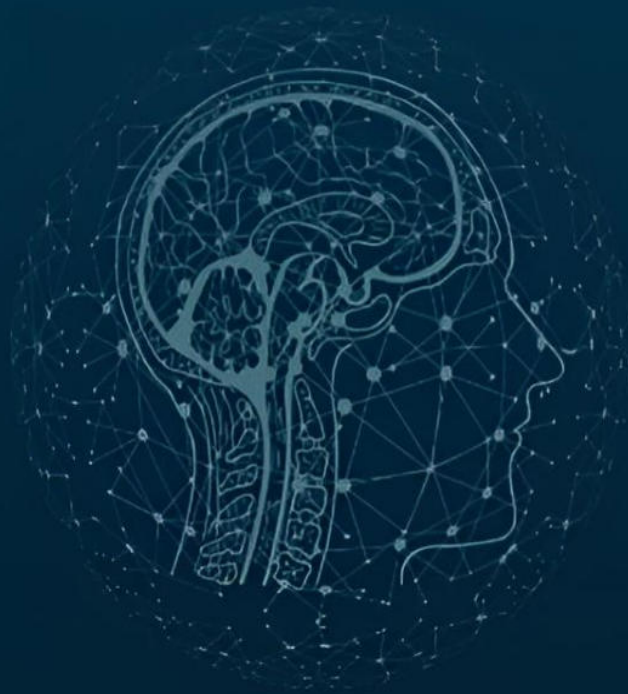


TECHBYTE

INSIGHT 2024

ANNUAL IT MAGAZINE



*Sustainable Employability in Next Gen:
Technological Provisions*



JAGAN INSTITUTE OF MANAGEMENT STUDIES

SECTOR-5, ROHINI

Jagan Institute of Management Studies (JIMS) in 30 years of its being has created a niche for itself in the fields of Management and Information Technology. The institute offers several undergraduate and postgraduate programmes i.e. BBA, BCA, B.A. Eco (H) affiliated from Guru Gobind Singh Indraprastha University.

Postgraduate programs namely, PGDM, PGDM (International Business), PGDM (Retail Management) in our campus are approved by the All India Council for Technical Education and are accredited from National Board of Accreditation (NBA). All programs have also been granted equivalence to MBA degree by Association of Indian Universities (AIU). The National Assessment and Accreditation Council (NAAC) has accredited JIMS with A grade.

Our MCA programme is affiliated to Guru Gobind Singh University, Delhi and is also accredited by National Board of accreditation. The institute is also among the top management colleges in Delhi to start a Fellow Programme in Management, a full-time doctoral level programme to equip students for careers in research and management.

The institute has been awarded the prestigious accreditation certification named South Asian Quality Standards (SAQS) from the Association of Management Development Institutions in South Asia, a SAARC recognized body. We are also a member institution of AACSB, a prestigious international accreditation body.

JIMS continues to remain in the list of ELITE B-SCHOOLS (Top 100) of India for the 7th year in a row in National Institutional Ranking Framework (NIRF) of the Ministry of Education, Govt. of India.

TECHBYTE 2024

20th Annual IT Symposium

“Sustainable Employability in Next Gen: Technological Provisions”

Department of Information Technology

Jagan Institute of Management Studies,
Sector-5 Rohini,
New Delhi-110085

CHAIRMAN

Mr. Manish Gupta

DIRECTOR

Dr. Pooja Jain

PATRON

Dr. Praveen Arora
Principal JIMS , IPU-Affiliated
Programmes

CONVENORS/EDITOR

Dr. Manjot Kaur Bhatia
Dr. Chetna Laroia

STUDENT EDITORS

Kanishk Arya (MCA),
Sahzuher Shah (MCA),
Anchal Jha (MCA),
Gaurav Sati (MCA),
Sachin Nayal (MCA)

DATE

Saturday, 03rd February 2024

VENUE

Auditorium, JIMS Sector-5, Rohini
New Delhi-85

Editor's Desk

As once stated by Swami Vivekananda, — “A knowledge gained through perception is more reliable than verbal understanding”. In this era of rapid technological advancements and digital transformation, it is essential for us to embrace and adapt to the changes that are occurring. To effectively do so, it is crucial to have a comprehensive understanding of the nature and impact of these changes. By recognizing and comprehending the transformations taking place, we can better navigate and leverage the opportunities presented by emerging technologies. We at JIMS, our goal is to provide students with the opportunity to excel academically while also emphasizing the importance of developing essential skills and attributes for their personal growth and development. By prioritizing these aspects, JIMS aims to equip students with the necessary tools to navigate and thrive in a rapidly changing world of technological advancements.

Techbyte serves as a transformative gateway for students, enabling them to develop a profound comprehension of current and upcoming technologies while staying updated on industry standards. This platform not only provides comprehensive training to prepare students for the industry but also offers valuable hands-on experience in event management through volunteering opportunities in organizing the prestigious symposium. We extend our heartfelt gratitude to the management of JIMS and the esteemed faculty members for their remarkable contributions. We highly value and appreciate the dedicated hard work and efforts demonstrated by our students.



Chairman's Message

Information Technology is ever growing industry which is producing millions of jobs for professionals from multiple fields. India has accomplished a terrific growth in this field and is being considered as hub for producing IT professionals for the world. Information Technology is providing the upcoming future that will change every face of Human existence.

Technology transformation is the integration of digital technology into all areas. It helped us to transform to reduce resource consumption and waste, increase efficiency, moderate environmental impacts and solved various business issues. When it comes to build a career in this technical era knowledge and skills are vital to make one stand out compared to others. JIMS, believe to impart the knowledge and nurture creative thinking.

Our objective is to provide quality IT education to students and skills needed by IT Professionals in the country. In this direction an annual IT symposium Techbyte is a platform for the students to interact with experts from the industry. Insight is a bunch of articles by our learning students and their learned faculty mentors.

I give my best wishes to the whole team of Techbyte and wish it to be a wonderful learning experience for all.

Manish Gupta



Director's Message

We have seen emergence of new technologies at a rapid pace since last decade. The advancement of disruptive technologies accelerates the reskilling requirements as well as the requirement of fresh talent.

As the world is experiencing digital transformation in Industry 5.0, we are experiencing a paradigm shift that has profound implications for the workforce and will affect talent management strategies, innovation, and business models. The future-ready workforce need to upskill and reskill on a continuous basis to accelerate their career progression and in order to be relevant for industry.

As the director of the institute, I'm extremely happy that JIMS also encourages, inspires and nurtures young students by training them on latest technologies, encouraging them to innovate and deliberate on new ideas to stay abreast with times.

An annual IT magazine is a renowned publication associated with Techbyte which provides a platform for students, IT professionals and academic experts to share their opinions on advances in the area of science and technology.

Wishing the entire team of Techbyte a great success!

Dr. Pooja Jain



Principal's Message

The world we're stepping into is buzzing with challenges and opportunities, none more crucial than building a future both thriving and sustainable. This isn't just about finding a job; it's about finding our purpose, a path where skills and passions align with a commitment to environmental responsibility and social justice.

That's what sustainable employability is all about: harnessing your potential to build careers that contribute to a healthier planet and a more equitable society. It's about becoming the architects of a greener future, the innovators, the leaders who bridge the gap between economic prosperity and environmental well-being.

This magazine "Insight" is a launchpad, a catalyst for igniting the inner eco-warrior, social innovator, and green genius. So dive in, explore, and let's rewrite the narrative together.

Remember, the future isn't just happening to us; we're the ones actively shaping it. Let's make it green, let's make it just, let's make it ours.

I wish TechByte 2024 a great success.

Dr. Praveen Arora

Principal, JIMS – IPU Affiliated Programme

CONTENTS

CLOUD COMPUTING TOOLS	01
SERVER VIRTUALIZATION: ORGANIZATION PERSPECTIVE	05
IMPORTANCE OF BIG DATA VIRTUALIZATION	07
FULL VIRTUALIZATION AND ITS TRANSFORMATIVE APPLICATIONS	11
DATA VIRTUALIZATION AND ETL	14
A GLIMPSE OF PARA-VIRTUALIZATION	19
CLOUD COMPUTING AND SECURITY ISSUES	21
EMULATION AND VIRTUALIZATION IN CLOUD COMPUTING	25
AI AND DATA VIRTUALIZATION: UNLEASHING THE POWER OF INSIGHTS	29
ABOUT DEPARTMENT OF INFORMATION TECHNOLOGY	32

DISCLAIMER

Great care has been taken in the compilation of information and every effort has been made to ensure that all Information is up-to-date at the time of going to press. The responsibility of the authenticity of articles lies with student writers. The institute or the editorial team is not responsible for error, if any, and their consequences.

CLOUD COMPUTING TOOLS

Faculty mentor:

Dr. Latika Kharb

Students Name:

Neha Singla (MCA - Year 1)

Harshita Sahni (MCA - Year 1)

INTRODUCTION

Cloud computing offers hardware as a service; This means that resources are owned and managed by the cloud rather than the end user. These sources may include browser-based software (such as TikTok or Netflix), third-party storage of photos and other digital media (such as iCloud or Dropbox), or third-party servers used to support computing. Business, research or personal study. Before cloud computing became widespread, businesses and casual computer users often had to purchase and manage the software and hardware they want to use. With the increasing availability of cloud-based applications, storage, services and technology, businesses and consumers can now access more information on demand, including network-based services. The shift from local software and hardware to remote connections and shared devices means cloud users don't need to invest in labour, capital, or smart skills to purchase and manage the computations. This unprecedented access to computing has directed in a new wave of cloud computing, transforming IT practices across businesses and replacing many computer-enabled applications. Thanks to the cloud tools, people can now collaborate with colleagues through video conferencing and other collaboration tools, access entertainment and learning content on-demand, communicate with home devices, hail taxis, stay at home, and book rooms using their mobile devices. IaaS, PaaS and SaaS are three most popular types of cloud service offerings (*IaaS vs. PaaS vs. SaaS* | IBM, n.d.). These services are specifically designed to increase the efficiency, scalability and adaptability of cloud computing environments. This versatile tool helps with everything from setting up virtual servers to manage data storage to simplify deployment and monitoring operations. It enables individuals and organizations to hold the power of cloud infrastructure to host applications, store data, and operate without investing in physical hardware

or infrastructure, solving the complexity of property management.

TYPES OF CLOUD COMPUTING TOOLS

Depending on the type of cloud service we use, the cloud service provider will manage the applications, operations, etc. It may provide some software and dependencies required to run it. But to truly understand what cloud computing tools are, what they are used for, and why they are important, it is important to understand where they fit into the cloud balance. Determining the best cloud computing tools is important and depends on specific needs, preferences, and the nature of the task at hand. Different tools perform well in different areas of cloud computing.

1. *Amazon Web Services*

Amazon Web Services (AWS) provides many and optional cloud computing tools and APIs for businesses that want to share resources. It provides a virtual environment where users can download and distribute various applications and services. Provides: network and content distribution, security, identity and compliance management, cloud storage and database, Amazon AppFlow, development intelligence, Upstream 2.0 and other AWS services.

2. *Google App Engine*

Google App Engine is a highly ascendable virtual environment that allows organizations to expand and run applications. It can be used for testing, high-performance web development and property maintenance. Its features: traffic segmentation and accelerating time to market, pay-as-you-go subscription, Python, PHP, .NET, JAVA and C# support,

- real-time cloud monitoring and cloud logging
3. **Cloud Hub**
Cloud Hub is a fully managed cloud service that allows deployment of aggregated applications across the cloud and allows users to create new API functions. It provides: multi-tenancy, flexibility and security, runtime admin console, high availability and connectivity to third-party applications and programs
 4. **Cloudability**
Cloudability evaluates an organization's budget, help identify opportunities to reduce costs and increase revenue, provides budget guidance. Features: Security API integration and multi-cloud calling, instant sync and compliance management, dashboard analytics, Pro and Enterprise subscription plans available (*Cloudability - Cloud Cost Management & Optimization - Apptio*, n.d.).
 5. **Cloudyn**
Cloudyn provides detailed information about its data, computing power, and data storage capabilities. It provides a dashboard view and prevents users from purchasing products through the Amazon cloud. Features include effective cost analysis, simple user interface and intuitive control panels, use of models and cost analysis, recommendations for unused resources, reduction, rework and rework
 6. **Informatica**
Informatica is a business tool for extracting, transforming and cleaning raw data. It provides fair understanding and transformation of raw data, has good data cleansing and maintenance, has automatic export, and can handle many operations simultaneously.
 7. **Chef**
Chef is used to set up a virtual machine in the cloud environment and also allow you to work on the workbook. Its features: data backup and recovery services, compliance training and security audits, infrastructure audits, automatic alerts and alerts across multiple platforms, multiple local and continuous deployments, automated testing, code review
 8. **Puppet**
Dummy applications and compute the rework and rapid deployment of systems' critical services (25+ *Best Cloud Computing Tools in 2023*, n.d.). Includes weed reporting compliance, open source and expansion, patch updates, application management and more.
 9. **AtomSphere**
AtomSphere integrates its functionality with other cloud applications and is a multi-tenant solution running as a single example. It allows the drag-and-drop workflows, has tons of connectors, is scalable, native, and provides higher learning.
 10. **RightScale**
RightScale allows easy deployment of applications in public, private and hybrid environments. It has a good management system, high performance and multi-cloud capabilities, it can move the office and provide data management in the cloud.
 11. **Enstratius**
Enstratius provides cross-level cloud infrastructure across public, private and hybrid cloud environments. It comes with features like self-service and single sign-on, role-based management tools, traceback and authentication to manage all cloud services.
 12. **Agility Platform**
Agility Platform is a mobile CMS solution that supports content creation, asset management and development. Features include SEO and marketing analysis, internal search and version control, rich text, user responsibility and privacy management
 13. **Netdata.Cloud**
Netdata.Cloud is a unique design tool by machine learning that helps users to figure out all the important metrics. It reduces the response times and improves resource management, and moreover relies on machine learning to detect anomalies, and provides real-time data visualization.
 14. **Cloudzero- Cloud cost Intelligence Platform**
Cloud Cost Intelligence is a broad and wide range of business-related cloud data related to understanding the leverage of the business cloud and its impact on business

finances. The right context for cloud pricing intelligence is important because without it, the impact of cloud pricing on business results cannot be properly analysed. Having cloud cost information means you can answer three important cloud-related questions: Where do the costs come from? How can I reduce costs? What are the future costs?

15. Lacework

Lacework provides continuous cloud security and managed compliance at scale. You can collect, analyze and correlate security threats across multiple cloud platforms such as AWS, Google Cloud Platform and Azure; so, you can prioritize addressing them. This security-based service provides vulnerability, security, and compliance management for all applications, processes, containers, processes, systems, accounts, and users in your environment (*Why Lacework - Lacework, n.d.*).

16. Notion

Notion is an all-in-one workplace, collaboration, knowledge and knowledge sharing platform. Teams can easily and frequently share information, tasks, schedules, project progress, and more without having to share multiple tools. Instead of starting from scratch, one can choose a template and change it whenever he/she want. Thinking is also known for its reporting, which allows you to record insights, ideas, and events from meetings, brainstorming sessions, and more and share them with others or correct yourself later.

17. Civis Analytics

If we are looking for a service that will help your team centralize, cleanse, support and use data at scale, Civis Analytics will be a good choice. The service allows you to collect data from trusted third parties and to gather experience with your data, share it across your organization, and inform decision-making. We can work with data in your favorite language, from SQL to Ruby to Python.

18. PagerDuty

PagerDuty's end-to-end incident response tool provides a unified workforce to manage incident management at scale.

PagesDuty has solutions for everyone, from DevOps and AIOps for security and customer service. Instant call management and notifications makes everything easy for teams to schedule the right people for the right job at the right time. Tools similar to PagerDuty include Big Panda, Splunk On-Call, ServiceNow's Lightstep Incident Response, and more.

19. Github

GitHub is the most popular cloud-based version control, collaboration and code repository hosting service for Git developers. Git's decentralized version control model allows distributed teams to build, distribute and develop software from anywhere. The platform is also popular among developers as a discussion, open-source project repository, and coding project management tool. GitHub alternatives include GitLab, BitBucket, AWS CodeCommit and more

20. Gremlin

Chaos engineering is a good strategy to prevent failures in cloud computing. Software developers use controlled testing to identify weaknesses, learn where things might go wrong, and practice what to do if something goes wrong. With Gremlin, you can use CPU spikes, server shutdowns, injection delays, process kills, and block DNS access to reveal vulnerabilities in your system. Additionally, you can test the disaster recovery process to avoid a false sense of security. Gremlin's competitors include Chaos Toolkit, Litmus, ChaosMesh, etc (*Chaos Engineering, n.d.*).

21. Cloudsfer

Cloudsfer is a cloud migration service that supports cloud-to-cloud and local migration. Features: Transfer data to 20 locations and backup and restore in the cloud.

22. Microsoft Azure

Microsoft Azure is Microsoft's cloud computing platform that provides many services similar to AWS. It includes computing, analytics, storage, database and other solutions. Azure is a popular choice for organizations using Microsoft technology because it integrates well with other Microsoft products. Features: Comprehensive cloud services, tight

integration with Microsoft products and various business solutions (*What Is Azure—Microsoft Cloud Services | Microsoft Azure*, n.d.).

23. **Oracle Cloud**

Oracle Cloud is the cloud services of Oracle Corporation. It provides a variety of cloud services including compute, storage, database, networking and business applications. It is generally preferred by organizations investing in Oracle software and technology. It comes with features like business-grade database services, a focus on security and compliance, and integration with Oracle Enterprise Applications.

24. **Terraform**

Terraform is an open source Infrastructure as Code (IaC) tool developed by HashiCorp ((22) *How Terraform Work as a IAC | LinkedIn*, n.d.). It allows users to define and configure the process using notification. Terraform supports many cloud providers, making it a popular choice for cloud management. It is an open-source Infrastructure as Code (IaC) tool that allows users to define and configure protocols using declarative expressions.

25. **DigitalOcean**

DigitalOcean is a cloud infrastructure focused on simplicity and green development. It is popular among beginners and individual developers due to its ease of use and cost-effective solution. It has a developer-centric cloud platform that focuses on simplicity, making it affordable and easy to use for startups and developers (*DigitalOcean | Cloud Hosting for Builders*, n.d.).

CONCLUSION

Cloud computing can be used to provide easy access to high performance, communications, and Internet storage. Future work should be directed towards data science/AI/ML services to protect user data and make data more secure. Cloud computing is a new technology that provides easy computing and access to high performance, communication, storage and internet infrastructure. Cloud computing has become an important part of IT and will be used by many large organizations. Some offer IaaS, PaaS, and some offer

SaaS. Amazon.com, Sun, and IBM provide storage services, while Google Apps provides software as a service. In the future, we will work on information science, artificial intelligence and machine learning services in the cloud to provide services to protect sensitive information documents, such as access to customers' credentials through encryption and other password protection technologies in the security team. increases efficiency and accuracy. safer information. Cloud computing will impact much of the computer industry, including software companies and Internet service providers. Cloud computing allows companies to easily deliver products to end users without having to worry about setting up server hardware and other requirements.

REFERENCES

1. *How Terraform Work as a IAC | LinkedIn*. (n.d.). Retrieved November 8, 2023, from <https://www.linkedin.com/pulse/how-terraform-work-iac-akash-patil-znyvf/>
2. *Best Cloud Computing Tools in 2023*. (n.d.). Retrieved November 8, 2023, from <https://www.knowledgehut.com/blog/cloud-computing/cloud-computing-tools>
3. *Chaos Engineering*. (n.d.). Retrieved November 8, 2023, from <https://www.gremlin.com/chaos-engineering/>
4. *Cloudability - Cloud Cost Management & Optimization - Apptio*. (n.d.). Retrieved November 8, 2023, from <https://www.apptio.com/products/cloudability/>
5. *DigitalOcean | Cloud Hosting for Builders*. (n.d.). Retrieved November 8, 2023, from <https://www.digitalocean.com/>
6. *IaaS vs. PaaS vs. SaaS | IBM*. (n.d.). Retrieved November 8, 2023, from <https://www.ibm.com/topics/iaas-paas-saas>
7. *What is Azure—Microsoft Cloud Services | Microsoft Azure*. (n.d.). Retrieved November 8, 2023, from <https://azure.microsoft.com/en-us/resources/cloud-computing-dictionary/what-is-azur>

Server Virtualization: Organization Perspective

Faculty Mentor:

Dr. Deepshikha Aggarwal

Students Name:

Kaushal (MCA - Year 1)

Radhika Khattar (MCA – Year 1)

INTRODUCTION

Server virtualization is a technology that allows multiple virtual servers to run on a single physical server. It enables organizations to consolidate their IT infrastructure, reduce hardware costs, and increase efficiency by running multiple applications and operating systems on a single server. Virtualization software, also known as a hypervisor, creates a layer of abstraction between the physical hardware and the virtual servers, allowing them to share the resources of the physical server. This technology has become increasingly popular in recent years due to its ability to improve server utilization, reduce energy consumption, and simplify IT management.

Why is Server Virtualization Important?

Server virtualization is important for several reasons. Firstly, it allows organizations to consolidate their IT infrastructure by running multiple virtual servers on a single physical server, which can help reduce hardware costs and increase efficiency. Secondly, virtualization can help organizations to improve server utilization, reduce energy consumption, and simplify IT management. Thirdly, it can help organizations to increase their flexibility and agility by enabling them to quickly provision new servers and applications as needed. Finally, virtualization can help organizations to improve their disaster recovery capabilities by enabling them to easily backup and restore virtual servers.

1. Brief Data Collected on Server Virtualization

Server virtualization has become increasingly popular in recent years. According to a report by Markets and Markets, the global server virtualization market size is expected to grow from USD 4.4 billion in 2020 to USD 7.7 billion by 2025, at a Compound Annual Growth Rate (CAGR) of 11.7% during the forecast period. The report cites factors such as the increasing need for server consolidation, reduced operational costs, and improved disaster recovery capabilities as key drivers of this growth. Additionally, virtualization software providers such as VMware, Microsoft, and Citrix continue to innovate and improve their offerings, making server virtualization an increasingly attractive option for organizations of all sizes.

2. How does Server Virtualization Works?

Imagine you have a big toy box with lots of toys inside. You want to share these toys with your friends, but you only have one toy box. So, what can you do? You can take out some toys from the toy box and put them in a smaller box, and then give that box to your friend. Now, your friend can play with those toys without having to share the big toy box with you.

In the same way, server virtualization works by taking a big computer (called a server) and dividing it into smaller parts (called virtual servers). Each virtual server acts like a separate computer, with its own operating system and applications. But all of these virtual servers share the resources of the big computer, like the processor, memory, and storage.

This means that you can run multiple virtual servers on one physical server, just like you can share toys with your friends using smaller boxes. Server virtualization is like having a magic wand that can make one computer act like many computers. It helps organizations save money by using fewer physical servers, and it makes it easier to manage all of the servers in one place. Server virtualization is a technology that allows you to run multiple virtual servers on a single physical server. Each virtual server acts like a separate computer, with its own operating system, applications, and resources. But all of these virtual servers share the resources of the physical server, like the processor, memory, and storage. So, how does this work? Server virtualization is accomplished through a software layer called a hypervisor. The hypervisor sits between the physical server hardware and the virtual servers, and it manages the allocation of resources to each virtual server. The hypervisor creates a virtualization layer that abstracts the physical hardware from the virtual servers, allowing them to run independently of each other.

There are two types of hypervisors: Type 1 and Type 2. Type 1 hypervisors run directly on the physical server hardware, while Type 2 hypervisors run on top of an existing operating system. Type 1 hypervisors are generally considered to be more efficient and secure, since they have direct access to the hardware. Once the hypervisor is installed, you can create multiple virtual servers on the physical server. Each virtual server can be configured with its own operating system, applications, and settings. You can also allocate resources to each virtual server, such as CPU, memory, and storage. Server virtualization offers many benefits, including improved server utilization, reduced hardware costs, simplified IT management, and increased flexibility and agility. It's no wonder that server virtualization has become such a popular technology in recent year

REFERENCES

1. www.techtargwt.com.
2. Amazon web services

IMPORTANCE OF BIG DATA VIRTUALIZATION

Faculty Mentor:
Dr. Suman Madan

Students Name:
Abhay Yadav (MCA- Year 1)
Amaanullah Iqbal (MCA-Year1)

1. INTRODUCTION

With the advent of technology, the world has witnessed an exponential growth of data. Every organization, no matter how big or small, collects massive amounts of data on a daily basis. However, the challenge lies in making sense of this data and using it to make informed decisions. This is where big data analytics comes into play. Big data analytics has become an essential tool for businesses to gain valuable insights and make informed decisions. However the sheer volume and complexity of data can pose significant challenges. This is where virtualization plays a crucial role. Virtualization allows for the efficient utilization of resources, scalability, and flexibility in handling big data analytics workloads. By virtualizing the infrastructure, businesses can optimize their data analysis processes, improve performance, and reduce costs. In this blog post, we will explore the importance of virtualization in big data analytics and its impact on business intelligence

2. UNDERSTANDING BIG DATA



Fig:1 Big Data Importance

In order to fully appreciate the importance of virtualization in big data analytics, it is crucial to understand what big data analytics entails. Big data analytics refers to the process of examining large and complex data sets to uncover patterns, trends, and correlations that can provide valuable insights for business decision-making.

Traditionally, analyzing big data was a laborious and time-consuming task, requiring specialized hardware and

Software resources. However, virtualization has revolutionized the way businesses approach big data analytics. By virtualizing their infrastructure, organizations can efficiently allocate computing resources, enabling them to handle the massive volumes of data associated with big data analytics.

Moreover, virtualization allows for the scalability and flexibility required in managing dynamic workloads. Businesses can easily expand or contract their computing resources based on the changing needs of their analytics processes. This agility empowers organizations to make real-time data-driven decisions with minimal delays.

In the next section, we will delve deeper into the specific benefits of virtualization in big data analytics and how it enhances business intelligence. Stay tuned!

3. WHAT IS VIRTUALIZATION?

Virtualization is a technology that enables the creation of multiple virtual machines (VMs) or virtual environments on a single physical server or computer. It allows for the separation of resources, such as processing power, memory, and storage, to be allocated to each virtual machine.

Each virtual machine operates independently, running its own operating system and applications. This means that multiple operating systems and applications can run simultaneously on a single physical server, maximizing its utilization and efficiency.

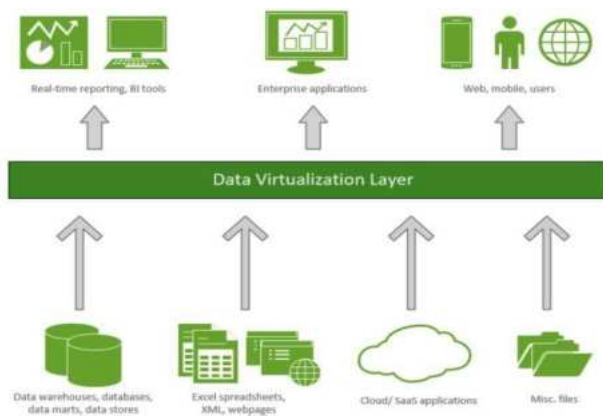


Fig:2 Data Virtualization Layer

Virtualization can be achieved through software applications called hypervisors. Hypervisors create and manage the virtual machines, ensuring that each VM operates smoothly and securely without interfering with other VMs.

The advantages of virtualization in big data analytics are immense. It allows businesses to consolidate their hardware infrastructure, reducing costs and improving resource utilization. Additionally, virtualization simplifies the deployment of big data analytics platforms, making it easier to scale and manage the infrastructure as data volumes grow.

In the following section, we will explore the benefits of virtualization in more detail, highlighting its impact on performance, scalability, and cost-effectiveness in big data analytics. Stay tuned to discover how virtualization can transform your organization's data analysis capabilities.

4. THE ROLE OF VIRTUALIZATION IN BIG DATA ANALYTICS

Virtualization plays a crucial role in big data analytics by providing a scalable and efficient infrastructure for processing and analyzing massive amounts of data. Let's delve deeper into the benefits that virtualization brings to the table.

One of the key advantages of virtualization in big data analytics is improved performance. By separating resources and running multiple virtual machines on a single physical server, organizations can effectively distribute workloads and optimize the utilization of computing power. This results in better performance and faster processing times, enabling businesses to derive valuable insights from their data in a shorter time frame.

Scalability is another area where virtualization shines. As the volume of data grows, businesses need to expand their analytics infrastructure to handle the increased workload. Virtualization simplifies this process by allowing organizations to add more virtual machines as needed, without the need for additional physical servers. This flexibility ensures that businesses can scale their analytics environment efficiently, in line with their growing data requirements.

Cost-effectiveness is a significant benefit of virtualization in big data analytics. By consolidating multiple virtual machines on a single physical server, organizations can reduce their hardware footprint, minimizing costs associated with maintenance, power consumption, and physical space requirements. Moreover, virtualization enables businesses to allocate resources dynamically, ensuring optimal utilization and avoiding wasteful overprovisioning.

In conclusion, virtualization plays a pivotal role in enhancing performance, scalability, and cost-effectiveness in big data analytics. By leveraging this technology, organizations can unlock the full potential of their data and gain a competitive edge in today's data-driven world. Stay tuned for the next blog section, where we will discuss some best practices for implementing virtualization in big data analytics environments.

5. BENEFITS OF VIRTUALIZATION IN BIG DATA ANALYTICS

Virtualization offers a multitude of benefits in the field of big data analytics. In this section, we will explore five key advantages that businesses can expect to gain by implementing virtualization in their analytics environments.



Fig:3 Data Virtualization Uses

5.1 Enhanced resource utilization: Virtualization enables organizations to make the most efficient use of their computing resources by running multiple virtual machines on a single physical server. This allows for better utilization of CPU, memory, and storage, resulting in improved performance and reduced costs.

5.2 Increased agility and flexibility: Virtualization provides businesses with the agility and flexibility to adapt to changing data requirements. By easily adding or removing virtual machines as needed, organizations can scale their analytics infrastructure in real-time, without the constraints of physical hardware.

5.3 Improved disaster recovery and high availability: Virtualization offers robust disaster recovery and high availability capabilities. By leveraging virtual machine snapshots, businesses can quickly and easily restore data in the event of a failure or disaster, minimizing downtime and ensuring business continuity.

5.4. Simplified management and infrastructure consolidation: Virtualization simplifies the management of an analytics environment by centralizing and streamlining the administration of virtual machines. This not only reduces administrative overhead but also enables infrastructure consolidation, resulting in reduced hardware costs and improved efficiency.

5.5. Enhanced security: Virtualization provides advanced security features, such as isolation and sandboxing that help protect sensitive data and applications from unauthorized access or breaches. By segregating virtual machines, businesses can isolate workloads and limit potential vulnerabilities.

In the next blog section, we will delve into the best practices for implementing virtualization in big data analytics environments, ensuring a successful and optimized deployment. Stay tuned to learn more about how to leverage virtualization effectively for your analytics needs.

6. BEST PRACTICES TO IMPLEMENT VIRTUALIZATION FOR BIG DATA ANALYTICS

Now that we have understood the numerous advantages of virtualization in big data analytics, it is essential to explore the best practices for implementing this technology successfully. By following these guidelines, businesses can ensure a seamless and optimized deployment of virtualization in their analytics environments.

First and foremost, it is crucial to assess the specific requirements and goals of your analytics infrastructure. Understanding your organization's needs will help you determine the appropriate virtualization platform and configuration that aligns with your business objectives.

Next, it is recommended to conduct thorough testing and performance evaluations before deploying virtualization in a production environment. This will allow you to identify and address any performance bottlenecks or compatibility issues, ensuring smooth operations once virtualization is implemented.

Another best practice is to establish a proper backup and recovery strategy. While virtualization offers robust disaster recovery capabilities, it is essential to have a comprehensive plan in place that includes regular backups, off-site storage, and testing the restoration process.

Additionally, organizations should prioritize security when implementing virtualization. This involves

regularly patching and updating virtualization software, segregating sensitive data and applications from other virtual machines, and implementing access control measures to prevent unauthorized access.

Lastly, ongoing monitoring and maintenance are critical to ensure the continued success of virtualization in big data analytics. Regularly monitoring performance, capacity, and resource utilization will help identify and resolve any issues promptly, ensuring optimal performance.

By adhering to these best practices, businesses can harness the full potential of virtualization in big data analytics and gain a competitive edge in the market. In the next blog section, we will explore some real-world use cases where virtualization has revolutionized the way organizations analyze and derive insights from their vast data sets. Stay tuned to learn more about the practical applications of this technology.



7. CONCLUSION

In conclusion, virtualization has become a critical technology for big data analytics, and implementing it effectively requires careful planning and adherence to best practices. By assessing the specific requirements of your analytics infrastructure, conducting thorough testing, establishing a backup and recovery strategy, prioritizing security, and regularly monitoring and maintaining virtualization, businesses can optimize their analytics environments and unlock the full potential of their data.

In the next blog section, we will delve into real-world use cases that highlight the practical applications of virtualization in big data analytics. These examples will demonstrate how organizations have leveraged virtualization to revolutionize their data analysis processes, drive innovation, and gain a competitive advantage in their industries. Stay tuned for these inspiring case studies that showcase the transformative power of virtualization in big data analytics.

REFERENCES

1. <https://simplified.com/>
2. <https://www.eckerson.com/articles/data-virtualization-in-business-intelligence-and-analytics>
3. <https://aws.amazon.com/what-is/virtualization/>
4. <https://datavirtuality.com/en/blog/data-virtualization-the-complete-overview/>
5. <https://hevodata.com/learn/data-virtualization-tools/>
6. <https://tdwi.org/articles/2022/04/04/diq-all-data-virtualization-use-cases.aspx>
7. <https://www.thedigitalwhale.com/blog-the-exponential-growth-of-data-and-its-impact-on-businesses/>

FULL VIRTUALIZATION AND ITS TRANSFORMATIVE APPLICATIONS

Faculty Mentor:
Dr. Aakansha Upadhyaya

Students Name:
Kanishk Arya (MCA- Year 1)
Nitin Singh Bisht (MCA- Year 1)

INTRODUCTION



Fig 1: Virtualization

What is Full Virtualization? Virtualization utilizes strategies that can make instances of a virtual environment. So, we can say that Full virtualization requires that every salient feature of the hardware be reflected into one of several virtual machines.

In layman's terms, virtualization is the process of creating a virtual computing environment that is simulated at the same level of abstraction as the physical environment, allowing the simulated environment to take advantage of the capabilities of a physical machine. It will be. In simple terms, it's basically like having a virtual computer inside your real computer, both of which do the same thing.

Co-working or shared office spaces are a popular real-world illustration of virtualization and resource sharing that most people can comprehend. Like virtualization in the IT industry, these areas are effectively virtualized work environments.

Advantages of Virtualization-

- Reduce costs
- Increase Scalability and flexibility
- Improve performance

- Support Collaboration
- Cloud Computing Enablement
- Simplified Management

Applications of virtualization

There are a wide range of real-world uses for virtualization in different industries. Several important uses of virtualization across several industries are as follows:

I. Transforming Science Education: The Impact of Virtual Labs and E-Learning

Virtual labs and e-learning have become very popular in science education these days. And we can say that main force behind this was COVID-19. Due to COVID the education infrastructure was completely shut and imparting education was became very tough. In order to address this circumstance, many remote learning strategies were implemented, such as virtual teaching, live streaming, online courses, and simulated labs.

That being said, it should not be confused with e-learning, which is conducted online and includes training courses, educational games, and simulations. E-learning is typically self-paced and does not call for in-person interaction with instructors or other students, whereas virtualization of education is a type of online learning that does. The majority of virtual education is synchronous and real-time, requiring students to attend at predetermined times

Now COVID times are gone but it has revolutionized the education system. You can say that virtual teaching is emerging as effective platform for science and engineering education and it has increased the spread of educational resources globally. It has become an elixir for the countries where the numbers of students

are more as compared to well-experienced teachers or highly equipped educational institutes and laboratories. Now these virtual classes alone cannot provide adequate skills or knowledge regarding how to implement the things in real. So addition of virtual labs added a new dimension to this internet based learning. Virtual labs are available 24/7 and provide students with an easy-to-use, simulated version of real-world labs that can teach them real-world science skills. In simple words we can say that these things have drastically reduced the number physical resources by increasing the reach of limited resources. [1]

II. Impact of Virtualization on revolutionizing Healthcare

Healthcare virtualization is the use of virtualization technology in various aspects of the healthcare industry, from infrastructure management to patient care. The future of healthcare will likely be a hybrid model where patients receive a combination of virtual and in-person care. This shift to hybrid medicine could alleviate some of the biggest problems in India's healthcare system. For example, a “digital front door,” a digital exit point that takes patients to their next point of care, is an advantage of virtual care.

Virtual health plans may include an AI platform or telehealth visits as a digital front door to provide members with the various benefits that the plan can offer. Virtualization change the methodology in the traditional teaching method. Medical students now have the opportunity to immerse themselves in 3D anatomical models, gaining a deeper understanding of human anatomy and pathology [2].

This hands-on approach increases the knowledge to more comprehensive learning experience, and tends towards the preparing future healthcare professionals for the challenges of real-world healthcare services.

Here are several ways in which virtualization is applied in healthcare:

Electronic Health Records (EHR): Virtualization helps in managing and securing electronic health records efficiently, ensuring accessibility while maintaining data integrity and confidentiality. **Telemedicine and Virtual Care:** Virtual Visits: Virtualization technologies play a key role in telemedicine, enabling virtual visits between patients

and healthcare providers. These include video consultations, remote monitoring, and virtual care platforms that improve access to health services.

Digital Health Platforms: Virtualization supports the infrastructure of digital health platforms, allowing patients to interact with healthcare providers, access medical records, and manage their health remotely [3].

Simulation and Training: Virtualization is used to develop medical training simulators that provide realistic scenarios for medical professionals to practice different procedures and improve their skills in a safe environment.

Continuing Education: Health care professionals can access virtual training modules and educational resources to facilitate continuous learning and professional development.

III. Virtualization in Automation: Streamlining Systems Deployment and Enhancing Testing Through Simulated Environments

Virtualization provides a way to create virtual environments, replicate hardware setups, and streamline the deployment and management of automation systems.

Here are several ways in which virtualization is applied in automation:

Industrial Control Systems (ICS): Simulated Environments: Virtualization allows for the creation of simulated environments to test and validate industrial control systems before deploying them in the physical environment. This allows operators and technicians to familiarize themselves with the user interface and practice operations without affecting the actual automation system.

Human-Machine Interface (HMI): HMI Virtualization: Virtualization can be used to create virtual HMIs for testing and training purposes. This allows operators and technicians to familiarize themselves with the user interface and practice operations without affecting the actual automation system.

Robot Simulation and Automation Testing: Virtualization used in robot simulation environments allows engineers to model robot movements and

interactions in virtual space before introducing robots into real manufacturing or logistics environments [4].

IV. Virtualization in Architecture and Construction: Transforming Project Management and Design

Paper documents are still often used in construction operations, which haven't altered much in recent decades. Furthermore, spreadsheets are still used in far too many operations. Spreadsheets are used in 45% of project management workflows, according to the 2021 JB Knowledge ConTech Survey. Excel is an extremely useful tool, but given the intricacy and dynamic nature of major building projects, it is also prone to error [5].

In architecture and construction, virtualization refers to the process of using computer software to generate digital representations of real-world things or settings. The construction process can then be simulated, possible issues can be found, and designs can be optimized using these models. In addition to providing customers and stakeholders with immersive experiences, virtualization may help them see a project through to completion.

Virtual reality (VR): By recreating the user's physical presence in a virtual environment, VR produces an immersive experience. With virtual reality (VR), one may explore various design possibilities, examine construction details, and navigate a virtual building.

Improved efficiency and productivity: Virtualization can improve efficiency and productivity by automating tasks, streamlining workflows, and reducing the need for physical prototypes.

Enhanced decision-making: Virtualization can provide decision-makers with more information and insights, allowing them to make better decisions about design, construction, and operation.

Conclusion:

The exploration of full virtualization and its diverse applications across various industries underscores the transformative power of this technology.

Virtualization, as a strategy for creating instances of virtual environments, has become an integral aspect of modern computing, revolutionizing the way resources are utilized and managed.

The advantages of virtualization, including cost reduction, increased scalability, flexibility, improved performance, support for collaboration, cloud computing enablement, and simplified management, showcase its versatility and adaptability in meeting the dynamic needs of today's technological landscape.

All these uses of virtualization ensure that virtualization is a promising technology that can transform the way we use different resources in our daily lives.

REFERENCES

1. <https://link.springer.com/article/10.1007/s42485-020-00038-7>
2. <https://www.sciencedirect.com/science/article/pii/S1110016819300511>
3. <https://constructible.trimble.com/construction-industry/what-is-virtual-design-and-construction>
4. <https://services.caddetails.com/blog/virtual-architecture-the-future-of-vr-and-the-construction-industry>
5. <https://www.mckinsey.com/industries/health-care/our-insights/virtual-health-a-look-at-the-next-frontier-of-care-delivery>
6. <https://www.wheelhouse.com/resources/top-10-virtualization-trends-a7090>
7. <https://www.process-worldwide.com/best-practices-for-virtualization-in-process-automation-a-426557/>

DATA VIRTUALIZATION AND ETL

Faculty Mentor:

Dr. Isha Singh

Students Name:

Anshu Kumar (MCA- Year 1)

Anshika Ghosh (MCA- Year 1)

Abhay Gupta (MCA- Year 1)

INTRODUCTION

Data virtualization and ETL offer different capabilities to access, integrate, and deliver data. Although they can offer competing approaches for some specific scenarios, they are both useful additions to the data integration toolset of any organization. In general, Data Virtualization is more agile, flexible, versatile, and cost-efficient than ETL. It has become a hot topic recently as a promise for better, faster more flexible and improved operations in the Business Intelligence world. On the other side As ETL technology evolved, both data types and data sources increased exponentially. Cloud technology emerged to create vast databases (also called data sinks). ETL tools have also become more sophisticated and can work with modern data sinks. They can convert data from legacy data formats to modern data formats.

People generate a vast volume of data daily due to the rapid expansion of information technology. As a result, dealing with such massive amounts of data and uncovering hidden patterns and insights to make useful decisions has become a problem for all companies. To address this issue, a data warehouse is necessary, which stores large amounts of data from many sources, and this sort of data is referred to as Big Data[1]. The Extract, Transform, and Load (ETL) method is the most typical method for collecting data from many sources and integrating it into a single source of information, i.e., a data warehouse[2]. During the ETL process, data is extracted from such a resource as a relational database, XML files, flat files, or spreadsheets, transformed to data warehouse standards, and then loaded into the warehouse[3]. Figure 1 depicts the fundamental architecture of the ETL process. Although ETL is required for data warehousing and analytics, not all ETL software packages are created equal. The best ETL tool may vary

depending on the circumstances and use cases[4]. This document compares the most common ETL 2022 tools [5]. In the modern era, this 21st century is the world of data also we can say that today is data world without it we cannot survive. In small word we can say that data is oxygen for everyone so that why most of the company work on how to acquire more data to the users.

This article includes an overview of Data virtualization in Section I, Section II contains the three principles of data virtualization, Section III contains Data virtualization Software, Section IV contains ETL-Extract Transform Load also ETL steps, ETL examples and Benefits of well engineered ETL. This article gives you the better understanding of Data virtualization and ETL with their working architecture.

DATA VIRTUALIZATION

Data virtualization is an umbrella term used to describe any approach to data management that allows an application locate to retrieve and manipulate data without requiring technical details about the data, such as how it is formatted or where it is physically located.[6]

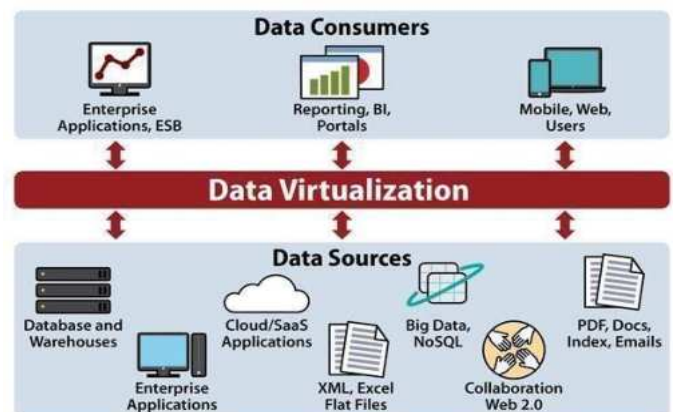


Fig 1: Data virtualization combines data from disparate

Example: - The photo's file path. Once it has been uploaded to Facebook, however, you can retrieve the Photo without having to know its new file path. In fact, you will have absolutely no idea where Facebook is Storing your photo because Facebook software has an abstraction layer that hides that technical information. This abstraction layer is what is meant by some vendors when they use the term data virtualization.

In the modern era, enterprise data comes in many forms and is stored in many locations. There is both structured and unstructured data, including rows and columns of data in a traditional database, and data in formats like logs, email, and social media content. Big Data in its many forms is stored in databases, log files, CRM, SaaS, and other apps.

I. Data virtualization - Three Cs Principles

- 1.1 **Connect-** Connect to any data sources (e.g. files, database, API's etc.)
- 1.2 **Combine-** The motive is to gather data from Multiple sources and combine them to fulfil a business requirement, this layer is meant to Cater exactly to that purpose. This layer defines the data transformation and combination to Meet business requirements.
- 1.3 **Consume-** Finally real-time data available to the data consuming platform.

II. How data virtualization works:



Fig 2: Data virtualization architecture

2.1 **Data sources-** This layer is responsible for accessing the information scattered across multiple source systems, with the help of connectors and communication protocols. Data virtualization platforms can link to different data sources including

- SQL and NoSQL databases like MySQL, Oracle, and MongoDB;
- CRMs and ERPs;
- Cloud data warehouses like Amazon Redshift or Google Big Query;
- Data lakes and enterprise data warehouses;
- Streaming sources like IoT and IoMT devices;
- SaaS(Software-as-a-Service) applications like Salesforce;
- Social media platforms and websites;
- Spreadsheets and flat files like CSV, JSON, and XML;
- Big data systems like Hadoop, and many more[7].

2.2 **Abstraction layer-** The cornerstone of the whole virtualization framework is the abstraction (sometimes referred to as *virtual* or *semantic*) layer that acts as the bridge between all data sources on one side and all business users on the other. This tier itself doesn't store any data: It only contains logical views and metadata needed to access the sources.

2.3 **Consuming Layer-** Another tier of the data virtualization architecture provides a single point of access to data kept in the underlying sources. The delivery of abstracted data views happens through various protocols and connectors depending on the type of the consumer. They may communicate with the virtual layer via SQL and all sorts of APIs, including access standards like JDBC and ODBC, REST and SOAP APIs, a most data virtualization software enables access for a wide range of business users, tools, and applications including such popular solutions as Tableau, Cognos, and Power BI.

III. Data Virtualization Software



Fig 3: Data virtualization Software.

3.1 Abstraction- It creates an abstraction layer that hides away the technical aspects, such as storage technology, system language, APIs, storage structure, and location, of the data[8].

3.2 Virtualized Data access- It allows you to break down data stores by establishing a logical data access point to disparate sources.

3.3 Transformation- It allows users to apply transformation logic on the presentation layer, thus improving the overall quality of data. It is the process of converting, cleansing, and structuring data into a usable format that can be analysed to support decision making processes, and to propel the growth of an organisation.

3.4 Data Federation- It is a software process that allows multiple databases to function as one. This virtual database takes data from a range of sources and converts them all to a common model. It works in unison with data federation software to provide integrated views of data sources from disparate databases.

3.5 Data Delivery- It enables you to publish datasets (requested by users or generated through client application) as data services or business data views.

3.6 Data Unification- It is the process of ingesting data from various Operational systems and combining them into a single source by performing transformations, schema integrations.

IV. ETL-EXTRACT TRANSFORM LOAD



Fig 4: ETL for Media and Entertainment Industry

ETL is the process of combining data from multiple sources into a large, central repository called a data warehouse. ETL uses a set of business rules to clean and organise raw data and prepare it for storage, data analytics, and machine learning (ML). ETL has ruled industry over 1 decade.[9]

1. ETL Process- STEPS

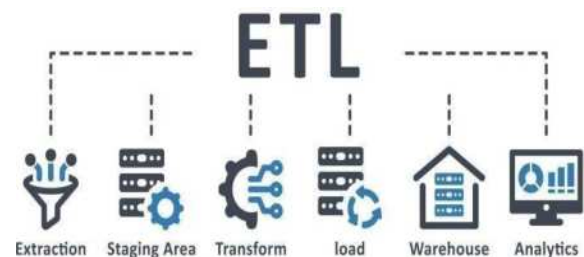


Fig 5:ETL Process-Steps

- ❑ **Extract-** Retrieves raw data from an unstructured data pool and migrates it into a temporary, staging data repository[10].
- ❑ **Clean-** Cleans data extracted from an unstructured data pool, Ensuring the quality of the data prior to transformation.
- ❑ **Transform-** Structures and converts the data to match the correct target source.
- ❑ **Load-** Loads the structured data into a data warehouse so It can be properly analyzed and used.
- ❑ **Analyse-** Big data analysis is processed within the warehouse, enabling the business to gain insight from the Correctly configured data.

Each step is performed sequentially. However, the exact nature of each step – which format is required for the target database – depends on the enterprise’s specific needs and requirements.

2. ETL Examples

- ❑ Connecting to a single or multiple operational data sources, including an ERP or CRM database[11].
- ❑ Extracting batches of XML, JSON, and flat files (or other formats) into rows according to one or more source system’s tables, based on certain criteria.
- ❑ Copying the data that was extracted to a staging area where data values can be standardised and writing the process outputs to log files for debugging.
- ❑ Beginning transformations on the staged data, which can range from being performed in-memory or in temporary tables on the disk.

3. Benefits of well Engineered ETL Process

- ❑ **Information clarity-** During ETL transformations, data is cleaned and joined across sources before it is saved in the database, where you can then analyze it. These operations allow to work with clear information and clarify raw data.
- ❑ **Information Completeness-** A well-designed ETL pipeline unites all business sources in a single place (the destination data warehouse/database). All of the information is complete, so there are no missing puzzle pieces.
- ❑ **Information Quality-** It validates data at extraction correct/discard data at transformation ensures that the quality of data is always controlled before it is analyzed, thus increasing trust in the analysis and helps to use data for business intelligence, data analytics, machine learning algorithms or other data science projects.
- ❑ **Information Velocity-** ETL processes can be designed to trigger the entire ETL pipeline whenever new data arises in the sources or when existing data is changed.
- ❑ **Novel Business Insights-** The entire ETL process brings structure to your company’s information. This allows you to spend more time analyzing novel questions and acquiring new insights, rather than trying to perform procedures to get valuable data at each stage.

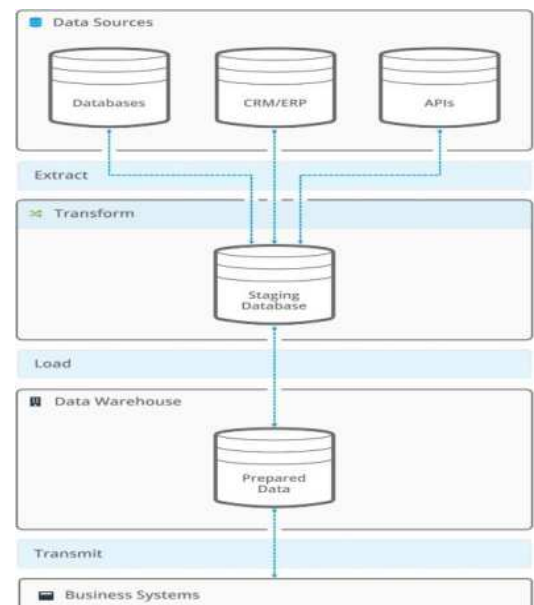


Fig 6: ETL Process Benefits

REFERENCES

1. F. d. A. Vilela and R. R. Ciferri, "A survey of real-time etl process applied to data warehousing environments," in ITNG 2022 19th International Conference on Information Technology- NewGenerations. Springer, 2022, pp. 91–96.
2. A. Dhaouadi, K. Bousselmi, M. M. Gammoudi,
3. S. Monnet, and S. Hammoudi, "Data warehousing process modeling from classical . approaches to new trends: Main features and comparisons," Data, vol. 7,no. 8, p. 113, 2022.
4. J. Y. A. Moura and B. Z. Cadessaib, "Effort estimation method for extract transfer load (etl) big data projects," in 2022 2nd International Conference on Information Technology and Education (ICIT&E). IEEE, 2022, pp. 160–167.
5. E. Mehmood and T. Anees, "Distributed real-time etl architecture for unstructured big data," Knowledge and Information Systems, vol. 64, no. 12, pp. 3419–3445, 2022.
6. B. Oliveira, M. Leite, O. Oliveira, and O. Belo,"A service-oriented fssrame-' work for etl implementation," in EPIA Conference on Artificial Intelligence. Springer, 2022, pp. 636–647.

A GLIMPSE OF PARA-VIRTUALIZATION

Faculty Mentor:
Dr. Deepti Sharma

Students Name:
Akansha Bhalla (MCA-Sem1)
Malvika Gaur (MCA -Sem1)

INTRODUCTION

Software instructions from the guest operating system running inside a virtual machine can use "**hyper-calls**" to communicate directly with the hypervisor in a type of virtualization known as para- virtualization. This offers an interface that is strikingly comparable to native host hardware software.

When instructions are incompatible with full virtualization or when faster access to the underlying hardware is needed for performance-related reasons, para-virtualization offers the most advantages. Para- virtualization can offer the performance of native code for timing-critical functions together with some virtualization advantages, such hardware sharing between several operating systems.

The goal of the updated interface is to minimize the amount of time the guest spends executing tasks that are significantly more challenging to complete in a virtualized environment than in a non-virtualized one. The tasks that would otherwise be performed in the virtual domain (where execution performance is inferior) can now be requested and acknowledged by the guest(s) and host, thanks to the para-virtualization, which offers specifically specified "**hooks**".

1. Virtual Machine Monitor (VMM)

A para-virtualized platform that is successful may lower the overall performance degradation of machine execution inside the virtual guest and simplify the virtual machine monitor (VMM) by moving key task execution from the virtual domain to the host domain. Para-virtualization necessitates explicitly porting the guest operating system for the para-API; a

traditional OS distribution that is not para- virtualization-aware cannot run on top of a para virtualization VMM. Even if the operating system cannot be updated, components that enable many of the important performance.

Diagnose Code

The Xen Windows GPLPV project provides a set of **GPL licensed** para-virtualization-aware device drivers designed to be installed into a Microsoft Windows virtual guest operating on the **Xen hypervisor**. Such applications are typically accessed via the para virtual machine interface environment. This ensures that multiple encryption algorithm models function in the same mode, allowing for smooth integration inside the para virtual framework. The phrase "para-virtualization" refers to a new name for an old concept. Since 1972 (and earlier as CP-67), IBM's VM operating system has provided such a feature. This is known as a "**Diagnose Code**" in the VM world since it employs an instruction code that is generally only utilized by hardware maintenance software and is hence undefined.

2. Full virtualization

Para-virtualization tries to address the problems identified with full virtualization. What sets para- virtualization apart from full virtualization is that it allows you to make changes to the host operating system. On the other hand, in full virtualization the host operating system is not aware that it is virtualized. In para- virtualization the unmodified operating system is

aware that it is being virtualised and sensitive operating system calls are captured and converted using binary translation.

3. QNX Hypervisor

Para-virtualization can help embedded car structures in sharing get admission to to peripherals. This can encompass the infotainment or device display, audio interfaces, contact interfaces, verbal exchange interfaces, and garage gadgets. The para-virtualization drivers will transfer get admission to among one of a kind running structures that require the usage of a peripheral. An car hypervisor along with QNX Hypervisor can aid get admission to to peripherals in numerous ways. It can emulate hardware gadgets in a totally virtualized way. It can by skip get admission to via directly, giving special control, with the digital device simply understanding that a reminiscence variety has been reserved. In para virtualization mode, the hypervisor can keep away from the limitations of complete virtualization. This may be carried out through the Organization for the Advancement of Structured Information Standards (OASIS) **VIRTIO** specification.

For example, the QNX Hypervisor can allow graphics, video, audio, and input/contact to be shared among QNX and Android running structures, presenting overall performance benefits. In an car environment, minimizing latency may be mission-critical. Safety capabilities need to be done without delay, and crucial infotainment capabilities along with navigation can be advanced through real-time execution.

In the early days, implementing platform Virtualization called for making adjustments to the operating systems running on the virtual machines. Para- virtualization is the term used to describe a system that requires modifications or active participation from a guest operating system. Systems with full virtualization, in contrast, do not require any modifications or active participation from the guest operating systems.

4. Benefits of Para- virtualization

It also presents numerous issues. In order to utilize a specific operating system in a paravirtualized environment, it is essential to verify whether the virtualization platform provides drivers for the guest operating system. In other words, the platform has the capability to handle **Linux kernel version 2.4** but it may not be compatible with **kernel version 2.6**. In addition to that, company 'A' has developed a version of the drivers, while open source group 'B' has also created their own set of drivers. Administrators can often find themselves in a state of confusion as they navigate the multitude of driver options, trying to discern the benefits and drawbacks associated with each implementation. Furthermore, the variations can be exceedingly intricate. The process of transferring data between hardware devices can vary significantly across different operating systems. Para virtualization can become exceedingly intricate, making it highly susceptible to errors. summary, portability and compatibility are crucial elements for enhancing productivity, accessibility, and connectivity in the modern digital age. Moreover, due to the sharing of resources in para virtualization, it lacks the portability that is commonly found in other virtualization architectures. It can be quite a challenge for individuals without training to modify an operating system in order to work together with a hypervisor. In conclusion, it should be noted that not every guest operating system is compatible with para virtualization. Consequently, tackling issues within these environments can prove to be quite challenging. The challenges of para- virtualization primarily lie in the technical domain, necessitating a skilled approach for deployment, management, and adeptly navigating intricate compatibility concerns. Technical requirements refer to the specifications or criteria that must be met in order to successfully develop or implement a technological solution. These requirements outline the necessary capabilities, functionalities, and performance standards of the solution. They serve as a guide for developers, engineers, and other stakeholders involved in the project, ensuring that the final product meets

the desired objectives and functions effectively. By clearly defining the technical requirements, any potential issues or limitations can be identified and addressed early on in the development process, resulting in a more efficient and successful implementation of the technology.

The current shortage in the market coincides with the advanced technical skills required for deploying, managing, operating, and monitoring Para virtualized guest OS. A Para virtualized OS may not perform optimally and could pose potential risks if not configured correctly. The concept of portability and compatibility refers to the ability of a system or device to be easily and seamlessly transferred or used across different platforms, devices, or environments. This aspect is essential in today's technology-driven world as it enables users to access and utilize their information and resources anywhere and at any time without constraints or limitations. Portability allows for flexibility and convenience, as users can carry their devices with them and switch between various systems without losing functionality. Compatibility ensures that different devices, software, and operating systems can interact and communicate effectively, resulting in a smooth and efficient user experience.

5. Types of para-virtualization

5.1 Cloud computing: Google, IBM, Amazon, Microsoft and different pinnacle companies all leverage para-virtualization to higher maximize their cloud computing environments, power efficiencies, and enhance scalability and overall performance. It is way to this version of virtualization that loads of hundreds of thousands of customers can seamlessly get admission to the cloud and its extensive variety of offerings and features.

5.2 IT and cellular: Large IT and cellular enterprise groups which have excessive workloads and high- overall performance needs are regarded to installation para virtualization environments to enhance operations, reliability, and resilience.

5.3 High-overall performance computing industries: From protection to pharmaceutical and different industries, agencies that want to perform high- overall performance computing environments typically check out para virtualization to lessen the overhead of the hypervisor of their digital infrastructures.

5.4 Legacy structures: With para-virtualization industries and groups can run legacy structures on current hardware while not having to alter the legacy structures themselves.

5.5 Research and improvement: Para-virtualization may be utilized in studies and improvement environments to test with new OS and programs while not having to alter the underlying hardware.

5.6 Security and privacy: Para-virtualization is likewise used to decorate cyber security and privacy, specifically in troubles associated with catastrophe healing solutions.

5.7 Software improvement: Para-virtualization presents DevSecOps groups withideal virtual environments to build, test, installation and screen their software,API, code, or apps.

6. Conclusion

Last but not the least there are numerous motives why pinnacle cloud carriers and main groups select to apply para virtualization. While para-virtualization isn't always for everybody or each use case, it's far a effective virtualization era which could provide some of benefits, consisting of stepped forward performance, efficiency, scalability, security, and decreased overhead. Ultimately whether or not user select to go together with para-virtualization, complete virtualization ,or partial virtualization will rely upon organization's precise needs, budget, and infrastructure.

CLOUD COMPUTING AND SECURITY ISSUES

Faculty Mentor:

Dr. Archana B Saxena

Students Name:

Aryan Malik (MCA-Sem1)

Vaibhav Singh(MCA -Sem1)

INTRODUCTION

Cloud computing has emerged as a recreation-changer inside the international of technology, providing convenience, scalability, and fee-effectiveness. But, as with all technological advancement, cloud computing also brings forth quite a number protection issues. This newsletter explores the various protection problems associated with cloud computing and highlights the measures that may be taken to mitigate those risks.

Understanding Cloud Computing and Its Benefits

Cloud computing refers to the delivery of computing offerings, which includes storage, servers, databases, software program, and networking, over the internet. It removes the want for physical infrastructure and permits users to get admission to their information and programs from everywhere, at any time. The benefits of cloud computing are to serve together with elevated flexibility, reduced charges, progressed collaboration, and greater scalability.

However, the very nature of cloud computing raises issues approximately information safety. let's delve into a number of the important thing protection demanding situations confronted inside the cloud computing landscape.

1. CLOUD COMPUTING MODELS

Cloud hosting deployment fashions are labeled by way of the proprietorship, size and get admission to. It tells approximately the character of the cloud. maximum of the agencies are willing to put in force cloud since it reduces the expenditure and controls price of operation

Public Cloud

it is a sort of cloud hosting in which the cloud offerings are brought over a network that is open for public utilization.

This model is honestly genuine representation of cloud hosting. on this the cloud version provider company presents services and infrastructure to numerous clients. Customers do now not have any manage over the vicinity of the infrastructure. There can be little or no or no difference among public and personal clouds structural design except the level of safety which might be supplied for various services given to the public cloud subscribers by means of the cloud website hosting providers. Public cloud is perfect for business which require managing load because o f the reducing capital overheads and operational price the general public cloud model is affordable. dealers might also provide the free service or license policy like pay in step with person. The price is shared by using all the users in public cloud. It income the clients by way of reaching economies of scale. Public cloud facilities may be to be had free of charge an e.g. of a public cloud is Google.

Private Cloud

It's also known as inner cloud. This platform for cloud computing is applied on cloud-primarily based cozy environment and it's miles safeguarded via a firewall that's governed by using the IT branch that belongs to a selected corporate. personal cloud allows best the legal users and offers the business enterprise extra manage over their facts. The physical computer systems may be hosted internally or externally they offer the sources from a awesome pool to the private cloud services. Organizations having unanticipated or dynamic wishes, assignments which might be vital management demands and uptime necessities are higher applicable to adopt personal cloud. In non-public cloud there may be no need for extra security policies and bandwidth barriers that can be found in a public cloud environment. customers and Cloud carriers have control of the infrastructure and stepped forward security, considering user's get right of entry to and the networks used are restrained. one of the great examples is Eucalyptus systems .

Hybrid Cloud

It is a kind of cloud computing, that's incorporated. it can constitute an association of or extra cloud servers of personal, public or community cloud this is sure together however remain character entities. Hybrid clouds are capable of crossing isolation and overcoming barriers with the aid of the company; consequently, it cannot be truly categorized into public, private or community cloud. It permits the consumer to increase the potential as well as the capability via assimilation, aggregation and customization with some other cloud bundle / provider. In a hybrid cloud, the resources are managed either in-house or by using external vendors. it's miles an adaptation among platforms wherein the workload exchanges between the non-public cloud and the public cloud as consistent with the desires and demand of agency. sources that are non-important like development and test workloads may be housed within the public cloud that belongs to a third- party company. whilst the workloads which can be vital or sensitive have to be housed internally. groups may use the hybrid cloud version for processing large records. Hybrid cloud website hosting has capabilities like scalability, flexibility and security.

Community Cloud

It's far a type of cloud web hosting in which the setup is mutually shared among loads of corporations which belong to a specific network like banks and buying and selling firms. it is a multi-tenant setup this is shared among many agencies that belong to a group which has comparable computing apprehensions. Theses network individuals generally proportion comparable performance and security worries. the principle goal of the communities is to obtain enterprise associated goals. community cloud can be controlled internally or may be controlled via third birthday celebration companies and hosted externally or internally. The value is shared via precise companies within the community, consequently, network cloud has cost saving ability. businesses have found out that cloud hosting has numerous capability. To be the best one should select the right type of cloud website hosting consequently, one want to realize the business and examine his/her demands. once the suitable kind of cloud website hosting is chosen, it is easy to achieve commercial enterprise associated desires without difficulty.

: Cloud computing service models

: Software as a service (SaaS)

Software as a provider (SaaS) is developing rapidly. SaaS makes uses the internet to provide packages which can be controlled by way of a 3rd-party seller and whose interface is accessed at the consumer facet. SaaS packages can be run from a web browser without the need to download or installation, but those require plugins. The cloud issuer gives the patron with the capability to set up an application on a cloud infrastructure.

: Infrastructure as a provider (IaaS)

Infrastructure as a provider, are used for tracking, and dealing with far flung datacenter infrastructures, which include compute (virtualized or bare steel), garage, users should purchase IaaS primarily based on consumption, similar to different software billing. IaaS users have the duty to be in rate programs, statistics, runtime and middleware.. vendors can nevertheless manage virtualization, servers, garage, and networking. IaaS vendors provide databases, messaging queues, and other offerings above the virtualization layer as well.

: Platform as a carrier (PaaS)

Platform as a provider (PaaS) is a form of cloud computing offerings that offers a platform that permits clients to broaden, run, and manage programs without the hassle of building and maintaining the infrastructure. One need not be troubled approximately decrease stage elements of Infrastructure, community Topology, security all this is accomplished for you by way of the Cloud service provider. With this generation, 0.33-birthday celebration providers can manipulate OS, virtualization, and the PaaS software itself. builders manipulate the packages.

3. Protection issues

Information Breaches: one of the number one concerns in cloud computing is the hazard of records breaches. Storing touchy information on far off servers increases concerns approximately unauthorized get right of entry to, statistics theft, and privateness breaches. Cloud carrier vendors must implement robust safety features to protect data from unauthorized get admission.

This consists of encryption, strong get right of entry to controls, and everyday protection audits. Facts Loss: whilst cloud companies typically have backup systems in location, facts loss can still occur because of hardware disasters, natural disasters, or human mistakes. organizations should ensure that their data is often sponsored up and enforce disaster recuperation plans to limit the effect of capability statistics loss. additionally, it is vital to pick out cloud carriers that offer reliable backup and recuperation answers.

Safety challenges and Mitigation strategies

Insider Threats:

Cloud computing includes more than one parties, together with cloud provider companies, third-celebration carriers, and customers. Insider threats can rise up from malicious actions via personnel or contractors with get right of entry to to touchy facts. Strict access controls, normal audits, and employee education are essential to mitigate this hazard. groups should also establish clean rules and techniques to address insider threats and screen user activities to come across any suspicious conduct.

Compliance and criminal problems:

Agencies running in regulated industries should follow particular information protection and privateness policies. whilst records is stored inside the cloud, ensuring compliance becomes extra complicated. businesses should cautiously pick out cloud providers that adhere to applicable policies and provide transparency concerning facts coping with practices. ordinary audits and exams can assist make sure compliance and discover any capability gaps.

Shared Infrastructure: Cloud computing entails the sharing of resources and infrastructure amongst multiple customers. This shared environment can introduce protection risks, inclusive of unauthorized get admission to to information or the ability for cross-tenant assaults. Cloud companies must enforce robust isolation mechanisms to make sure that every user's facts stays separate and relaxed. users should additionally consider enforcing extra safety features, along with encryption and get right of entry to controls, to protect their information inside the shared Surroundings.

Conclusion

Cloud computing offers vast advantages, however it additionally affords particular security challenges. groups should be proactive in addressing these issues to protect their facts and preserve client consider. via implementing sturdy security features, inclusive of encryption, get entry to controls, everyday audits, and cautious dealer choice, groups can navigate the cloud computing panorama with self-assurance and obtain the rewards of this transformative technology. As cloud computing keeps to adapt, it's far crucial for companies to live updated at the brand new safety practices and adapt their techniques for that reason to make sure the safety of their valuable information.

EMULATION AND VIRTUALIZATION IN CLOUD COMPUTING

Faculty Mentor:

Ms. Ankita Chopra

Students Name:

Jeevesh Kumar (MCA-1st year)

Ghratika Gupta (MCA -1st year)

Priya Joshi (MCA -1st year)

1. What is Virtualization?

Virtualization is a technology that allows multiple operating systems (OS) or applications to run on a single physical machine. It creates a virtual (rather than actual) version of a resource, such as a server, storage device, network, or even an operating system. This is achieved by using a layer of software called a hypervisor or virtual machine monitor (VMM).

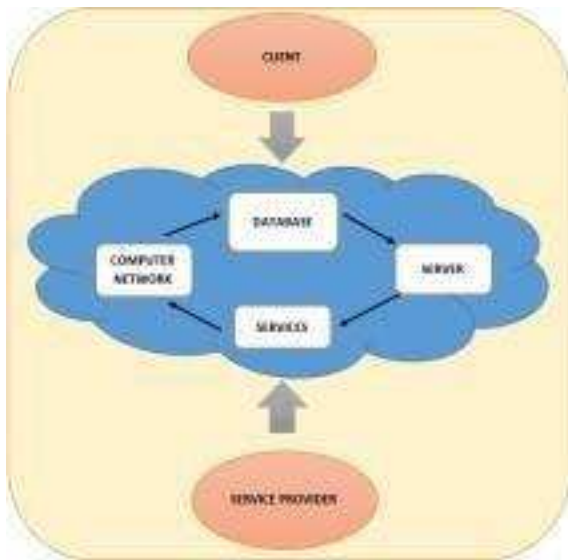


Figure 1: Relationship between Service Provider and End User

Server-based system architecture (SBSA) features are the same as the host system with a hypervisor or

Virtual Machine Monitor (VMM). It is much more efficient and has various stages of implementation.

1.1 Benefits of Virtualization-

- Cost savings through efficient resource utilization.
- Flexibility with dynamic allocation for changing workloads.
- Isolation ensures enhanced security and system stability.
- Energy efficiency by reducing the need for physical servers.
- Scalability, easily adding or removing virtual resources as needed.
- Improved hardware lifespan through optimized workload distribution.

1.2 Drawbacks of Virtualization:

- Operating costs may reduce performance in some applications.
- Complexity may pose challenges for small-scale deployments.
- Dependency on host system; failure affects all virtual machines.
- Initial setup costs can be high for small businesses.

1.3 Characteristics of Virtualization

- **Resource Pooling:** Physical resources are pooled and dynamically allocated to virtual machines as needed.

- **Hardware Independence:**

Virtualization enables the running of multiple operating systems on the same physical hardware.

- **Sharing:** Virtualization allows the creation of a separate computing environment within the same host.

- **Centralized Management:** Virtualization provides centralized interfaces for managing multiple virtual machines from a single point.

- **Isolation:** Virtual machines operate independently, providing isolation and preventing interference between them.

- **Encapsulation:** Entire virtual machines and their configurations are encapsulated in files, facilitating easy deployment and migration.

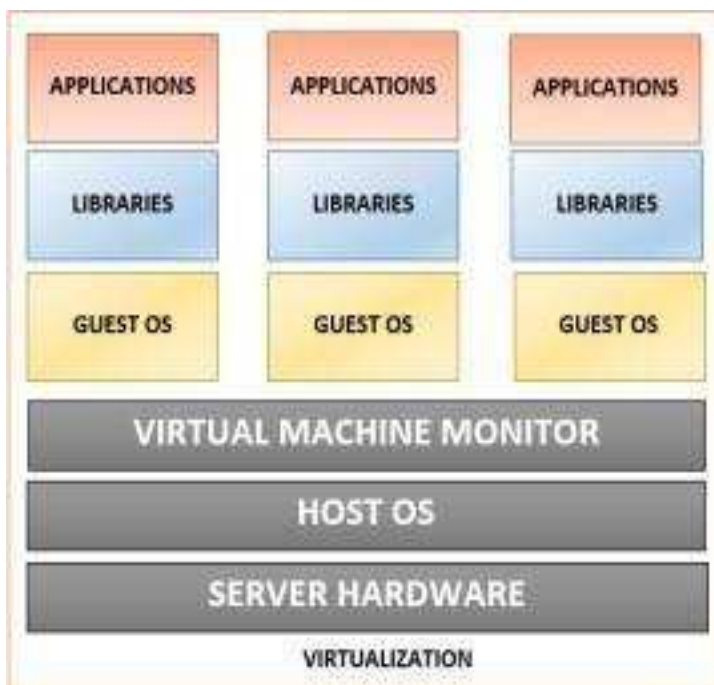


Figure 2: Overview of Virtualization

1.4 Types of Virtualization

Application Virtualization: In this technique, a software or program is not actually installed on the workstation on which it will be used but rather it is deployed on the server. Although the program behaves and thinks like it directly interacts with the workstation's OS but in reality it does not. A virtualization barrier is installed between OS and the program for this trick to work. An example would be virtualizing Microsoft PowerPoint to run on Ubuntu over an Opera browser.

a) Network Virtualization:

Network virtualization is like creating a virtual version of a physical network. In other words, having a "network on your computer" that behaves like a real network, allowing you to run and test different networking scenarios without needing a physical setup. This can be useful for various purposes, such as testing new network configurations, simulating complex network setups, or isolating networks for security purposes, all in a virtual space rather than physically rewiring or reconfiguring hardware.

b) Desktop Virtualization:

Desktop virtualization is a technology that allows you to run multiple virtual desktops on a single physical computer. It's like having several computers inside one physical computer. Each user gets their own virtual space that looks and feels like a regular desktop, even though it's running on a shared machine. This approach can make it easier to manage and deploy desktop environments, as well as improve resource utilization and flexibility in an organization.

c) **Storage Virtualization:** It's like having a giant, virtual storage space that combines the capacity of multiple physical storage devices. This virtualization layer enables easier management, better utilization of storage resources, and the

ability to allocate or expand storage without directly dealing with the complexities of individual physical devices. It provides a more streamlined and adaptable approach to handling data storage for computers and networks.

- d) **Server Virtualization:** It is like dividing a single powerful computer into several smaller, independent computers. Each virtual server operates as if it has its own dedicated hardware, even though it's sharing the resources of the underlying physical server. This approach improves the efficient use of server resources, allowing for better scalability, flexibility, and cost-effectiveness in managing computing workloads.
- e) **Data virtualization:** It is a technology that allows you to access and use data from different sources as if it's all in one place, even if it physically resides in various locations and formats. It creates a layer of abstraction that hides the complexities of where the data is stored, making it easier to retrieve and use information without needing to know its exact location.

2. What is Emulation?

An emulator is a tool, whether in hardware or software, that allows one device (the Host) to mimic the functions of other systems (the Guest). It provides a means to run both the hardware and software of one system on a different one. While emulation introduces some additional processing overhead, it offers several advantages. It's cost effective, easily accessible, and enables the execution of programs that might be obsolete in the current system.

Emulators work by translating CPU instructions meant for one architecture into instructions that can be successfully executed on a different architecture. These emulation systems can be accessed remotely and are user-friendly. They are particularly valuable for embedded and operating system development, allowing testing without affecting the

underlying operating system. Emulation effectively contains the size of the design under test (DUT), making it a versatile tool regardless of the host system's capabilities.

2.1 Difference between Emulation and Virtualization



Figure 3: Emulation vs Virtualization

Here are the key distinctions between emulation and virtualization:

1. Software Bridge vs. Direct

Hardware Access: Emulation requires a software bridge, while virtualization directly accesses hardware.

2. Execution of Code in Various Domains:

Virtual machines execute code directly in the language of use, unlike emulators.

3. Interpreter for Basic Emulation:

Emulation relies on an interpreter to translate source code for processing on the system.

4. Execution Environment and Speed:

Emulators don't run the guest OS on physical hardware, and they are generally slower compared to virtual machines, which utilize the CPU more efficiently.

5. Layer Between Hardware:

Virtualization introduces a layer between hardware for control, while emulation doesn't. Virtualization allows efficient resource sharing among guest machines.

	EMULATION	VIRTUALIZATION
DEFINITION	Refers to mimicking the functions of one system on another.	Refers to creating virtual instances of resources.
PURPOSE	Run software designed for one system on another.	Efficiently utilize physical resources.
ISOLATION	Provides less isolation, as it replicates the behavior of the original system.	Provides strong isolation between virtual instances.
RESOURCE UTILIZATION	Less efficient, as it replicates functions on a different system.	Maximizes resource efficiency by sharing resources among virtual machines.
PERFORMANCE	High overhead due to emulation layer.	Provides better performance as it runs directly on hardware or hypervisor.
USE CASE	Useful for compatibility or migration purposes.	Ideal for server consolidation, resources optimization, and workload isolation.
COMMON TECHNOLOGIES	QEMU, bochs, DOSBox, Appetize.io, Riasetta 2, BlueStacks.	VMware, Hyper-V, KVM, Xen, Docker (containerization).
EXAMPLE	Running legacy applications on modern hardware.	Creating multiple virtual machines on a single physical server.

Table 1: Differences between Emulation and Virtualization

The core similarity between virtualization and emulation lies in their common objective of mimicking hardware through different methods. Both enable the execution of a program in an environment intended for another, albeit with distinct techniques. Both virtualization and emulation provide solutions for deploying multiple isolated services without the need for separate physical platforms. These approaches serve specific standards and operate at different levels within the realm of computing technologies.

Bibliography

1. (Difference between Emulation and Virtualization, 2020)
2. (Virtualization in Cloud Computing and Types, 2023)
3. (What is network virtualization?)
4. (Data Virtualization: Process, Components, Benefits, and Available Tools, 2021)

Conclusion

In summary, the distinction between emulation and virtualization can be muddled, as emulation" is sometimes used interchangeably with a virtual environment in server virtualization. Emulation involves replicating both the hardware and software of the system we want to mimic on the host device. In contrast, virtualization emulates specific hardware components based on specifications, facilitated by the guest OS to run seamlessly on the same architecture.

AI and Data Virtualization: Unleashing the Power of Insights

Faculty Mentor: Dr. Neha Punetha

Students Name:

Vaibhav Singh (MCA- sem 1)

Priyanka Mahto (MCA-sem 1)

INTRODUCTION

Artificial Intelligence (AI) and Data Virtualization are two rapidly evolving fields that have revolutionized the way we analyze and interpret data. In this article, we will explore the significance of AI in tackling data challenges and the role of data virtualization in presenting complex information in a visually appealing manner (Sathyan et al., 2022).

1. What is Data Virtualization?

Data virtualization is the process of representing complex data sets visually, enabling users to comprehend patterns, trends, and insights more effectively. It is a modern data integration technique that integrates data in real time, without having to physically replicate it. It transforms raw data into intuitive charts, graphs, and interactive dashboards, making it easier for decision-makers to understand and communicate information.

Data virtualization tools like Tableau, Power BI, and D3.js have gained popularity due to their ability to present data in a visually appealing and interactive manner (Pang et al., 2020). By visualizing data, organizations can uncover hidden patterns, identify outliers, and communicate complex information in a more accessible way. Data virtualization combines data much faster than traditional way of data abstract and then publish data in different applications like AI tools such as Python, R programming (Pang et al., 2020).

1.1 The Data Challenge for AI Projects

AI projects heavily rely on vast amounts of data for training and decision-making. 87% of data science projects never make it into production because the quality, variety, and volume of data pose significant challenges. Ensuring data accuracy, relevance, and accessibility becomes crucial for successful AI implementation (Namasudra et al., 2023).

implementation of AI and data virtualization. The other major task is how to obtain accurate or relevant data? As we all know AI comes into implementation so that it can help human better, for this, it need to understand humans better and how their mind and action work. A human have it's five sense which predict or decide their action. Nowadays AI have power for visualization, sound production, sound detection but it still lacks. A machine can't feel and detect odor which leads to insufficient and inaccuracy in data for understanding human better.

1.2 Tackling the Data Pipeline Problem with Data Virtualization

The data pipeline problem refers to the challenges organizations face in integrating and accessing data from various sources. Data virtualization offers a solution by providing a unified view of data, regardless of its location or format. It allows organizations to access and analyze data in real-time without the need for complex data integration processes. By reducing data redundancy and improving data agility, data virtualization enables faster decision-making and enhances the efficiency of AI projects. With data virtualization, organizations can overcome the limitations of traditional data integration methods and gain a holistic view of their data landscape (Lin et al., 2023). Company's model with large volumes of combined data, including information about customer preferences, purchase habits, demographics, product ingredients, and product qualities, to arrive at the most effective recommended formulas. **Now days, Healthcare Company** uses the Denodo Platform to combine data about individual patients with historical patient data, and then the Denodo Platform feeds the combined data to an AI/ML system, which helps determine the appropriate course of action for individual cases and sends alerts when necessary.

There is not a single problem which comes into the path of

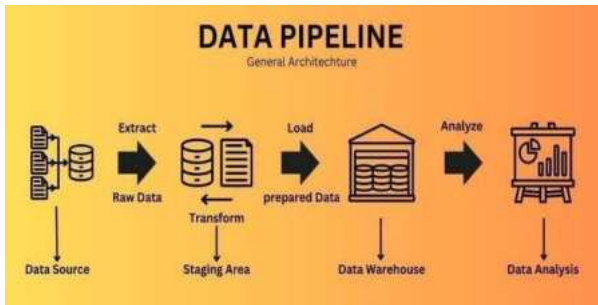


Fig1: DATA PIPELINE – Types, Architecture, & Analysis

2. Customer Case Study

One company that has successfully leveraged AI and data visualization is **Netflix**. By analyzing user data, **Netflix** is able to make personalized recommendations to its users, increasing user engagement and retention. **Netflix** also uses data visualization to identify patterns in user behavior, and uses this information to optimize its content offerings.

The Denodo Platform uses data virtualization to integrate data in real time, without having to physically replicate it. It can seamlessly combine views of data from a wide variety of different data sources and feed AI/ML engines with data from a common data services layer. (Kang & Park, 2014)

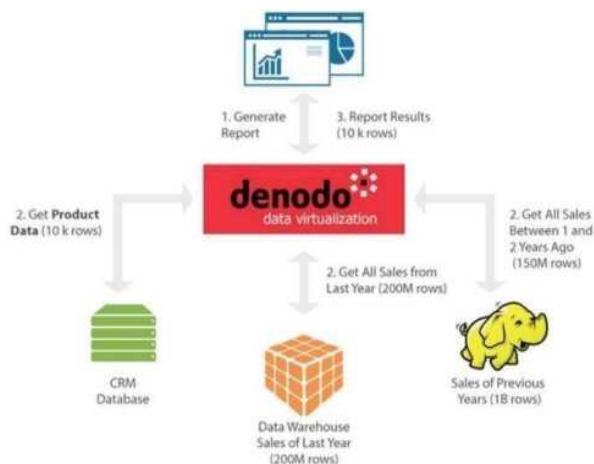


Fig2: Denodo Query Optimizations for the Logical Data Warehouse

Many leading Food Flavoring Company leveraged the Denodo Platform to build an AI model for predicting what kinds of flavors would appeal to customers in different demographics. The Denodo Platform feeds the



Fig3: Data Virtualization Market Size & Share | Segment Review- 2030

3. Data Virtualization need to work in certain fields

With data virtualization, the connection to all necessary data sources must be operational as there is no local copy of the data, which is one of the main drawbacks of the approach. Connection problems occur more often in complex systems where one or more crucial sources will occasionally be unavailable. Smart data buffering, such as keeping the data from the most recent few requests in the virtualization system buffer can help to mitigate this issue.

A single point of failure frequently emerges from the virtualization server's sole access point to all data sources. All operational systems risk losing their data feeds if the server goes down (Sohrabi & Azgomi, 2019). When data is delivered for analysis, data virtualization can help to resolve privacy-related problems. Virtualization makes it possible to combine personal data from different sources without physically copying them to another location while also limiting the view to all other collected variables. However, virtualization does not eliminate the requirement to confirm the security and privacy of the analysis results before making them more widely available. Regardless of the chosen data integration method, all results based on personal level data should be protected with the appropriate privacy requirements. There is no batch data support. The integration method does not support transporting data in batches or in bulk, which may be necessary for various circumstances.

3.1 Key Takeaways

- AI and data virtualization are powerful tools that complement each other, enabling organizations to extract valuable insights from complex data.
- Effective data management is crucial for successful AI implementation, emphasizing the need for accurate, relevant, and accessible data.
- Data virtualization simplifies complex information, making it easier for decision-makers to understand and communicate insights.
- Data virtualization addresses the data pipeline problem, providing a unified view of data and enhancing the efficiency of AI projects.

3.2 Next Steps

As AI and data virtualization continue to evolve, organizations should focus on:

- Investing in advanced data management systems and data science expertise.
- Exploring innovative data visualization techniques and tools to enhance data comprehension.
- Embracing data virtualization to overcome data integration challenges and improve AI project efficiency.
- Continuously monitoring and adapting AI models and data visualization strategies to stay ahead in the rapidly changing technological landscape.

4. CONCLUSION

AI and data virtualization have emerged as indispensable tools for organizations seeking to unlock the power of data. By addressing data challenges, simplifying complex information, and enabling real-time data access, AI and data visualization empower decision-makers to make informed choices. As we move forward, the integration of AI and data visualization will continue to shape industries, revolutionizing the way we analyze, interpret, and present data. By harnessing the potential of AI and data virtualization, organizations can gain a competitive advantage and drive innovation in the data-driven world.

REFERENCES

- Lin, S. S., Zhou, A., & Shen, S. L. (2023). Safety assessment of excavation system via TOPSIS-based MCDM modelling in fuzzy environment. *Applied Soft Computing*, 138, 110206. <https://doi.org/10.1016/j.asoc.2023.110206>
- Namasudra, S., 4.0, K. A.-B. and its A. in I., & 2023, undefined. (2023). Introduction to Blockchain Technology. *Springer*, 119, 1–28. https://doi.org/10.1007/978-981-19-8730-4_1
- Pang, X., Zhou, Y., Wang, P., Lin, W., & Chang, V. (2020). An innovative neural network approach for stock market prediction. *Journal of Supercomputing*, 76(3), 2098–2118. <https://doi.org/10.1007/S11227-017-2228-Y/FIGURES/20>
- Sathyan, R., Parthiban, P., Dhanalakshmi, R., & Sachin, M. S. (2022). An integrated Fuzzy MCDM approach for modelling and prioritising the enablers of responsiveness in automotive supply chain using Fuzzy DEMATEL, Fuzzy AHP and Fuzzy TOPSIS. *Soft Computing* 2022 27:1, 27(1), 257–277. <https://doi.org/10.1007/S00500-022-07591-X>
- Sohrabi, M. K., & Azgomi, H. (2019). Evolutionary game theory approach to materialized view selection in data warehouses. *Knowledge-Based Systems*, 163, 558–571. <https://doi.org/10.1016/J.KNOSYS.2018.09.0>

ABOUT DEPARTMENT OF INFORMATION TECHNOLOGY

Vision & Mission of the Institute

OUR VISION

To be an Institute of Academic Excellence with total commitment to quality education and research in Management and Information Technology with a holistic concern for better life, environment and society.

OUR MISSION

To serve the society and improve the quality of life by imparting high quality education in management and information technology, providing training and development services, fostering research, giving consultancy services to industry and disseminating knowledge through the publication of books, journals and magazines.

Department of Information Technology

Vision

- To be a department of excellence in technical education, widely known for the development of competent and socially responsible IT professionals, entrepreneurs and researchers

Mission

- To impart established and contemporary technical knowledge.
- To synchronize concepts, logic and skills for effective decision making.
- To encourage entrepreneurial environment and nurture innovative ideas.
- To foster research and provide consultancy service to the corporate.
- To utilize technical knowledge of students towards social issues through various group activities.



JAGAN INSTITUTE OF MANAGEMENT STUDIES
3, Institutional Area, Sector-5, Rohini
Email: techbyte@jimsindia.org
Website: <https://www.jimsindia.org>