ISSN: 2754-6659

Journal of Artificial Intelligence & Cloud Computing

Review Article

Open Access

Enhancing Virtual Desktop Infrastructure with Hybrid Cloud Environments and GPU Acceleration

Sriramaraju Sagi

NetApp, USA

ABSTRACT

This research paper focuses on the implementation and testing of a cloud based Virtual Desktop Infrastructure (VDI) that combines both cloud technology and GPU acceleration. The paper explores how various technologies, such, as Citrix Virtual Apps, Windows 11 Virtual Desktops, VMware vSphere and NetApp Storage can be integrated within a data center framework. The study emphasizes the importance of hybrid cloud environments in addressing scalability, security and disaster recovery challenges faced by VDI systems. In particular it highlights the significance of NVIDIA A40 cards for GPU accelerated Virtual Desktop Infrastructure (VDI) and the use of SPECviewperf 2020 to evaluate performance. The validation process involves conducting tests, on disaster recovery scenarios using NetApp SnapMirror while also evaluating the performance of NetApp AFF A400 storage systems. The results demonstrate that this system effectively handles latency issues and performs well in large scale VDI deployments.

***Corresponding author**

Sriramaraju Sagi, NetApp, USA.

Received: June 03, 2023; **Accepted:** June 12, 2023; **Published:** June 22, 2023

Keywords: Virtual Desktop Infrastructure (VDI), Hybrid Cloud Environments, GPU Acceleration, Citrix Virtual Apps

Introduction

Virtual Desktop Infrastructure or VDI plays a role, in computer environments. It allows for control and distribution of desktops to users offering flexibility, scalability and enhanced security. By virtualizing workstations VDI reduces hardware costs and simplifies maintenance. VDI is a technology in today's computer landscape providing adaptability and efficiency in managing and accessing resources. Its ability to separate the desktop environment from hardware enables administration and improves security standards. This shift towards virtualization meets the growing demand for computing resources catering to computational tasks ranging from data intensive applications to complex simulations. Furthermore VDI seamlessly integrates with emerging technologies like cloud computing and GPU acceleration further enhancing its significance. It offers solutions that can adapt to the evolving needs of both industry and academia. This integration not only optimizes computing resources, it also promotes sustainability by reducing the need for frequent hardware updates. In summary VDI is a component in advancing computer infrastructures. It embodies efficiency, scalability and security. All aspects, in today's era.

Hybrid Cloud VDI Environments

Hybrid Cloud VDI setups are combinations of, on premises and cloud based infrastructures that create a platform for delivering Virtual Desktop Infrastructure. This integration allows businesses to leverage the clouds scalability and adaptability while still maintaining control and security within their infrastructure.

By combining VDI with hybrid cloud architectures enterprises effectively address challenges like data sovereignty, latency issues and fluctuating demands for computational resources. The utilization of Hybrid Cloud VDI presents a solution for enterprises aiming to optimize their IT operations while adhering to regulatory and performance standards. It combines the strengths of on premises and cloud environments in a merger that not enhances operational efficiency but also facilitates the implementation of innovative application deployment scenarios, such as GPU accelerated computing workloads. This strengthens the role of VDI, in computing environments.

Incorporating VDI into hybrid cloud infrastructures is a step, in how businesses manage their computer systems. This setup allows for integration of both private resources giving organizations the flexibility to allocate desktop services based on real time needs and workloads. By using VDI in environments it opens up a wide range of practical applications. These include managing high demand periods by utilizing cloud resources ensuring data security and compliance through, on site infrastructure. Additionally this hybrid approach enables the use of cloud based GPU acceleration enhancing the performance of applications and providing an exceptional user experience. The adaptability and scalability offered by Hybrid Cloud VDI environments make them a progressive choice that aligns with the requirements of computing tasks and overall digital transformation goals.

As the discussion surrounding Hybrid Cloud VDI progresses it becomes evident that the flexibility provided by these environments greatly empowers enterprises to navigate the complexities of workplaces. This agility is supported by the ability to dynamically adjust resources as needed ensuring fulfillment of compute and storage requirements. Consequently this reduces wastage. Maximizes cost effectiveness. Moreover the hybrid approach enables a disaster recovery strategy by utilizing cloud resources for backup and failover capabilities thereby enhancing business continuity planning. By incorporating VDI into hybrid cloud infrastructures organizations can meet both operational requirements while strategically positioning themselves to leverage advancements in cloud services. This ensures their competitiveness, in a interconnected global environment.

The challenges associated with scalability, in cloud VDI implementations usually revolve around finding the balance between on premises infrastructure and cloud resources. It can be difficult to manage data consistency, network capacity and latency across scenarios especially when demand fluctuates. In installations the main focus is on security, which involves safeguarding data controlling access and mitigating threats for both cloud and on premises components. To address these challenges effectively certain measures can be taken. These include implementing encryption using factor authentication conducting regular security audits and adopting a zero trust security framework to ensure comprehensive protection and compliance with legal requirements. These procedures are critical for maintaining the integrity and confidentiality of data within hybrid cloud VDI systems.

To handle scalability in hybrid cloud VDI deployments successfully a dynamic resource allocation framework is necessary. This framework efficiently manages the distribution of resources between on premises infrastructure and the cloud components to ensure performance even as user demands increase. Achieving this requires monitoring capabilities and predictive analytics to adjust resources and prevent potential bottlenecks. In terms of security measures it is essential to go beyond precautions by implementing threat detection systems and response mechanisms. Additionally conducting evaluations of policies and configurations plays a role, in maintaining robust security measures.

These solutions work together to ensure that hybrid cloud VDI deployments are both secure and easily scalable enabling enterprises to leverage the advantages offered by these infrastructures.

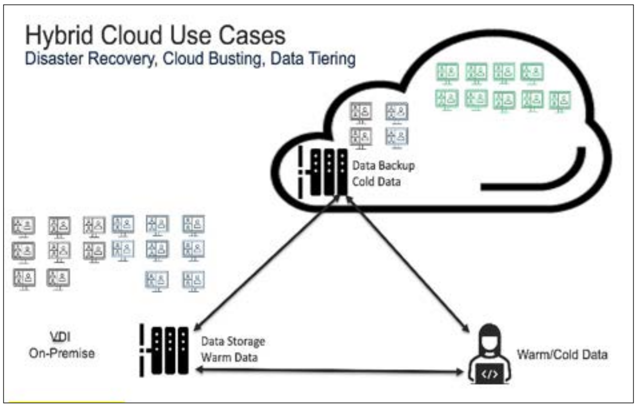

Figure 1: VDI Hybrid Coud Use Case

GPU Acceleration VDI

The importance of GPU acceleration, in applications cannot be overstated. It significantly enhances efficiency in tasks that involve handling large amounts of data. In the context of Virtual Desktop Infrastructure (VDI) GPUs are essential for improving performance delivering high quality graphics instantly and supporting applications. This enhancement is particularly vital in industries that rely on computing power for tasks involving graphics, such as data analysis, artificial intelligence and 3D design. By offloading these activities to GPUs VDI environments can offer users an responsive experience when dealing with resource intensive applications. This highlights the impact of GPU acceleration in computing. Incorporating GPUs into VDI setups not boosts performance but also enables faster computing processes making it crucial for a wide range of modern applications. Acceleration plays a role in achieving real time analytics deep learning algorithms and advanced scientific simulations underscoring the contribution of GPUs to enhancing the capabilities and efficiency of VDI platforms. The integration of GPU acceleration and VDI will continue to play an role, in meeting growing computational demands by providing enhanced resources conveniently and efficiently.

In order to optimize GPU acceleration, for desktops we need to address technological challenges. These include ensuring allocation of GPU resources effectively managing the scheduling of GPUs and reducing latency for end users. It is crucial to enhance performance in order to achieve powerful computing experiences for graphics intensive applications. These challenges involve the complexity of integrating GPUs into virtualized environments distributing workloads across desktops and ensuring compatibility across different hardware and software configurations. Overcoming these obstacles is essential to leverage the benefits of GPU acceleration in VDI environments resulting in improved performance and user satisfaction. To enhance GPU acceleration for desktops it is important to implement solutions that specifically target these challenges. Dynamic resource allocation methods can be utilized to adjust GPU resources based on real time demand while employing tools for managing GPUs can effectively distribute their capabilities, across multiple workstations. Additionally optimizing network protocols and infrastructure can significantly reduce latency. Enhance the user experience in GPU accelerated VDI setups. To achieve efficient utilization of GPUs an understanding of both the virtualization layer and the underlying physical hardware is essential.

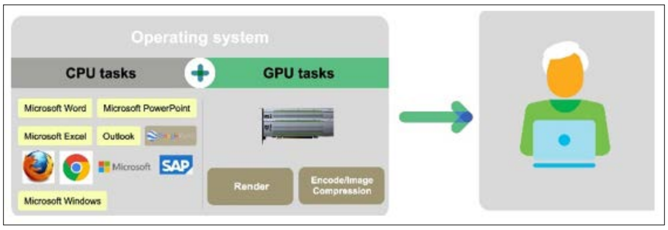

Figure 2: VDI CPU vs GPU Tasks

Literature Review

Several research studies have explored the possibilities and obstacles associated with cloud virtual desktop infrastructure. Chirivella Perez and Dutta put forward cloud solutions that combine the advantages of both public clouds [1, 2]. Chirivella Perez emphasizes the potential, for growth while Dutta focuses on resource management. The research suggests implementing a Virtual Physical deployment for the OpenStack cloud management platform. The goal of this approach is to enhance the scalability of the cloud architecture and simplify the process of deploying services all while reducing time and complexity. Thorough testing has been conducted on this proposed infrastructure under demanding conditions. The management services have been improved to include real time analysis of network traffic within the cloud infrastructure showcasing the scalability of this suggested framework. Dutta's paper suggests a mechanism for managing resources in a Hybrid Cloud model to handle VM requests. This strategy has proven to be cost efficient. Reduces company spending. Future work may involve developing a dynamic resource allocation technique aimed at lowering the energy consumption rate, in Cloud data centers.

In the research conducted by Xu and Loreti they explored ways of using computing resources and implementing clusters in hybrid clouds [3,4]. Their findings highlight the evolution of cloud computing, towards hybrid cloud environments and the significant impact of cloud bandwidth on performance. They also suggest that dynamic off premise provisioning could enhance execution. Hoecke and Sotomayor Borja have developed management tools for clouds. Hoeckes tool focuses on load balancing and dynamic upscaling while SotomayorBorjas solution addresses virtual infrastructure management [5, 6]. Gillette and Tzanakaki have conducted studies examining the use of cloud computing and virtualization in contexts such as infrastructures and wireless optical networks [7,8]. These studies emphasize how cloud virtual desktop infrastructure can improve resource management, scalability and efficiency. By integrating cloud computing with virtualization technologies it becomes possible to enhance interoperability reduce network footprints and improve resource elasticity. Additionally this integration can mitigate limitations in data transmission capacity by enabling storage, distribution and retrieval of data within a location. Tzanakakis article presents two findings; the proposal of an optimized virtual infrastructure and the direct relationship between service volume/characteristics, with energy consumption and resource requirements.

Various studies have explored the use of GPU accelerated applications, in desktop infrastructure (VDI). According to Chang and Yang advanced VDI solutions that integrate GPUs on OpenStack offer promising possibilities [9,10]. Changs work primarily focused on evaluating and comparing performance with CPUs while Yang concentrated on improving the display performance of machines. The study suggests that combining a Graphics Processing Unit (GPU) with OpenStack through PCI pass through can result in a high performance Virtual Desktop Infrastructure (VDI). This integration brings benefits such as reduced latency improved speed and enhanced functionality. By leveraging GRID GPU technology multiple virtual machines (VMs) can utilize a GPU of the traditional one to one configuration between the GPU and user. This allocation of GPU resources has the potential to significantly improve user satisfaction and optimize performance in desktop environments. Yangs research highlights findings including the recommended integration of GPU technology with OpenStack for enhanced virtual machine display performance categorization of three processing drivers, for virtual desktops and the importance of delivering a seamless cloud based virtual desktop experience.

In their studies Shi and Vasilas developed frameworks to improve the performance of graphics processing units (GPUs), in machines. Shis framework, known as vCUDA focuses on enhancing high performance computing tasks while Vasilas framework, VGVM aims to integrate general purpose GPU (GPGPU) capabilities into virtualized settings [11,12]. The vCUDA framework enables the utilization of hardware acceleration for high performance computing applications within machines showcasing its practicality and competitiveness compared to non virtualized contexts. The report also identifies the source of overhead. Proposes potential future enhancements. Key findings from the study highlight the performance advantages that GPUs offer for data parallel applications the growing need for integrating GPU benefits and virtualization in environments and the introduction of VGVM as a mechanism to enable GPU accelerated application execution within Virtual Machines (VMs).

Li and Liu both explored GPU accelerated VDI platforms in their research [13]. Lis study involved comparing virtualization technologies while Lius focus was on video processing [14]. Schmitt on the hand proposed methods to provide CUDA acceleration and GPGPU capabilities within virtualized environments [15]. Duato implemented rCUDA for this purpose while Schmitt introduced support, for 3D graphics.

The user did not provide any text. The research paper introduces an approach that effectively utilizes GPGPU capabilities, in virtual machines. This is achieved by utilizing the rCUDA framework. The experiments demonstrate the practicality and potential for expansion of this method indicating that rCUDA can offer CUDA based acceleration support to virtual machines running on a single physical server equipped with multiple GPUs. Furthermore the trials showcase scalability suggesting that the rCUDA framework can be efficiently implemented in a system, with numerous concurrent virtual machines.

Validation and Analysis

The goal of this research is to explore the application of a designed data center architecture that incorporates various technologies such, as Citrix Virtual Apps and Desktops Remote Desktop Server Hosted (RDSH) sessions Windows 11 Virtual desktops VMware vSphere, Cisco Unified Computing System (Cisco UCS) Cisco Nexus 9000 switches and NetApp Storage AFF A400 All Flash array with iSCSI storage access.

This paper specifically focuses on using Hybrid Cloud VDI for two purposes; disaster recovery (DR) and backup. Businesses encounter types of disasters, which makes safeguarding data through disaster recovery (DR) crucial for ensuring business continuity. With disaster recovery (DR) companies can seamlessly transfer their operations to a location when faced with a failure and then restore and return to the site in a reliable manner. NetApp SnapMirror is a solution designed specifically for disaster recovery scenarios where primary data needs to be replicated to storage located at a remote geographical location. SnapMirror serves as a replication tool that creates a copy or mirror of data on secondary storage. This ensures that you can continue providing data services even if there is a disaster, at the site. The VDI components of this disaster recovery system were implemented using Citrix Cloud.

NetApp BlueXP and CVO were used to protect the data and virtual machines, on premises. Microsoft Azure was utilized for ensuring the functionality of cloud based infrastructure components such as Active Directory, SQL, DNS and Virtual Machines. This solution included implementing and verifying a disaster recovery plan. To confirm the success of a disaster recovery (DR) scenario start by transferring data from an ONTAP volume to Cloud Volumes ONTAP using SnapMirror. Then retrieve the data, from the Microsoft Azure cloud compute instance. Proceed with verifying data integrity.

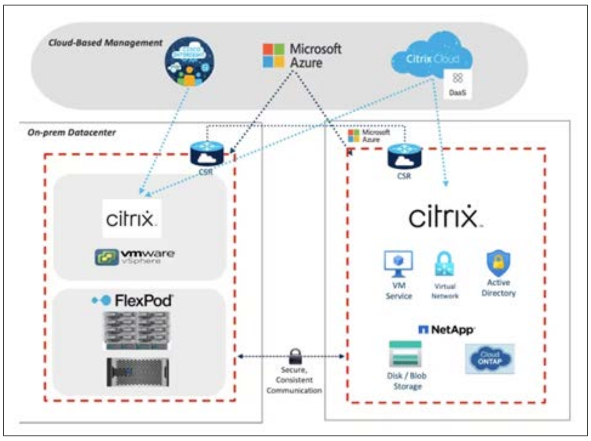

Figure 3: Hybrid Cloud to Azure

We conducted tests on VDI workloads using NVIDIA A40 cards to evaluate their performance. The GPU test workload consisted of a solitary server equipped with two A40 GPUs. Every Windows 11 computer was equipped with an 8GB NVIDIA Grid profiles. The inclusion of NVIDIA A40 GPUs in the solution has a beneficial effect on the end-user experience, particularly for high-end power users. By employing SPECviewperf 2020, we conducted a performance assessment of NVIDIA GPUs in the Cisco UCS X210c M7. The results revealed exceptional Composite scores for executing resource-intensive graphic-accelerated tasks.

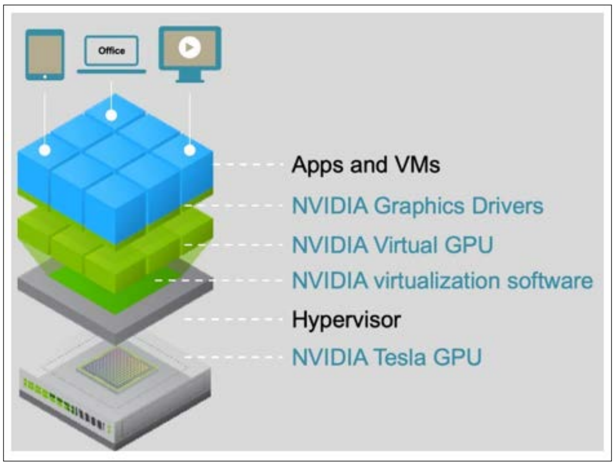

Figure 4: NVIDIA GPU for Virtual Desktops

We analyzed the performance results of the NetApp AFF A400 storage system for each of the Citrix software modules (HSD, PVS, Persistent) during cluster testing. To ensure an excellent end-user experience, it is crucial to maintain a latency of less than one millisecond from a storage standpoint, regardless of the IOPS and bandwidth being utilized. The test results demonstrate that the AFF A400 storage effectively maintains a minimal amount of latency, even when accommodating thousands of desktops on the system. The screenshots below shows the AFF A400 storage test data with information about IOPS, bandwidth, and latency at the peak of each use case test.

Figure 5: Volume Average Latency, Volume Total Throughput and Volume Total IOPS for 2500 RDS

Conclusion

This article effectively highlights the implementation and validation of a cloud Virtual Desktop Infrastructure (VDI) that incorporates GPU acceleration. By utilizing technologies, like Citrix Virtual Apps Windows 11 Virtual Desktops VMware vSphere and NetApp Storage in conjunction with NVIDIA A40 GPUs it has been proven that this VDI system is robust scalable, secure and highly efficient. The verification process includes testing disaster recovery scenarios using NetApp SnapMirror and conducting performance analysis of NetApp AFF A400 storage systems to ensure performance during delays. This study not showcases the feasibility of deploying VDI on a scale but also emphasizes the importance of GPU acceleration in modern computing environments. It represents an advancement, in the utilization of cloud VDI applications.

References

1. Chirivella PE, Juan GA, José MC, Jose MAC (2013) Hybrid and Extensible Architecture for Cloud Infrastructure Deployment. 2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing 611-617.
2. Dutta D, Kaustav C, Kasturi S (2013) Resource Management in a Hybrid Cloud Infrastructure. International Journal of Computer Applications 79: 41-45.
3. Xu B, Chunyan L (2011) A Cloud Computing Infrastructure on Heterogeneous Computing Resources. J Comput 6: 1789-1796.
4. Loreti D, Anna C (2015) A Hybrid Cloud Infrastructure for Big Data Applications. 2015 IEEE 17th International Conference on High Performance Computing and Communications 1713-1718.
5. Hoecke S, Tom W (2011) Efficient Management of Hybrid Clouds. IEEE International Conference on Cloud Computing.
6. Borja S, Rubén SM, Ignacio ML, Ian F (2009) Virtual Infrastructure Management in Private and Hybrid Clouds. IEEE Internet Computing 13: 14-22.
7. Gillette SE (2012) Cloud Computing and Virtual Desktop Infrastructures in Afloat Environments. NPS https://apps.dtic.mil/sti/tr/pdf/ADA562746.pdf.
8. Tzanakaki A, Anna T, Markos PA, Georgios SZ, Bijan RR, Reza N (2013) Virtualization of heterogeneous wireless-optical network and IT infrastructures in support of cloud and mobile cloud services. IEEE Communications Magazine 51: 155-161.
9. Chang CH, Chao TY, Jheng YL, Chuan LL, Chia CK (2020) On Construction and Performance Evaluation of a Virtual Desktop Infrastructure With GPU Accelerated. IEEE Access

8: 170162-170173.
10. Chao TY, Jung CL, Jheng YL, Chih HC, Chuan LL, et al. (2018) The Implementation of a Virtual Desktop Infrastructure with GPU Accelerated on OpenStack. 2018 15th International Symposium on Pervasive Systems, Algorithms and Networks (I-SPAN) 366-370.
11. Lin S, Hao C, Jianhua S (2009) vCUDA: GPU accelerated high performance computing in virtual machines. IEEE International Symposium on Parallel & Distributed Processing 1-11.
12. Dimitrios V, Stefanos G, Nectarios K (2016) VGVM: Efficient GPU capabilities in virtual machines. 2016 International Conference on High Performance Computing & Simulation (HPCS) 637-644.
13. Dan L, Huiran Z, Jichao Z, Wenfeng S, Min C, et al. (2016) Video processing using GPU-accelerator under desktop virtualization environment. 2016 International Conference on Audio, Language and Image Processing (ICALIP) 766-770.
14. Jose D, Antonio JP, Federico S,Juan CF, Rafael M, et al. (2011) Enabling CUDA acceleration within virtual machines using rCUDA. 2011 18th International Conference on High Performance Computing 1-10.
15. Thilo S, Alexander W, Christian H, Peter S (2010) Towards GPGPU Assisted Computing in Virtualized Environments. Euro-Par Workshops https://link.springer.com/chapter/10.1007/978-3-642-21878-1_80.