Data in the Cloud

Defining the Cloud

What Is The Cloud?

The cloud has changed the way we think about, store, and analyze data. But what is it? Is it just a hard drive in the sky? Is it even *a thing, or many things*? While the term "cloud" is used to describe and market many different products, the concept is simple enough…The cloud is computing that does not take place on your devices.

One example of computing that doesn't take place on your device is using Google Sheets to create a spreadsheet. While you are using your device to access the file, it is stored and processed in the cloud. You only need to have a web browser installed. Google takes care of storing the document and creating the web-based editor that you use to edit values. You can share the document and allow others to edit without requiring them to have a program installed or download a file. Other examples of the cloud in everyday life include online gaming platforms and streaming services.

While end users benefit from these cloud offerings, there is a lot of technology that goes into making these products a reality. As users become more accustomed to the convenience and benefits of cloud offerings, more companies need to leverage these technologies to offer innovative products and remain competitive. This has led to a large-scale adoption and integration of cloud services.

Cloud Vendors

While it is often referred to as "the cloud" there are many vendors offering cloud products. The largest of these vendors maintain their own versions of "datacenters" which are large buildings that house massive amounts of physical hardware, like CPUs, storage, and networking devices. These buildings are located throughout the world and are networked together to offer greater availability and increased flexibility of where the computing takes place. Users then access the products and services hosted on these cloud platforms via the internet.

Some of the biggest operators in the cloud are Amazon Web Services (AWS), Google Cloud Platform (GCP), Microsoft Azure, Oracle Cloud Infrastructure (OCI), and IBM Cloud. These companies have large enough datacenters that they can rent hardware and services to other companies. Those services are then used by companies who want to use cloud technology to enhance their own products and services but either don't want or need to create their own datacenter. To better understand exactly what these vendors are offering, it can be helpful to understand some details about the internet.

Internet

By its nature, the cloud depends on the internet to provide end users with access to hardware in remote locations. So what is the Internet? A simple way to think of it is as a series of connected networks which make it possible for connected devices to communicate with each other, even if

they aren't in close proximity to each other. Over time, the internet has become an increasingly vital technology that connects a larger, more diverse population of people and devices from around the world.

As with any form of communication, things go smoother when everyone speaks the same language, but there are lots of different hardware and software vendors who supply the many devices and applications connected by the internet. Fortunately, there are several organizations that help to establish and maintain "protocols" which allow the Internet to function despite the diversity of the connected devices.

One of the fundamental requirements for communicating on the internet is being able to identify and distinguish each connected device. This is key to ensuring you are communicating with the intended device. Think about sending a letter to someone. No matter what you write in the letter, you need to describe where the postal service should deliver it. To help with this, we use a standardized method of describing a location. When sending a letter, you write out the name of the recipient, then the components of their address in a specific format.

A similar process happens when sending messages over any network, but instead of using a physical address they use something called an internet protocol address, or IP address. Similar to physical addresses, IP addresses are written in a standard format to indicate how to locate a specific device. They are presented as a series of four numbers ranging from 0-255, which are separated by periods. These numeric values actually represent four 8-bit binary numbers but are presented in the other format to make them easier for humans to read. Each device on the network has a unique IP address which is used to "route" information to the correct destination.

IP addresses are also how websites and applications on the internet are identified. When you enter a URL to visit a site, a DNS (or Domain Name System) provides "domain resolution" to translate the registered domain name you entered, into the correct IP address of the site's web server.

Packets

While IP addresses identify the devices connected to the network, they don't deliver the messages. To do that, information is split into "packets", which you can think of as envelopes that hold the messages. Packets contain data (sometimes called the payload) and a "header" which describes some key information about the packet. The header includes things like the IP addresses of the sender and recipient, the size of the packet, and the order of the packets in the message (e.g. box 2 of 8). This information allows the packets to be routed without being "read' and allows the recipient to recombine the packets in the correct order after they arrive.

As files become larger and more traffic occurs on the internet, dividing files into packets helps to decrease the size of each "delivery". It also allows for the distribution of data delivery over several routes, meaning the information can be sent more quickly and doesn't rely on a single pathway.

Data in the Cloud

Servers

Another key component of "the cloud" is the concept of a server. A server is a dedicated computer that performs tasks for other devices (called clients). There are many types of servers which perform various tasks like file storage, hosting a database, hosting a website, or even hosting an application. Server hardware is often more robust than hardware found in personal devices because they serve many clients at the same time, often handle larger amounts of data, and are rarely powered down. To take advantage of this hardware, many servers run specialized operating systems.

In order for a server to perform tasks for clients, they need to be connected via a network. The network can be private (aka intranet) or public like the internet. A web server is a specific type of server that uses the internet and the Hypertext Transfer Protocol (aka HTTP) to receive and respond to requests from clients. For instance, when you want to visit a web page, your device sends a request for the page content via the internet. Your request is then routed to the correct server based on the URL you entered. That server processes the request, locates the appropriate files, and sends the requested information back to your device.

as a Service

as a Service

Now that we've talked about servers and the internet, we can talk more about what cloud vendors are actually offering. There are many cloud vendors and therefore lots of cloud products available, but they generally fall into three categories: Infrastructure as a Service (IaaS), Platform as a Service (PaaS), or Software as a Service (SaaS).

Infrastructure as a Service

Infrastructure as a Service offers the ability to "rent" hardware in order to host an environment of your choosing. This service is commonly used to upgrade web servers or expand storage into the cloud. IaaS still requires companies to install and maintain their environment (OS, applications, drivers, and updates) but allows them to benefit from modern hardware without having to purchase the newest hardware after each release.

Platform as a Service

Platform as a Service provides the same hardware options included with IaaS offerings but includes additional software like an operating system to create a "ready to go" environment. These products are popular for developers who need to operate in various platforms for testing or hosting an application. These options are highly customizable but also offer pre-packaged solutions for convenience, which ensures all hardware and software in the environment is compatible and up to date.

Software as a Service

Data in the Cloud

Software as a Service (or SaaS) delivers a final product to the end-user, usually a web application (e.g. an online document editor, calendar service, streaming service, etc.). To host these applications and allow end-users to compute in the cloud environment, the SaaS offering must appropriate hardware from the pool of resources, maintain a compatible software environment that hosts the front-end assets (like the website or application's interface), and perform any required tasks with the end-user's data.

Most people will only interact with SaaS products in their everyday life, but the IaaS and PaaS are often leveraged by the SaaS vendor to provide those products. Most of the cloud vendors sell individual services which allow customers to configure exactly what they need in the cloud, regardless of the service category those offerings fall under.

Benefits of the cloud

The cloud isn't just for public facing products. The flexibility and breadth of resources also appeals to businesses as they seek new ways of enhancing the systems and processes on which they've built their companies. In some instances, companies will simply migrate an existing solution to the cloud, meaning the same processes are simply "relocated". In other instances, those solutions need to be refactored or enhanced to offer additional capabilities in the cloud. Still other scenarios require an entirely new solution to be developed in a cloud-native environment. Each business evaluates their current processes and makes decisions on what, when, and how they can leverage the cloud going forward. So why are they spending resources to make this transition?

There are numerous benefits of cloud computing for both individuals and companies.

- Availability – By storing files in the cloud, any device can be allowed to access them. This also protects files from being lost if a personal device is lost or destroyed.
- Performance – rather than being limited by the hardware of a device, using the cloud allows systems to take advantage of more powerful hardware when it's required. Cloud vendors regularly update the hardware they offer which means improved technology without having to buy a new device every year.
- Flexibility – the cloud offers different operating systems, versions of software, and varying hardware configurations to achieve optimal results.
- Pricing – many cloud vendors offer exceptional value with pay as you go pricing for services which reduces the initial cost of starting a project when compared to buying a device up front.

**this is not an all inclusive list of benefits.*

Cloud-based Servers

Many of the tasks performed by servers are being moved into the cloud to take advantage of the larger pool of hardware resources available. Rather than being constrained by the limited physical hardware that a given company can afford to purchase and maintain, renting hardware

Data in the Cloud

from a cloud vendor and hosting from the cloud offers a wider variety of hardware and functionality. The datacenters created by cloud vendors are huge investments that are "shared" by many companies/users, which lowers the cost of running operations on the cloud. Because they are located around the world and interconnected, the cloud makes it easier to safeguard against outages or overburdening the hardware as might be the case with a private, physical server owned by a single company.

Scalability

Before cloud offerings became viable, each company or individual had to purchase hardware, install software and adjusting settings, then maintain that environment to ensure the system worked as intended and was up to date. This model is referred to as On-Premises (or on-prem) and it is still used for a variety of reasons, like regulatory requirements or data sensitivity.

However, hosting an On-Prem environment is a lot of work and can be expensive. Businesses need to find and hire knowledgeable workers, budget for hardware and labor, maintain physical and virtual security, ensure compatibility with other software and users in the company, etc. One of the big obstacles for hosting on-prem is "scalability" which refers to the system's ability to accommodate increases in workload over a short time.

Companies who host their own environment have to design their system to accommodate their busiest times. For instance, a company with a web site that gets an average of 5000 visitors per day may see over 50000 visitors when they have a promotion for the holidays. If their hardware is only capable of hosting 5000 users, their site will crash, and they will miss out on all the extra business for the holiday special. For this reason, they choose to invest in a system capable of handling more than 50000 users at once. This may seem like a good solution, but for the majority of the year their system is overpowered, and they've spent more than they would like to on hardware that they aren't using most of the time.

Scalability is one of the biggest benefits of cloud environments. By having datacenters with lots of hardware all networked together, it's easy to change your environment based on what you need. The cloud makes it easy to scale vertically (using more powerful equipment, like upgrading to a cpu with a faster clockspeed) or to scale horizontally (using more equipment, like adding an additional cpu). The ability to scale up and down is referred to as "elasticity". This elasticity can offer cost savings by scaling up hardware at busy times and scaling back down when appropriate.

Elasticity, Scalability, & Serverless

You may be wondering how the system knows when to scale up or down? Instead of relying on a system administrator or cloud engineer to log in and reconfigure the system, many systems utilize automated scaling solutions provided by cloud vendors. These solutions measure the workload processing and resource consumption and allow the admins to set up "triggers" that run code when conditions are met.

Data in the Cloud

There are also options to have the system take care of all hardware allocation without needing to set up triggers. These services simply use the resources they need when they need them, without any need for human intervention/approval.

While some cloud services are referred to as "serverless", there is still a server in operation. The difference is that the server is managed by the cloud vendor and doesn't require the administrators to perform the maintenance and administrative tasks that traditional servers require.

Virtual Server

As mentioned, the cloud offers flexibility of hardware to fit your use case via networked datacenters but having a bunch of hardware connected together isn't the same as a traditional server. This requires a way to separate out the resources allocated to each customer from the greater pool of hardware. To host multiple servers on the same hardware (known as virtualization), we need a "hypervisor" which is additional software that creates separate environments from the hardware resources. The resulting environments are known as "Virtual Machines" (or VMs).

In the cloud, these virtual server environments are logically isolated which prevents other cloud customers from seeing your data despite being hosted by the same cloud hardware. Hosting multiple instances from a shared environment is known as "multi-tenancy". Even though these virtual servers are logically isolated, additional layers of security best practices are required to ensure information doesn't fall into the wrong hands.

Security

Security

Security is a primary concern when discussing any cloud technology. Because of the cloud's reliance on the internet and sharing technologies with other parties, there is always a risk of exposure. If all of our valuable data or code is on the internet, what keeps it from falling into the hands of unintended people? The answer is complicated…so let's discuss a few of the key terms involved.

Users

The term user refers to anyone who will interact with an asset, whether it be a device, database, or website. Users are an important concept in computer systems because they allow us to keep track of who is interacting with what assets, and setup security to ensure only appropriate users can access sensitive items. It's common to store a list of users and their details in a database. These details can be captured when signing up for a service or website, or when you start a new job.

Data in the Cloud

In business settings, many organizations use something called a directory to keep an accurate list of employees. These directories are often tied to the company's organizational structure or hierarchy. Active Directory (or AD) is a common example of a directory which can be queried by/connected with other applications using the Lightweight Directory Access Protocol (or LDAP).

Authentication & Authorization

While often used interchangeably, Authentication and Authorization represent fundamentally distinct functions. In simple terms, authentication is the process of verifying who a user is, while authorization is the process of verifying what they have access to. Comparing these processes to a real-world example, when you go through security in an airport, you show your ID to authenticate your identity. Then, when you arrive at the gate, you present your boarding pass to the flight attendant, so they can authorize you to board your flight and allow access to the plane.

Whitelist (allow-list) authentication is another cyber security strategy, among many others that is leveraged to authenticate more than one user, service, domain name or applications while denying others.

MFA

While passwords are still a common form of authentication, they are often leaked or bypassed. Multi-factor authentication, or MFA, is a way to require additional authentication before providing access to secure assets. MFA usually requires a password and another method of authentication. One of the most common methods is to use a code generating application (e.g. Microsoft Authenticator, Google Authenticator) which generates a unique, one-time code that is only accepted for a limited time. Another common approach is to use physical tokens which often take the form of a usb drive (yubikey). These devices require the user to know the password and physically possess the device in order to gain access. Some environments may also offer biometric security, like fingerprint readers or face scanners.

Permissions

Another common security feature that is not exclusive to the cloud is permissions. Permissions are exactly what they sound like, the ability to do something. In computer environments, permissions are a great way to ensure each person has the ability to do what they need to while minimizing the chance that they accidentally cause an issue. One common example is someone sharing a link to view a document, but not edit. In this example, the recipient has permission to view the document but doesn't have permission to make any changes. This concept extends to databases, where the administrator has different permissions from those granted to end users. A security best practice is to limit user permissions to the least amount of access they require to perform their job. You may also hear the term "access" used to describe permission to view or edit assets.

Data in the Cloud

Roles and Groups

While permissions can be granted to individuals, systems usually allow for the creation of roles and groups which bundle standard permissions together. Roles and groups are similar but there are some slight differences.

Roles are usually designed to provide the permissions and access required for a specific job role. For instance a data analyst may only need read permission on a database and the ability to post reports in a separate folder. By setting up a role with the predetermined permissions all new data analysts can easily be given the access they require.

On the other hand, groups often bundle permissions by functional area. For example, all employees in the finance department may be assigned to a group that includes permissions for shared assets within the department like a database or server.

Users can be assigned to multiple roles or groups, and it is possible for the permissions they grant to overlap. When this happens, it is up to the organization to decide how to handle role and group permissions that contradict each other (i.e. which supersedes the other).

IAM

Many of the security features we've discussed are referred to as IAM which stands for Identity Access Management. These policies allow administrators to control access to assets in the cloud and comply with regulatory requirements for keeping data secure.

Subnetting

Another way of ensuring security in the cloud is by hiding significant components through "subnetting" which means only allowing certain types of traffic to interact with the assets. When architects create components in the cloud, they can limit the exposure of critical components (like a database) by assigning a non-internet facing IP address. This ensures that it only communicates with trusted sources.

SSO

While usernames and passwords are still widely used, it can be frustrating for users to have to remember different passwords for each of the sites and services they use. Instead of simply reusing a password (which is bad), many services offer SSO (or Single Sign On) as a way of reducing the friction for their users. Single Sign On works by using SAML (aka Security Assertion Markup Language) to sign into one service and then using that service to verify the user's identity in another service. Essentially, you are authenticated once, then that system can "vouch" for you so you don't have to log into the next service.

Encryption

Data in the Cloud

Encryption refers to encoding data, so that it isn't readable without decryption which usually requires a specific set of instructions (known as an algorithm). This is a commonly used tactic designed to make the data unusable in the event that the wrong person sees or acquires it (known as a breach). There are many types of encryption that use different methods to encode the data and vary in their difficulty to "crack". Some rearrange characters, replace them with different characters, or both. Encryption can also be deployed in different ways. For instance, a local computer's hard drive can be encrypted, data can be encrypted when stored in the cloud, and/or data can be encrypted for travel. One of the downsides to encryption is that it typically requires additional time to save or read files because there are additional steps required to encode and decode the data.

VPNs

Encryption is a key factor in creating a VPN (or virtual private network) which allows users to connect to a private network over public network infrastructure like the internet. VPNs are commonly used by organizations as a secure way for employees to access files from public internet connections. When logged into a VPN, an encrypted connection is created with the secure network (known as a tunnel). Some organizations protect their critical infrastructure by requiring users to be on the VPN in order to access certain assets.

APIs

APIs

Given that "the cloud" relies on the internet and that security is a primary concern, access to assets is often limited in order to keep them secure. However, one of the main benefits of the cloud is the ability for others to access assets. APIs, or application programming interfaces, are often used as a way for software and applications to communicate with each other without allowing "entry" into the backend or requiring human intervention.

APIs use endpoints which are "gateways" that allow communication with the "back end" of the system (like a database). Instead of landing on a web page and signing in to get access to data, you simply send over all the relevant pieces of information (like your credentials and the information you are requesting) and then you are sent back the information you requested**.** Since APIs are meant to be automated, they often require that the relevant pieces of information be formatted in a particular order. They are also usually encrypted to keep all communication secure. To help end-users understand and meet the required format of each API, they include detailed documentation on how information should be arranged.

There are many types of APIs which differ in the structure and protocols used. Two of the most common descriptors are REST (aka Restful) and SOAP APIs. They differ in the flexibility they offer for things like language and protocols used to send communications. SOAP is stricter because it's protocol requires XML, while REST APIs are often faster and can use JSON.

Data in the Cloud

APIs are critical for cloud-native application development because they allow for communication between “microservices” which are application sub-services. The application’s overall functionality is broken down into smaller functional components, each of which is accomplished as a microservice. These microservices are built, scaled, and maintained independently, which provides greater flexibility and elasticity than traditional architecture.

ODBC

One specific API that is used with databases is ODBC which stands for Open Database Connectivity. This API allows end users to request information from databases in a structured format, regardless of the database engine or operating systems in use. ODBC is a very common way for data analysts to access a database

ODBC is a standard and in order to communicate with the wide variety of database engines, it relies on engine specific drivers to “translate” the standard requests of ODBC into actionable commands in the database environment.

The Cloud

As you’ve seen, the cloud offers new capabilities and benefits to both end users and organizations, including availability, elasticity, cost effectiveness, and a global footprint. These improvements require a shift in how some technologies are leveraged and implemented when compared with traditional solutions. We can expect more capabilities to be developed as data and services continue to migrate to the cloud. If these concepts and emerging technologies excite you, there are many opportunities for growth, specialization, and careers within this space.