SHRI RAMDEOBABA COLLEGE OF ENGINEERING AND MANAGEMENT

Department of Computer Science & Engineering
Session: 2021-22

CLOUD COMPUTING (CST452-4)
ASSIGNMENT

VII Semester, B.E.

Shift-I/II

Group Members

Atharva Bisen (B-16)

Jatin Jangir (B-47)

Soham Mundhada (B-78)

Course Co-ordinator

Prof.Heena Agrawal

Table of contents

_____

Data Center and Cloud as a Service

_____

Introduction

A Data Center is a physical facility which companies use to store their information as
well as other applications which are integral to their functioning. While a Data Center is
thought to be one thing, in reality, it is often composed of technical equipment
depending on what requires to be stored – it can range from routers and security devices
to storage systems and application delivery controllers. A data center's design is based
on a network of computing and storage resources that enable the delivery of shared
applications and data. To keep all the hardware and software updated and running, a
Data Center also requires a significant amount of infrastructure. The key components of
a data center design include routers, switches, firewalls, storage systems, servers, and
application-delivery controllers. These facilities can include ventilation and cooling
systems, uninterruptible power supplies, backup generators, and more.

With the growth of population increasing exponentially by the hour, key data that a firm
works on, which is used to draw important insights about the audience and potential
clients, has reached dimensions of the scale of petabytes. It requires huge data centers in
the form of huge data warehouses to store and manage such a humongous amount of
data properly. Data is the most critical asset of any organization and businesses are
faced with the imminent challenges of managing and governing data while ensuring
data compliance. Data management is critical for every company to improve business
agility with up-to-date information available anywhere, anytime to the employees who
need it most. There are entire ecosystems that grow perennially around Big Data and
Data Analytics, which make enterprises aim for significantly critical tools to manage
everyday data.

With businesses realizing the dynamism of what can be done with their data, they are
moving on from their existing resources to well-equipped Data Centers to aid better data
management. Data Centers have become top priority for businesses across the globe to
measure up their IT infrastructure requirements. With this shift in addressing
information, Data Centers have moved beyond being just an additional storage facility.
Infact, they have emerged as a key business parameter.

Role of a Data Center

Data centers are an integral part of the enterprise, designed to support business applications and provide services such as:

- Data storage, management, backup and recovery
- Productivity applications, such as email
- High-volume e-commerce transactions
- Powering online gaming communities
- Big data, machine learning and artificial intelligence

Types of Data Centers

Many types of data centers and service models are available. Their classification depends on whether they are owned by one or many organizations, how they fit (if they fit) into the topology of other data centers, what technologies they use for computing and storage, and even their energy efficiency. There are four main types of data centers:

1. **Colocation data centers** : In colocation ("colo") data centers, a company rents space within a data center owned by others and located off company premises. The colocation data center hosts the infrastructure: building, cooling, bandwidth, security, etc., while the company provides and manages the components, including servers, storage, and firewalls.
2. **Cloud data centers** : A cloud data center is the type of data center in which the cloud company manages and takes care of the actual hardware with the help of a third-party managed services provider. It gives the clients the liberty to run applications and manage websites and data within the ambits of a virtual infrastructure running on the cloud servers.
3. **Managed services data centers** : These data centers are managed by a third party (or a managed services provider) on behalf of a company. The company leases the equipment and infrastructure instead of buying it.
4. **Enterprise data centers** : These are built, owned, and operated by companies and are optimized for their end users. Most often they are housed on the corporate campus.

In-House Data Center Vs Cloud

Some of the major key differences are mentioned below:

- The major difference between the Data Center and Cloud is that the applications are offered locally and are accessible by users whenever needed without an internet connection. While in the cloud, the applications are online and a network connection is necessary to access the same.

- Data is stored in a public repository in Cloud whereas data is stored in the local repository in the Data Center. The repository is taken care of by service providers in the cloud and those are taken care of by developers in the Data Center.
- Investment is needed in the Data Center to set up and maintain the repository. Small companies and startups may find it hard to find investment. No large investments are needed for clouds as it is a subscription and companies can easily find the money invested for cloud deployment.
- Larger organizations go forth with Data Center. Small companies and startups go with cloud deployment. Less budget and resources force the companies to go with cloud services offered by third party services. Organizations find it easy to fund the repository and target the resources than to risk the security.
- Organizations feel hard and remorseful while recovering the data or application from the Data Center as it takes a long time and the data is not fully recovered. It is not hard to recover data from the Cloud if something goes wrong. Data is distributed in several remote servers and the data is easily recovered when called for once at least.
- The infrastructure of the cloud can be extended whenever needed and hence extra storage is not a burden for companies using Cloud. However, when extra storage is needed in the Data Center, it means extending the local servers, which in turn results in huge investment, again.
- Cloud is faster than the Data Center. This is because all the data is stored in different servers and it will not result in any cacophony while using the application. Data Center's speed depends on the network of the organization and the amount of data stored in the servers.
- Just as our application is being updated in the mobile phones or systems, cloud updates the software of the company. This does not happen automatically in the Data Center and developers need to take care of the same.
- Security is a concern for larger companies and hence they store the application in-house. This results in the Data Center. Since the applications are stored in Cloud (maybe in public), security cannot be offered as a benefit in Cloud though the service providers offer the same.

Need for Remote Data Centers to be accessible through Cloud

In the remote workplace, data centers play a significant role, businesses that previously did not have in-house data centers or IT departments needed to migrate to cloud-based services quickly. Working with a colocation data center is always a smart option as they will assist you in managing the IT infrastructure. The most significant aspect of remote work is the data center. They handle the traffic for cloud storage, meetings, VOIP traffic, and web-based applications. They also have access to your mission-critical business systems. It's critical to safeguard your data center's environment. Staff will access your network on a daily basis. As a result, they will be utilizing the IT assets

housed in colocation data centers. It is critical that this procedure be streamlined and secure.

Workers often prefer remote work, which may inspire loyalty to the company. It is expected that the COVID-19 pandemic has caused a permanent shift in working practices for many people. Despite these clear advantages, moving from a largely office-based workforce to a remote workforce poses a number of practical obstacles. Business owners may have difficulty accessing data or apps and avoiding security breaches/intrusions to their network infrastructure if they do not have a manned and managed centralized data center on their premises. Fortunately, there are third-party companies that can handle these concerns and ensure that a decentralized company model runs well. These companies are known as colocation data centers, and their popularity has skyrocketed in recent months.

Cloud Data Centers

Cloud storage is a system that provides functions such as data storage and business access. It assembles a large number of different types of storage devices through the application software which are based on the functions of the cluster applications, grid techniques, distributed file systems, etc. Cloud storage can be simply understood as the storage in cloud computing, and also can be considered to be a cloud computing system equipped with large capacity storage. Cloud storage system architecture mainly includes storage layer, basic management layer, application interface layer and access layer.

Cloud storage technology of enterprises

1. **Amazon Web Services (AWS)**

Amazon Web Services (**AWS**) is a subsidiary of Amazon that provides on-demand cloud computing platforms and APIs to individuals, companies, and governments, on a metered pay-as-you-go basis. These cloud computing web services provide distributed computing processing capacity and software tools via AWS server farms. One of these services is Amazon Elastic Compute Cloud (EC2), which allows users to have at their disposal a virtual cluster of computers, available all the time, through the Internet. AWS's virtual computers emulate most of the attributes of a real computer, including hardware central processing units (CPUs) and graphics processing units (GPUs) for processing; local/RAM memory; hard-disk/SSD storage; a choice of operating systems; networking; and pre-loaded application software such as web servers, databases, and customer relationship management (CRM).

AWS services are delivered to customers via a network of AWS server farms located throughout the world. Fees are based on a combination of usage (known as a "Pay-as-you-go" model), hardware, operating system, software, or networking features chosen by the subscriber required availability, redundancy, security, and service options. Subscribers can pay for a single virtual AWS computer, a dedicated physical computer, or clusters of either. Amazon provides select portions of security for

subscribers (e.g. physical security of the data centers) while other aspects of security are the responsibility of the subscriber (e.g. account management, vulnerability scanning, patching).

2. **Azure**

The Azure Storage platform is Microsoft's cloud storage solution for modern data storage scenarios. Azure Storage offers highly available, massively scalable, durable, and secure storage for a variety of data objects in the cloud. Azure Storage data objects are accessible from anywhere in the world over HTTP or HTTPS via a REST API. Azure Storage also offers client libraries for developers building applications or services with .NET, Java, Python, JavaScript, C++, and Go. Developers and IT professionals can use Azure PowerShell and Azure CLI to write scripts for data management or configuration tasks. The Azure portal and Azure Storage Explorer provide user-interface tools for interacting with Azure Storage.

Microsoft Azure Architecture : Like other cloud platforms, Microsoft Azure depends on a technology called virtualization, which is the emulation of computer hardware in software. This is made possible by the fact that most computer hardware works by following a set of instructions encoded directly into the silicon. By mapping software instructions to emulate hardware instructions, virtualized hardware can use software to function like "real" hardware. Cloud providers maintain multiple data centers, each one having hundreds (if not thousands) of physical servers that execute virtualized hardware for customers. Microsoft Azure architecture runs on a massive collection of servers and networking hardware, which, in turn, hosts a complex collection of applications that control the operation and configuration of the software and virtualized hardware on these servers. This complex orchestration is what makes Azure so powerful. It ensures that users no longer have to spend their time maintaining and upgrading computer hardware as Azure takes care of it all behind the scenes.

3. **Google File System (GFS)**

4. **Hadoop Distributed File System (HDFS)**

Hadoop is hosted by the Apache Software Foundation, which provides support for a community of open source software projects. Although Hadoop is best known for MapReduce and its distributed file system (HDFS), the other subprojects provide complementary services, or build on the core to add higher- level abstractions. The detailed contents refer to the document.

The full name of HDFS is Hadoop Distributed File System. HDFS is run on large clusters of commodity hardware and is like GFS of Google. The architecture of HDFS is master/slave and a HDFS cluster has one namenode and multiple datanodes. Namenode is the central server, equivalent to master in GFS. It is responsible for the

namespace operation of file systems. Datanode is similar to the chunkserver of GFS which is responsible for managing storage on datanodes, creating blocks, deleting blocks, copying blocks etc. The files in HDFS are divided into one or multiple blocks which are stored in datanodes. Namenode and datanodes can be run on the low-cost Linux computer. HDFS is developed in the Java language.

Difference between Cloud data Centers and normal Data Centers

Features	Traditional data centers	Cloud data centers
Server	Co-located dependent failure	Integrated Fault-Tolerant
Resources	Partitioned Performance Interrelated	Unified performance Isolated
Management	Separated manual	Centralized full control with automation
Scheduling	Plan ahead overprovisioning	Flexible Scalable
Renting	Per Physical machine	Per logical usage
Application/ Services	Fixes on designated servers	Runs and move across all VMs

How does cloud integration of data centers work

Datacenter management is required to administer many different topics related to the data center, including:

- **Facilities Management**. Management of a physical data center facility may include duties related to the facility's real estate, utilities, access control, and personnel.
- **Datacenter inventory or asset management**. Datacenter features include hardware assets and software licensing, and release management.
- **Datacenter Infrastructure Management**. DCIM lies at the intersection of IT and facility management and is typically accomplished by monitoring data center performance to optimize energy, equipment, and floor use.
- **Technical support**. The data center provides technical services to the organization, and as such, it should also provide technical support to the end-users of the enterprise.

● Datacenter management includes the day-to-day processes and services provided by the data center.

Data Center Virtualization

Data center virtualization is the transfer of physical data centers into digital data centers using a cloud software platform, so that companies can remotely access information and applications.

In a virtualized data center, a virtual server, also called a software-defined data center (SDDC) is created from traditional, physical servers. This process abstracts physical hardware by imitating its processors, operating system, and other resources with help from a hypervisor. A hypervisor (or virtual machine monitor, VMM, virtualizer) is a software that creates and manages a virtual machine. It treats resources such as CPU, memory, and storage as a pool that can be easily reallocated between existing virtual machines or to new ones.
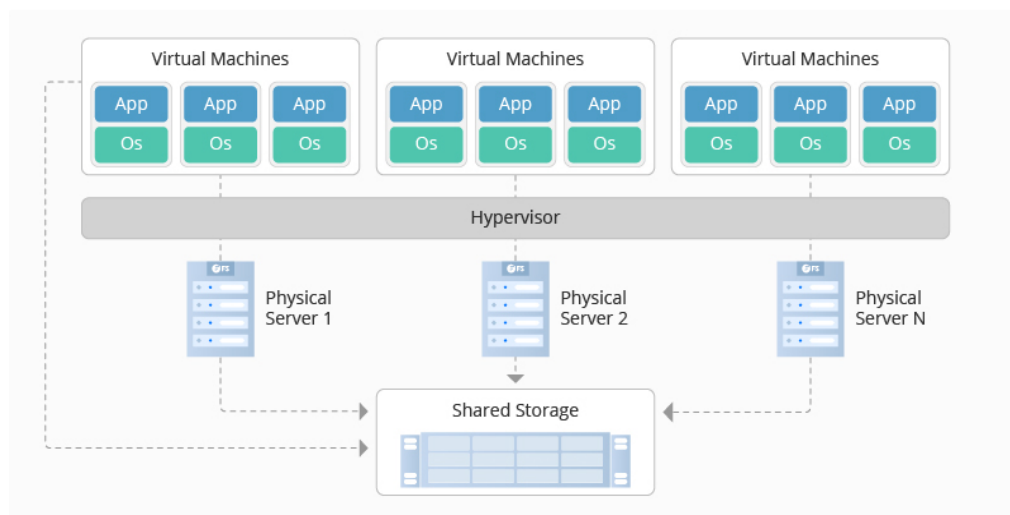
Benefits of Data Center Virtualization

Data Center as a Service (DCaaS)

Data center as a service is the provisioning of physical data center facilities and infrastructure to a client. Though it sounds very much like cloud computing's infrastructure as a service (IaaS) and software as a service (SaaS), it is more like renting out a house or an office.

DCaaS infrastructure and facilities are uniquely tailored according to a client's needs. For example a client can opt to use their own hardware and the DCaaS provider can

simply provide the space required. The space can be constructed, even with sensors and controls enabling the client to supervise changes in the computing environment. A major benefit of DCaaS providers is that they can offer expansion for data centers that can no longer do so for various reasons such as lack of physical space, power, cooling or capital. This remotely rented data center is usually controlled remotely via WAN, making WAN downtime or disruption one of the biggest concerns with DCaaS providers, as it interrupts access to services and availability of applications.

Need a Data Center as a service

1. DCaaS solves all the problems such as Lack of power, cooling, rack space, capital, bandwidth, and experienced IT staff.
2. DCaaS saves costs compared to building a private data center.
3. DCaaS provides flexibility and a cost-effective solution.
4. DCaaS provides an effective solution for companies using hybrid (Public cloud and Private cloud) and multi-cloud solutions.
5. With DCaaS it is easier to scale up and down resources than an on-site **Data Center**.
6. DCaaS offers a Pay as you go service that helps in saving the cost. Moving to the OpEx model from CapEx saves the money for more strategic initiatives.
7. DCaaS can be customized or tailored as per the customer's requirement.
8. DCaaS also provides backup and redundancies to keep the data up and running
9. A DCaaS provider makes sure of the uptime as per the SLA and also provides security to its clients.
10. DCaaS allows you to get started right away rather than waiting for capital expenditure and then setting that up.
11. DCaaS incorporates compliance assurances into operations and policies.

Key benefits of DCaaS

Whether a company is facing physical limitations, such as space in a server room, or limitations relating to the time and costs involved in infrastructure upgrades, DCaaS solves the problem. For this reason, many organizations are choosing DCaaS to prepare for the future, particularly if they've got big plans to expand into new markets, or they're starting to introduce more interconnected digital services for their customers.

Efficiency is key for any company looking to compete. As the capabilities of competitors grow, it's becoming ever more important that brands are able to go one step further, continuously improving their own offering. Opting for DCaaS is a good way of doing this, as it rapidly boosts efficiency in ways that customers will definitely notice.

Cost is another key consideration for the vast majority of companies, and DCaaS offers benefits here too. Compared to the costs involved in setting up and maintaining private IT infrastructure, DCaaS is far more affordable. And with DCaaS, this affordability

doesn't mean compromises in terms of the quality of the services companies have access to.

Unexpected downtime can be a major source of disappointment for both brand owners and their customers. When brands rely on cloud-based services, these periods of downtime can quickly become a problem. But DCaaS services seek to eliminate frustrations like these, by offering continuous uptime that companies can always rely on. The flexibility of these services mean that providers are able to get companies back up and running at a moment's notice, so customers are never left in the lurch.

DCaaS Advantages and Disadvantages

The three primary advantages of DCaaS include:

- Scalability Pay-as-you-go (SaaS) model allows flexibility to scale up/down based on business needs, only paying for what is used.
- Expandability Data centers experiencing limitations can leverage additional resources.
- ProductivityIn-house IT staff can spend more time focusing on other mission-critical needs.

A few of the biggest concerns with DCaaS providers surround availability and business continuity.

- WAN disruptions or provider downtime can leave applications inaccessible.
- Although SLAs should address downtime/availability issues, a business should carefully consider the potential ramifications of unexpected downtime.
- Remember that DCaaS providers are still just a business themselves; staffing issues, mergers, and even going out of business are all challenges that could potentially require mitigation in the future.

Difference between DCaaS and other services

Data Center as a Service works differently to other popular data center services. This option enables companies to access the full capabilities of their chosen data center, without having to physically visit the center.
Choosing DCaaS is a great way of future proofing, as the services that the chosen data center provides can expand in line with a company's growth. DCaaS enables companies to avoid both the physical and budgetary limitations of expanding private infrastructure, meaning they can grow quickly as and when they need to.

For companies looking to outsource their IT requirements, DCaaS is a great option. Packages include everything that a company might need, and can be quite flexible depending on the size and budget of the organization in question. DCaaS solutions will usually include a wide range of resources, from data storage and servers to networking. When a company chooses DCaaS, it benefits from a fast and easy setup process, which is ordinarily handled by the provider in question. There's therefore very little work involved for the company's own time-pressed teams, so it's a good option for companies that are having to expand quickly to meet customers' demands.