✓   100 XP   ▶

What is cloud computing?

8 minutes

Cloud computing is renting resources, like storage space or CPU cycles, on another company's computers. You only pay for what you use. The company providing these services is referred to as a cloud provider. Some example providers are Microsoft, Amazon, and Google.

The cloud provider is responsible for the physical hardware required to execute your work, and for keeping it up-to-date. The computing services offered tend to vary by cloud provider. However, typically they include:

- **Compute power** - such as Linux servers or web applications
- **Storage** - such as files and databases
- **Networking** - such as secure connections between the cloud provider and your company
- **Analytics** - such as visualizing telemetry and performance data

Cloud computing services

The goal of cloud computing is to make running a business easier and more efficient, whether it's a small start-up or a large enterprise. Every business is unique and has different needs. To meet those needs, cloud computing providers offer a wide range of services.

You need to have a basic understanding of some of the services it provides. Let's briefly discuss the two most common services that all cloud providers offer – *compute power* and *storage*.

Compute power

When you send an email, book a reservation on the Internet, pay a bill online, or even take this Microsoft Learn module you're interacting with cloud-based servers that are processing each request and returning a response. As a consumer, we're all dependent on the computing services provided by the various cloud providers that make up the Internet.

When you build solutions using cloud computing, you can choose how you want work to be done based on your resources and needs. For example, if you want to have more control and responsibility over maintenance, you could create a *virtual machine* (VM). A VM is an emulation of a computer - just like your desktop or laptop you're using now. Each VM includes an operating system and hardware that appears to the user like a physical computer running Windows or Linux. You can then install whatever software you need to do the tasks you want to run in the cloud.

The difference is that you don't have to buy any of the hardware or install the OS. The cloud provider runs your virtual machine on a physical server in one of their datacenters - often sharing that server with other VMs (isolated and secure). With the cloud, you can have a VM ready to go in minutes at less cost than a physical computer.

VMs aren't the only computing choice - there are two other popular options: *containers* and *serverless computing*.

What are containers?

Containers provide a consistent, isolated execution environment for applications. They're similar to VMs except they don't require a guest operating system. Instead, the application and all its dependencies is packaged into a "container" and then a standard runtime environment is used to execute the app. This allows

the container to start up in just a few seconds, because there's no OS to boot and initialize. You only need the app to launch.

The open-source project, Docker, is one of the leading platforms for managing containers. Docker containers provide an efficient, lightweight approach to application deployment because they allow different components of the application to be deployed independently into different containers. Multiple containers can be run on a single machine, and containers can be moved between machines. The portability of the container makes it easy for applications to be deployed in multiple environments, either on-premises or in the cloud, often with no changes to the application.

What is serverless computing?

Serverless computing lets you run application code without creating, configuring, or maintaining a server. The core idea is that your application is broken into separate *functions* that run when triggered by some action. This is ideal for automated tasks - for example, you can build a serverless process that automatically sends an email confirmation after a customer makes an online purchase.

The serverless model differs from VMs and containers in that you only pay for the processing time used by each function as it executes. VMs and containers are charged while they're running - even if the applications on them are idle. This architecture doesn't work for every app - but when the app logic can be separated to independent units, you can test them separately, update them separately, and launch them in microseconds, making this approach the fastest option for deployment.

Here's a diagram comparing the three compute approaches we've covered.
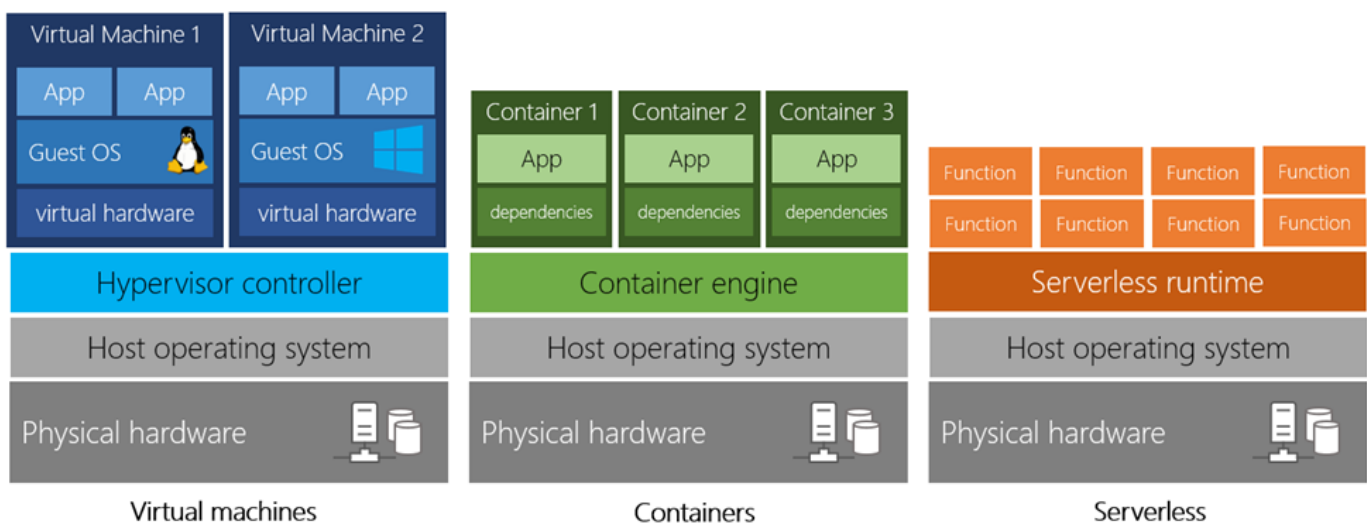
Storage

Most devices and applications read and/or write data. Here are some examples:

- Buying a movie ticket online
- Looking up the price of an online item
- Taking a picture
- Sending an email
- Leaving a voicemail

In all of these cases, data is either *read* (looking up a price) or *written* (taking a picture). The type of data and how it's stored can be different in each of these cases.

Cloud providers typically offer services that can handle all of these types of data. For example, if you wanted to store text or a movie clip, you could use a file on disk. If you had a set of relationships such as an address book, you could take a more structured approach like using a database.

The advantage to using cloud-based data storage is you can scale to meet your needs. If you find that you need more space to store your movie clips, you can pay a little more and add to your available space. In some cases, the storage can even expand and contract automatically - so you pay for exactly what you need at any given point in time.

Summary

Every business has different needs and requirements. Cloud computing is **flexible** and **cost-efficient**, which can be beneficial to every business, whether it's a small start-up or a large enterprise.

Next unit: Benefits of cloud computing

Continue >