ISSN 2616-7433 Vol. 7, Issue 8: 23-27, DOI: 10.25236/FSST.2025.070804

Solutions to Storage Performance Mismatch in the Digital Transformation Applications of Higher Education

Fanrui Lin[a], Liangfang Wang[b,*]

Wenzhou Medical University, Wenzhou, Zhejiang, China
[a]lfr@wmu.edu.cn, [b]wangliangfang@wmu.edu.cn
**Corresponding author*

Abstract: *Against the backdrop of the accelerated digital transformation in higher education, the storage performance of traditional data centers is limited by factors such as vertical scaling, resulting in a mismatch with the dynamic resource demands and peak data traffic in application scenarios, making it difficult to meet business requirements. This paper proposes a solution that adopts the Hyper-Converged Infrastructure (HCI) distributed storage architecture. It addresses the issue of storage performance mismatch by reconstructing the data center architecture, optimizing both storage and business continuity migration strategies, and enhancing technical and cultural training,etc. It can effectively adapt to the dynamic load changes in future digital transformation application scenarios and provides a reusable technical path for the construction of data centers in similar institutions.*

Keywords: *Higher Education; Digital Transformation; Storage Architecture; Hyper-converged; Data Center*

1. Introduction

The application scenarios of digital transformation in higher education have introduced new demands for storage performance. During the semester opening period, tens of thousands of students simultaneously use online course selection systems to make course selections, check class schedules, etc. At this time, the system will face a peak traffic volume during peak periods, and the sudden surge in instantaneous traffic is prone to causing stalls or even crashes. The sudden surge in data volume often leads to lagging or even crashes, Moreover, the allocation of teaching and research resources such as online teaching platforms and exam systems also exhibit seasonal fluctuations[1][2]. Meanwhile, new application scenarios such as real-time analytics of big data are rapidly becoming popular, and the total amount of educational data is showing an exponential upward trend. For innovative applications such as big data analysis and AI training, multi-node parallel computing is required to accelerate processing and exhibit high throughput characteristics. Only when deployed in a distributed storage architecture environment can they fully exert their functions[3]. As a key component of new infrastructure, data centers are the core infrastructure for the digital transformation of higher education, providing computing power and storage support for it. The ability to process and analyze massive amounts of educational data in real time through high-performance computing and storage resources directly determines the efficiency of application scenarios for the digital transformation of higher education. The traditional centralized SAN (Storage Area Network) storage architecture is limited by its expansion capabilities. It difficult to meet the core requirements of elastic scaling for the digital business in higher education, and unable to support the distributed parallel computing needed for AI training or the high throughput requirements for big data analysis[4]. Faced with these challenges, it has become urgent to build a new generation of data center architectures with elastic scalability, rapid deployment characteristics, and intelligent operation and maintenance. This architectural provides a solid technical foundation for the application scenarios of digital transformation in higher education.

2. The Phenomenon of Mismatch Between Traditional Architecture and Application Scenarios

Currently, the data centers of some domestic universities still utilize centralized SAN storage architecture. They construct dedicated networks through independent Fibre Channel (FC-SAN) or

ISSN 2616-7433 Vol. 7, Issue 8: 23-27, DOI: 10.25236/FSST.2025.070804

iSCSI protocol (IP-SAN) to isolate business traffic from the storage data flow, ensuring transmission bandwidth and low latency. Moreover, they provide block-level storage services for critical business and important data with high performance and stability as key factors[5]. Combined with virtualization technology, the efficiency of system deployment and management has increased by several times or even more.

However, with the changing performance requirements of digital transformation application scenarios in higher education, the traditional centralized SAN storage architecture, limited by its scalability, has become difficult to support and meet the new business demands. The traditional architecture uses dual controllers as nodes, whose performance does not improve with node expansion. The computing and storage are separated and adopt a vertical scaling approach, making the storage controllers becoming performance bottlenecks, with the maximum IOPS reaching only the order of ten thousand. With the rapid growth of virtual machines and the increase in load on storage arrays, it can lead to read-write bottlenecks. For example ,online teaching platforms, online examination systems, and online course selection systems may experience significant increases in I/O latency during peak traffic periods. At this time, the system accesses the database centrally, causing the I/O latency of the traditional centralized SAN storage to sharply increase from milliseconds to hundreds of milliseconds. The larger the access volume, the more obvious the performance bottleneck becomes, ultimately leading to the phenomenon of system loading lag. When tens of thousands of students are simultaneously selecting courses or take exams, traditional centralized SAN storage suffers drastic throughput drops and cliff-like performance declines. In addition, there will also be a phenomenon of out-of-control resource contention. when multiple virtual machines compete for shared storage resources via FC network, it can lead to storage lock conflicts, resulting in login timeouts during peak periods. For other innovative application scenarios, such as AI training and big data analysis, can only fully exert their function when deployed in a distributed storage architecture environment. However, the centralized storage architecture is not fully compatible with them and the performance is not satisfactory. Due to the hardware comes from different manufacturers, the operation and maintenance management interfaces vary. It is necessary for professionals with many years of relevant experience to operate it. The deployment is difficult and takes a long time, making it difficult to adapt to the rapidly changing business scenarios. Meanwhile, the maintenance costs of dedicated hardware are high, and the costs for vertical upgrades (such as high-end controllers, SAN switches) are also high, which are not friendly to universities with limited funds.In conclusion, traditional centralized storage is no longer capable of meeting the requirements of new application scenarios.

3. Technical Paths to Solve the Mismatch Between Storage Performance and Application Scenarios

This paper proposes a solution based on an emerging HCI distributed storage architecture. It addresses the challenge of storage performance mismatch by reconstructing the data center architecture, optimizing storage strategies and business continuity migration strategies, enhancing network continuity, strengthening data security barriers, and improving technical and cultural training, etc. The solution has been implemented in the data center of Wenzhou Medical University.

3.1. Performance Adaptation of HCI Architectures and Application Scenarios

The HCI architecture based on the hardware resource pooling architecture and a unified intelligent management platform, has gradually become a key technology for the digital transformation in the higher education industry. As a new generation of cloud data center solution, HCI achieves deep integration of computing, storage, network and security services through software defined technology. With the standard x86/ARM servers as the hardware carrier, it deeply integrates physical computing resources (CPU/memory), storage space and network functions via the virtualization engine, and solves the resource silo issues caused by the three-tier separation of "compute - storage - network" in the traditional IT architectures[6]. It supports dynamic pooling and on-demand allocation of computing and storage resources, combined with automated deployment tools, and can achieve minute level delivery of business systems. Through "Horizontal expansion, vertical performance enhancement and rapid deployment", the storage performance can be temporarily and rapidly improved to meet the challenges of peak traffic such as online teaching platform, online examination system, online course selection system, etc. The distributed architecture is horizontally scalable, with a maximum performance of up to millions of IOPS to meet the requirements of high concurrency . In conjunction with fully automated deployment, it can rapidly expand capacity on demand before the peak period. Meanwhile, innovative

ISSN 2616-7433 Vol. 7, Issue 8: 23-27, DOI: 10.25236/FSST.2025.070804

services such as AI training and big data analytics, etc. running them in a distributed storage architecture can fully leverage their effectiveness. The total cost of ownership (TCO) is low, it can be purchased on demand,and is more friendly and operational to universities with limited funds. In conclusion, HCI distributed storage architecture can fully meet the requirements of new application scenarios.

3.2. Reconstruct the Data Center Architecture Design

The core goal of Software-Defined Data Centers (SDDCs) is to enable customers to achieve agility, elastic scalability, and intelligent management experience at lower costs through abstraction, pooling, and automation technologies. The virtualization of data center mainly includes computing node virtualization, storage virtualization, network virtualization, and security virtualization, etc[7]. The renovation goal of the data center of Wenzhou Medical University is to establish a comprehensive virtualization infrastructure with high concurrency, stability, easy expansion, easy management and sustainable development, and to plan the business deployment over the next five years to meet the long-term and sustainable expansion needs of digital transformation in higher education. Therefore, a solution based on a the HCI distributed storage architecture has been designed, utilizing virtualization technology to comprehensively reintegrate the computing pool, storage pool, network pool ,etc. of the data center.

Integrates computing, storage, network and other resources into a unified and efficient pool to achieve flexible scheduling and allocation of resources. We utilize VMware vSphere for unified server virtualization management, deploy ESXi hosts, and manage them through vCenter Server to achieve on-demand allocation of hardware resources[8].

The computing pool is configured with 9 H3C R4900 G5 x86 rack server nodes (CPU: 9×2 Intel Gold 6330, memory: 9×512G), and the ESXi operating system is installed to form a resource pool for CPU and memory. The resource pool supports virtual machine real-time migration and dynamic resource scheduling. It can automatically achieve a balanced allocation of CPU and memory according to load conditions.

The storage pool is configured with a combination of 9*2*3.2TB NVMe SSD cache layers and 9*4*16TB NL SAS capacity layers, forming a mixed-flash pool. A virtual storage area network (vSAN) for a software-defined distributed storage system is established to replace the external disk array. In order to meet the requirements of different application scenarios for storage performance and capacity, the storage is partitioned into volumes. Specifically, fully SSD volumes are used to construct a high-performance storage pool to meet the requirements of high-performance applications, while mixed volumes of SSD and HDD are used to construct a large-capacity storage pool to meet the capacity needs.

The network pool is equipped with 36*10-gigabit physical network cards, which are connected by two stacked aggregation switches to ensure the continuity of the core network. The kernel ports are configured for ESXi management network, IP storage, migration and other functions. These ports are deployed as distributed vSwitch virtual switches, which divide into multiple Virtual Local Area Networks (VLANs) are respectively used for management, services and storage. Different services utilize different VLANs, the resources of the 10 Gbps service network, 10Gbps storage network, and 1Gbps management network are isolated to avoid conflicts. It not only ensures the transmission rate but also facilitates management and troubleshooting. The virtual machines accesses the vSwitch ports via the virtual network cards(vNICs), the uplink is connected to the external network via physical network cards of the server, thereby enabling communication between the virtual machines and the campus network.

3.3. Optimize Storage Strategy

Data classification is based on usage frequency and is divided hot data, warm data, and cold data. The tiered storage strategy selects different storage resource pools based on data classification. For example, in the online course selection system, high-frequency hot data is stored in high-performance storage pool, the historical course selection records, which are warm data, are stored in high-capacity storage pool, and the archived cold data are stored in object storage. By leveraging the dynamic expansion capability provided by vSAN, the storage capacity can be expanded on demand. The expanded resource pool serves as the physical foundation for data classification. Through the elastic expansion of storage resource pool and automatic data classification, the IOPS performance can be

ISSN 2616-7433 Vol. 7, Issue 8: 23-27, DOI: 10.25236/FSST.2025.070804

flexibly expanded according to the requirements of the application scenarios.the issue of system crashes during peak traffic periods is mitigated.

3.4. Enhance Network Continuity

The core backbone switch of the data center serves as the network convergence point for the entire data center. Its uplink port is connected to the campus core switch, while its downlink port is connected to the computing node servers. All the above network nodes are composed of two stacked switches.This solution provides mutual backup redundancy for switches in the core network links, significantly enhancing the network continuity.

3.5. Strengthen Data Security Protection Barriers

The main switch in the data center is connected to the WAF (Web Application Firewall), serving as the first layer of protection. The automatic recommendation micro-isolation strategy is configured for virtual machines, which generates the most suitable micro-isolation rules based on access patterns, historical traffic and user configurations. The operating system is configured with antivirus software, firewall, security settings, and dynamic access control ,etc. to prevent data leakage from tens of thousands of teachers and students. This serves as the second layer of protection. All of the above together form the entire security protection system for the data center.

3.6. Business Continuity Migration Strategy

As the centralized SAN storage architecture and the HCI distributed storage architecture are completely different, virtual machines cannot be migrated without shutting down. It is necessary to utilize VMware's two core technologies, vMotion (host computing resource migration) and Storage vMotion (storage resource migration), to achieve business continuity migration. Firstly, migrate the computing resources from the traditional architecture to the HCI architecture, and then migrate the storage part. We classify businesses according to their importance level and impact scope and adopt a phased migration strategy. We first prioritize migrating non-core, low-impact businesses, and then gradually migrate key and core businesses with broader impact, thereby ensuring a smooth migration.

3.7. Strengthen Personnel Skills and Cultural Training

Although the application of digital technology in the field of education has been relatively mature, higher education institutions lack professional and technical personnel[9].Teachers and staff generally exhibit resistance towards new technologies, preferring to remain in their comfort zones. This conservative culture has hindered innovation and developmet. On the one hand, they are already very familiar with the original technical process and can apply it with ease. On the other hand, the time and effort required to learn new technologies constitute a significant resistance to change in the context of heavy workloads. To address such issues, it is necessary to strengthen the training of operation and maintenance personnel in digital capabilities such as distributed architecture and software-defined storage, etc. When technical capabilities are improved, they will naturally appreciate the advantages brought by new technologies, operation and maintenance efficiency will be enhanced, and resistance will naturally disappear. Meanwhile, it is necessary to optimize the talent system and attract more experts who are proficient in new technologies[10].

4. Conclusion

The construction plan for university data centers based on the HCI distributed storage architecture has solved the problem that the traditional centralized SAN architectures struggling to meet the rapidly evolving needs of the digital transformation scenarios in higher education. The previous issue of frequent system crashes during peak traffic surges has been alleviated through elastic scaling of storage resource pool and automatic data classification. The concurrent performance of big data analytics, AI training, and other services has also significantly improved. The HCI architecture reduces the number of storage array devices, simplifies operation and maintenance management, and greatly improves the efficiency of problem-solving for IT personnel. It provides a reusable technical path for the construction of data centers in similar institutions.

ISSN 2616-7433 Vol. 7, Issue 8: 23-27, DOI: 10.25236/FSST.2025.070804

References

[1] Wen Jie, Wu Qingjie, Chen Yi, et al. Performance Optimization of Elective System at Fudan University[J]. Acta Scientiarum Naturalium Universitatis Sunyatseni, 2009,48(S1):95-99.DOI:CNKI: SUN:ZSDZ.0.2009-S1-031.

[2] Ma Biao, Liu Mingyan, Wang Shihao, Li Li. Digital Transformation of Educational Examination: Problems, Elements and Objectives——Take the Practive of Jiangsu Province for Example[J]. Modern Educational Technology, 2023,33(04):24-31.

[3] Rihong H .Rihong H . Research on the Practical Effectiveness and Innovative Path of Digital Transformation in Higher Education[J].Adult and Higher Education,2024,6(8): DOI:10.23977/ADUHE.2024.060810.

[4] Zhu Kanglin. Application of Distributed Virtualized Storage in Public Security College [J]. Computer Science, 2016,43(1):571-576.

[5] Zhu Guangju. Study and Practice on the Construction of Storage Network in Digital Library[J]. Journal of Modern Information, 2010,30(12):77-80.

[6] Wang Jiabing and Wang Chengyan. Exploration of Hyper-Converged Infrastructure in Library Data Center Construction [J]. Chinese Journal of ICT in Education, 2019(19):90-92.

[7] Meng Fanli, Sun Rong, Xu Ming. Research and Exploration in Laboratory[J]. Research and Exploration in Laboratory, 2012, 31(12):62-66, 71.DOI:10.3969/j.issn.1006-7167.2012.12.017.

[8] Li Yong, Yang Huafen. Virtualization construction of university data center based on vSphere[J]. Experimental Technology and Management, 2018, 35(01): 133-135. DOI: 10.16791 /j.cnki.sjg.2018.01.031.

[9] Liu Z. Liu Z. Research on the Path of Digital Transformation in Chinese Higher Education [J].Adult and Higher Education,2023,5(18):DOI:10.23977/ADUHE.2023.051814.

[10] Chen Yuli. Current Status, Challenges and Improvement Strategies of Educational Data Center Construction in China [J]. Straits Science, 2024(07):97-100.