**Understand the difference between on-premises and cloud-based servers**

Completed100 XP

* 5 minutes

When traditional hardware and infrastructure components near the end of their life cycle, many organizations consider digital transformation projects. Here we'll consider options for those transformations. We'll look at features of both on-premises and cloud environments. We'll also cover the factors that businesses must consider as they explore each option.

**On-premises environments**

**Computing environment**

On-premises environments require physical equipment to execute applications and services. This equipment includes physical servers, network infrastructure, and storage. The equipment must have power, cooling, and periodic maintenance by qualified personnel. A server needs at least one operating system (OS) installed. It might need more than one OS if the organization uses virtualization technology.

**Licensing**

Each OS that's installed on a server might have its own licensing cost. OS and software licenses are typically sold per server or per CAL (Client Access License). As companies grow, licensing arrangements become more restrictive.

**Maintenance**

On-premises systems require maintenance for the hardware, firmware, drivers, BIOS, operating system, software, and antivirus software. Organizations try to reduce the cost of this maintenance where it makes sense.

**Scalability**

When administrators can no longer scale up a server, they can instead scale out their operations. To scale an on-premises server horizontally, server administrators add another server node to a cluster. Clustering uses either a hardware load balancer or a software load balancer to distribute incoming network requests to a node of the cluster.

A limitation of server clustering is that the hardware for each server in the cluster must be identical. So when the server cluster reaches maximum capacity, a server administrator must replace or upgrade each node in the cluster.

**Availability**

High-availability systems must be available most of the time. Service-level agreements (SLAs) specify your organization's availability expectations.

System uptime can be expressed as three nines, four nines, or five nines. These expressions indicate system uptimes of 99.9 percent, 99.99 percent, or 99.999 percent. To calculate system uptime in terms of hours, multiply these percentages by the number of hours in a year (8,760).

| **AVAILABILITY** | | |
| --- | --- | --- |
| **Uptime level** | **Uptime hours per year** | **Downtime hours per year** |
| 99.9% | 8,751.24 | (8,760 – 8,751.24) = 8.76 |
| 99.99% | 8,759.12 | (8,760 – 8,759.12) = 0.88 |
| 99.999% | 8,759.91 | (8,760 - 8,759.91) = 0.09 |

For on-premises servers, the more uptime the SLA requires, the higher the cost.

**Support**

Hundreds of vendors sell physical server hardware. This variety means server administrators might need to know how to use many different platforms. Because of the diverse skills required to administer, maintain, and support on-premises systems, organizations sometimes have a hard time finding server administrators to hire.

**Multilingual support**

In on-premises SQL Server systems, multilingual support is difficult and expensive. One issue with multiple languages is the sorting order of text data. Different languages can sort text data differently. To address this issue, the SQL Server database administrator must install and configure the data's collation settings. But these settings can work only if the SQL database developers considered multilingual functionality when they were designing the system. Systems like this are complex to manage and maintain.

**Total cost of ownership**

The term *total cost of ownership* (TCO) describes the final cost of owning a given technology. In on-premises systems, TCO includes the following costs:

* Hardware
* Software licensing
* Labor (installation, upgrades, maintenance)
* Datacenter overhead (power, telecommunications, building, heating and cooling)

It's difficult to align on-premises expenses with actual usage. Organizations buy servers that have extra capacity so they can accommodate future growth. A newly purchased server will always have excess capacity that isn't used. When an on-premises server is at maximum capacity, even an incremental increase in resource demand will require the purchase of more hardware.

Because on-premises server systems are very expensive, costs are often *capitalized*. This means that on financial statements, costs are spread out across the expected lifetime of the server equipment. Capitalization restricts an IT manager's ability to buy upgraded server equipment during the expected lifetime of a server. This restriction limits the server system's ability to accommodate increased demand.

In cloud solutions, expenses are recorded on the financial statements each month. They're monthly expenses instead of capital expenses. Because subscriptions are a different kind of expense, the expected server lifetime doesn't limit the IT manager's ability to upgrade to meet an increase in demand.

**Cloud environments**

**Computing environment**

Cloud computing environments provide the physical and logical infrastructure to host services, virtual servers, intelligent applications, and containers for their subscribers. Different from on-premises physical servers, cloud environments require no capital investment. Instead, an organization provisions service in the cloud and pays only for what it uses. Moving servers and services to the cloud also reduces operational costs.

Within minutes, an organization can provision anything from virtual servers to clusters of containerized apps by using Azure services. Azure automatically creates and handles all of the physical and logical infrastructure in the background. In this way, Azure reduces the complexity and cost of creating the services.

On-premises servers store data on physical and virtual disks. On a cloud platform, storage is more generic. Diverse storage types include Azure Blob storage, Azure Files storage, and Azure Disk Storage. Complex systems often use each type of storage as part of their technical architecture. With Azure Disk Storage, customers can choose to have Microsoft manage their disk storage or to pay a premium for greater control over disk allocation.

**Maintenance**

In the cloud, Microsoft manages many operations to create a stable computing environment. This service is part of the Azure product benefit. Microsoft manages key infrastructure services such as physical hardware, computer networking, firewalls and network security, datacenter fault tolerance, compliance, and physical security of the buildings. Microsoft also invests heavily to battle cybersecurity threats, and it updates operating systems and firmware for the customer. These services allow data engineers to focus more on data engineering and eliminating system complexity.

**Scalability**

Scalability in on-premises systems is complicated and time-consuming. But scalability in the cloud can be as simple as a mouse click. Typically, scalability in the cloud is measured in compute units. Compute units might be defined differently for each Azure product.

**Availability**

Azure duplicates customer content for redundancy and high availability. Many services and platforms use SLAs to ensure that customers know the capabilities of the platform they're using.

**Support**

Cloud systems are easy to support because the environments are standardized. When Microsoft updates a product, the update applies to all consumers of the product.

**Multilingual support**

Cloud systems often store data as a JSON file that includes the language code identifier (LCID). The LCID identifies the language that the data uses. Apps that process the data can use translation services such as the Bing Translator API to convert the data into an expected language when the data is consumed or as part of a process to prepare the data.

**Total cost of ownership**

Cloud systems like Azure track costs by subscriptions. A subscription can be based on usage that's measured in compute units, hours, or transactions. The cost includes hardware, software, disk storage, and labor. Because of economies of scale, an on-premises system can rarely compete with the cloud in terms of the measurement of the service usage.

The cost of operating an on-premises server system rarely aligns with the actual usage of the system. In cloud systems, the cost usually aligns more closely with the actual usage.

In some cases, however, those costs don't align. For example, an organization will be charged for a service that a cloud administrator provisions but doesn't use. This scenario is called *underutilization*. Organizations can reduce the costs of underutilization by adopting a best practice to provision production instances only after their developers are ready to deploy an application to production. Developers can use tools like the Azure Cosmos DB emulator or the Azure Storage emulator to develop and test cloud applications without incurring production costs.

**Lift and shift**

When moving to the cloud, many customers migrate from physical or virtualized on-premises servers to Azure Virtual Machines. This strategy is known as *lift and shift*. Server administrators lift and shift an application from a physical environment to Azure Virtual Machines without rearchitecting the application.

The lift-and-shift strategy provides immediate benefits. These benefits include higher availability, lower operational costs, and the ability to transfer workloads from one datacenter to another. The disadvantage is that the application can't take advantage of the many features available within Azure.

Consider using the migration as an opportunity to transform your business practices by creating new versions of your applications and databases. Your rearchitected application can take advantage of Azure offerings such as Cognitive Services, Bot Service, and machine learning capabilities.

Types of storage in cloud:

Azure Blob storage, Azure Files storage, and Azure Disk Storage.

As a data engineer you might use additional technologies like Azure HDInsight and Azure Cosmos DB. To manipulate the data in big-data systems, you might use languages such as HiveQL or Python.