software management.⁷ In exchange for a subscription fee (though some, especially consumer, SaaS services might appear free because they are supported by advertising), customers can access applications that allow them to store data, work collaboratively on documents, email one another, or use such shared services as contact-management⁸ or enterprise financial services.⁹ For this arrangement to work, however, customers must give up ownership of the software and potentially some control over back-end functions like data processing and formatting. Their fee buys them confidence that the software provider will continue to provide maintenance, security, and innovative functionality to the application as technology evolves, customer needs change, and new security threats arise.

While many cloud providers specialize in one or two of these services, “hyperscale” cloud providers—like Microsoft, Google, Amazon, Oracle, Alibaba, and Tencent—offer all three.

These services are made available to customers through one of three deployment models. Each carries its own implications for security, responsibility, and liability in the event of service disruption, data breaches, privacy violations, and other cloud incidents.

- **Public cloud** is available to users over the public internet. The cloud’s digital and physical infrastructure is shared by all public cloud customers and managed by a cloud provider. Both individuals and companies can operate on the public cloud. For example, both can use Google Drive over the public internet. Cloud providers balance demand for cloud storage, computing power, and networking resources by setting up many data centers across geographic regions. A customer’s data might reside in any of these locations. As the cloud operates across borders, national laws can affect how data flows between these locations. For example, data localization laws can force a cloud provider to store a customer’s data in their home jurisdiction. Under this model, customers have little control over how their data is secured, where it is stored, and how it is moved between jurisdictions. As a result, in the event of a cloud incident, the cloud provider assumes almost total responsibility for its effects.
- **Private cloud** deployments are designed for a single customer, often government agencies or major companies. The cloud infrastructure is hosted in a data center owned by the customer or in a dedicated data center managed by the cloud provider.¹⁰ Operating on a private cloud allows customers to exercise more control over how their data is secured, stored, and moved. It also allows for greater customization in how cloud

infrastructure resources are set up and used. Under this model, responsibility for securing and governing the use of the cloud is shared between the customer and cloud providers based on terms they have agreed in their private contracts.

- **Hybrid cloud** deployments combine the public and private cloud models, allowing customers to choose the optimal combination of public-private cloud, as their circumstances and needs change.¹¹ This allows customers to store sensitive data and run critical workloads in private cloud environments, where they have more control over security while still making use of the efficiency and flexibility of the public cloud model for other functions.

The varied service and deployment models illustrate that the cloud is not a monolith nor is it entirely under the control of cloud providers.

The Evolution of Cloud Computing and the Cloud Market

The cloud market has grown rapidly in recent years. As a result, cloud infrastructure has been expanding and evolving around the world as more companies and governments integrate this technology into their daily operations and as individuals embrace consumer cloud services. Enterprise demand is also a key driver of this new construction, as enterprise cloud customers seek options to locally store sensitive data in order to comply with government policies on data localization.¹² These policy concerns have forced cloud providers to revise their calculations about how to most efficiently store and route data.

Another major trend in the cloud environment is the transition from its original, centralized approach to cloud computing (wherein data is stored, processed, and managed in large data centers around the world) to edge computing, which brings cloud functions, such as data storage and processing, closer geographically to end users. Edge computing aims to reduce the time between a user's keystroke (or the measurement taken by a sensor) and the retrieval or processing of data in the cloud. New applications, such as connected and autonomous vehicles¹³ and remote surgery,¹⁴ can benefit from these low response times. They will also benefit from deeper integration between cloud providers and telecommunications companies, which will make possible the local, real-time processing of data transmitted over 5G networks.

Over the next three years, buyers are expected to shift their cloud strategies to incorporate more edge computing. The hyperscale providers as well as smaller providers are doubling down on edge's flexibility to augment their existing services.

Notes

1 Cloud Providers: An organization whose primary function is to provide a component of [cloud computing service](#) — typically IaaS, PaaS, and/or SaaS — to other businesses, organizations, or individuals.

² See: Ryan Vlastelica, “Cloud Computing Seen as Tech Haven Amid Pandemic Uncertainty,” Bloomberg, April 7, 2020, <https://www.bloomberg.com/news/articles/2020-04-07/cloud-computing-seen-as-tech-haven-amid-pandemic-uncertainty>

³ Adam Stone, “2020 Puts Cloud Computing in Government to the Test,” Government Technology, September 1, 2020, <https://www.govtech.com/computing/2020-puts-cloud-computing-in-government-to-the-test.html>

⁴ See: Rani Molla, “Why your free software is never free,” Vox, January 29, 2020, <https://www.vox.com/recode/2020/1/29/21111848/free-software-privacy-alternative-data>

⁵ “Cloud Computing Market by Service Model (Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS)), Deployment Model (Public and Private), Organization Size, Vertical, and Region – Global Forecast to 2026,” MarketsandMarkets, 2021, <https://www.marketsandmarkets.com/Market-Reports/cloud-computing-market-234.html>.

⁶ Scott Carey, “What is PaaS? A simpler way to build software applications” InfoWorld, July 6, 2021, <https://www.infoworld.com/article/3223434/what-is-paas-a-simpler-way-to-build-software-applications.html>.

⁷ Keerthi Rangan, “What is SaaS? What You Need to Know for Future Success,” G2, July 15, 2021, <https://www.g2.com/articles/what-is-saas>.

⁸ The core of SaaS pioneer Salesforce’s business is a contact management service.


⁹ For example, widely-used software from Workday Inc. supports a variety of timekeeping and human-resources tasks.

¹⁰ Blair Felter, “The Different Types of Cloud Computing and How They Differ,” VXCHNGE, February 3, 2021, <https://www.vxchnge.com/blog/different-types-of-cloud-computing> .

¹¹ Muhammad Raza, “Public vs Private vs Hybrid: Cloud Differences Explained,” BMC, August 31, 2020, <https://www.bmc.com/blogs/public-private-hybrid-cloud/>.

¹² Aditya Karla, “Exclusive: India panel wants localization of cloud storage data in possible blow to big tech firms,” Reuters, August 4, 2018. <https://www.reuters.com/article/us-india-data-localisation-exclusive-idUKKBN1KP08J>.

¹³ George Anadiotis, “Why autonomous vehicles will rely on edge computing and not the cloud,” ZDNet, November 4, 2019 <https://www.zdnet.com/article/why-autonomous-vehicles-will-rely-on-edge-computing-and-not-the-cloud/>.



¹⁴ Dewanand A.Meshram and Dipti D.Patil, "5G Enabled Tactile Internet for Tele-Robotic Surgery," Procedia Computer Science, Volume 171 (2020), 2618-2625.
<https://www.sciencedirect.com/science/article/pii/S1877050920312771>.