ISSN: 2350-0328
International Journal of AdvancedResearch in Science, Engineering and Technology
Vol. 10, Issue 11, November 2023

Grid/cloud computing and IoT for medical and healthcare applications

Giorgos Giannakakis

Postdoctoral Research Associate, Institute of Computer Science, Foundation for Research and Technology Hellas, Heraklion, Greece

ABSTRACT: This article investigates the integration of grid and cloud computing technologies in medical and healthcare applications. Over the last years, there has been a significant advance in distributed processing power and scalability in addition to telecommunication advances and IoT technologies. These technologies are applied and improve many biomedical applications. The clinical and the research community benefit from distributed data processing and management, increased computational capabilities, enhanced diagnostic modelling, improved error correction and real-time data processing. The interoperability of health information systems is enhanced, ensuring secure and seamless data exchange. Besides, the integration of grid and cloud computing with emerging technologies like artificial intelligence (AI) and machine learning (ML) provides promising results for effective and enhanced capabilities in medical and healthcare applications.

KEY WORDS: grid computing, cloud computing, medical applications, healthcare applications, Machine Learning, AI.

I. INTRODUCTION

The field of medicine and healthcare is undergoing a notable transformation as a result of the latest technological breakthroughs, with the incorporation of grid computing taking the forefront in this revolution. The rapid and exponential growth of medical data, along with the need for distributed processing facilities, has brought about a time when conventional computing paradigms become inadequate. The demand for robust computational power and distributed processing skills has become essential as a result of the increasing complexity of medical data.

Grid computing emerged as a response to address the growing demand for scalable and distributed computational capacity and to overcome the limitations of traditional centralized computing systems in handling the complexities of current data-intensive applications. Its aim was to leverage the combined capabilities of a network of interconnected resources. The grid computing paradigm emerged in the late 20th century with the goal of creating a flexible and decentralized infrastructure. This architecture enabled the computer resources to be pooled together and operate as a unified system.

Grid computing divides computer operations into smaller, controllable parts that are dispersed over a network. It is characterized by its decentralization, which enables the execution of multiple tasks simultaneously through parallel processing. This aspect sets it apart from traditional computing models. Grid computing greatly enhances job execution speed by harnessing the processing capabilities of numerous machines in parallel, which would otherwise be challenging for a single system to manage.

Cloud computing, which has evolved from grid computing, offers the advantages of grid computing with expanded capabilities [1]. It introduced client-server computing architecture, higher scalability and flexibility in relation to grid computing, service-oriented approach and accessibility to computing resources through standard web protocols.

The continuous evolution of healthcare applications led to the integration of the Internet of Things (IoT), which brought a new era of connectivity and data streaming. IoT devices, ranging from wearable health trackers to smart medical instruments, have become integral components of patient care. These devices generate real-time data, providing healthcare professionals with continuous insights into patient health and treatment efficacy. In the context of medical IoT, cloud computing serves as the backbone that processes and analyzes the vast streams of data generated by these devices. Wearable devices, for example, can continuously monitor vital signs, and the data they generate can be securely transmitted to cloud platforms for

ISSN: 2350-0328
International Journal of AdvancedResearch in Science, Engineering and Technology
Vol. 10, Issue 11, November 2023

analysis. This real-time connectivity facilitates proactive healthcare interventions, timely diagnostics, and personalized treatment plans.

The evolution of computing architectures is presented in Fig.1.

Fig.1: Evolution of computing architectures

This paper investigates the role of grid, cloud and IoT computing architectures in addressing the evolving needs of healthcare applications. It specifies the evolving needs, it describes how these computing architectures address these needs and provides a list of the highlighted technologies in medical and healthcare applications.

II. LITERATURE SURVEY

In this section, the healthcare section needs/requirements for grid computing and the healthcare/medical applications utilizing grid/cloud computing are investigated according to the relevant literature.

A. Medical applications needs/requirements in the healthcare section

The medical computational applications in healthcare have demonstrated their beneficial impact on patients and healthcare professionals [1]. Their requirements become even more demanding in recent years to address the volume/complexity of medical data as well as to meet the standards for high-quality and consistent health services. We have identified the following significant medical applications needs/requirements that grid computing addresses as follows.

Parallel Processing: Grid computing allows for the parallel and simultaneous processing of algorithms in different datasets reducing the computational time. During clinical trials and epidemiological investigations, a substantial volume of data frequently requires processing and analysis. Grid computing enables the distribution of this task among several computers, resulting in an acceleration of the whole process.

Distributed Computing: Grid computing facilitates the dispersion of computer tasks among various geographically dispersed locations. This is especially advantageous in extensive investigations where data is gathered from multiple locations.

Efficiency: Using grid's combined computational capabilities, researchers can greatly diminish the duration needed for data processing and analysis. Timely results are vital for decision-making in clinical trials and epidemiological investigations.

Decentralization: Grid computing can enhance data security and privacy by distributing the processing of sensitive information over multiple nodes. Ensuring patient data confidentiality is of utmost importance, particularly in healthcare research.

Scalability (adaptability to workload): Grid computing allows the system to adjust and accommodate variations in workload. It has the ability to deploy resources in a flexible manner to meet the changing demands of a clinical trial or epidemiological investigation.

Interoperability: Grid computing can enhance the interoperability of health information systems, facilitating the smooth interchange of patient data among various healthcare providers and systems.

Secure Data Sharing: Grid computing can be utilized with strong security protocols to facilitate the safe exchange of confidential healthcare data among authorized entities.

Integration with Machine Learning and AI algorithms: As technology continues to advance, the integration of grid computing with emerging technologies like artificial intelligence (AI) and machine learning (ML) is likely to further enhance its capabilities in medical and healthcare applications.

Meeting these requirements is essential to create a robust and compliant grid computing infrastructure in the healthcare domain.

B. Applications of grid/cloud computing in the healthcare sector

eHealth involves the utilization of Information and Communication Technologies (ICT) to create a smart environments for the management and monitoring of individuals' health conditions [2]. Grid computing has been applied effectively in the medical and healthcare domains, leading to significant progress in research, diagnosis, treatment and health services management. Healthgrids is the term used to describe grid infrastructures that consist of applications, services, or middleware components designed to address the processing of biomedical data utilizing resources such as medical devices, databases, computational power and medical expertise [2]. Healthgrids have been applied effectively mainly in the following domains:

Biomedical Research and Simulations: Biomedical research aims to model human bodily functions utilizing complex models (in vivo, in vitro models) and to simulate human biological processes. Due to this complexity which a common task especially in the era of precision medicine substantial computing capabilities are required that can be addressed by grid/cloud architectures as it is referred in the related literature [3, 4].
Over the latest years, there is a growing research utilizing wearable body sensors recording vital biosignals (eg ECG, EEG, EMG, breath pattern. blood pressure, temperature, body activity etc) which through Internet of Things (IoT) or Internet of m-Health Things (mIoT) [5] services connect to grid or cloud computing services to analyze recordings and support health-related applications.

Genomic Research: Grid computing is utilized for extensive sequence analysis derived from genomic databases, enabling researchers to investigate genetic variants and comprehend disorders at a molecular scale.

Medical Imaging Processing: Grid/cloud computing has been utilized widely in the area of medical imaging to process and analyze medical images of different modalities (such as MRI, CT, SPECT, X-rays, etc). The grid infrastructure enables multilevel services such as the data storage across multiple nodes in the grid, the parallel processing during algorithm execution [6] (eg heavy computational tasks such as Medical Image 3D Reconstruction), the scalability of the computing infrastructure, the security of the sensitive medical data, the sharing of medical data [7] and the creation of create virtual machines for specific computational tasks.

Clinical Trials and Epidemiological Studies: Grid computing facilitates the management and analysis of large datasets derived from clinical trials, empowering researchers to discern patterns, associations, and prospective therapeutic results.
Extensive epidemiological investigations, which entail evaluating data from several sources, can take advantage of the distributed computational capabilities of grids.

Drug Discovery: Grid computing facilitates the simulation and analysis of drug interactions, helping in the discovery of potential pharmaceutical compounds and accelerating the drug development process.

The studies of grid/cloud computing for medical applications along with technologies used, healthcare domain and contribution are summarized in Table 1.

ISSN: 2350-0328
International Journal of AdvancedResearch in Science, Engineering and Technology
Vol. 10, Issue 11, November 2023

Table 1: Overview of medical grid/cloud computing applications studies along with technologies used, healthcare domain and contribution

Study	Technology	Domain	Contribution
Ramos et al. [4]	Grid computing	Biomedical Research and Simulations	Software framework incorporating Machine Learning classifiers.
Arbona et al. [8]	Grid computing	Biomedical Research and Simulations	Advanced computational and data services
Kumar et al. [9]	Grid computing	Medical Imaging Processing	Good performance in very large datasets, advantage of nodes over high-bandwidth WAN
Huang et al. [6]	Grid Computing	Medical Imaging Processing	It does not require to significant rewrite the image processing codes, computational effectiveness
Lebre et al. [7]	Cloud Computing	Medical Imaging Sharing	Proposes a shared medical imaging repository and a secure accounting mechanism that can be integrated with modern DICOMWeb services for storage, search, and retrieve medical imaging data
Sinnott et al. [10]	Grid Computing	Clinical Trials	Accessing and using clinical datasets in a flexible, dynamic and secure manner
Kyriazis et al. [11]	Grid Computing	Clinical Trials	Yields very good results for the conducted virtual trial in agreement with the outcome of the real clinical study
Wegener et al. [12]	Grid Computing	Clinical Trials	Provides integration of R in a grid architecture executing complex analysis scenaria
Wegener et al. [13]	Grid Computing	Clinical Trials	Provides an update to the ACGT GridR environment that enables the parallelization of R scripts loops for their distributed execution on a computing grid
Singh et al. [14]	Grid Computing	Genomic Research	Achieves significant speedups on grid prototype, easily programmable mass-produced graphics hardware
Hemandezl et al. [15]	Grid Computing	Genomic Research	Reducing the computational time of the Blast process for thousands of sequences, increased performance
Nguyen et al. [16]	Cloud Computing	Genomic Research	CloudAligner achieves good performance, covers most primary features, is more accurate in next generation sequencing (NGS) analysis.
Langmead et al. [17]	Cloud Computing	Genomic Research	Parallelizes efficiently sequence alignment and genotyping algorithms. Allows large datasets of DNA sequences to be analyzed rapidly without sacrificing accuracy or requiring extensive software engineering efforts
Matsunaga et al. [18]	Cloud Computing	Genomic Research	CloudBlast performs efficiently massive sequence alignment
Langmead et al. [19]	Cloud Computing	Genomic Research	Myrna, cloud-computing pipeline for calculating differential gene expression in large RNA-Seq datasets
Habegger et al. [20]	Cloud Computing	Genomic Research	VAT, functionally annotate variants from multiple personal genomes, provides a novel means of clearly visualizing the functional impact of variants across different transcript isoforms of a given gene
Kim et al. [21]	Grid Computing	Drug Discovery	DrugScreener-G can hide technical details of Grid computing from users and also support multiple Grid middleware

C. Computational grid fault tolerance

Grid/cloud computing systems, despite their overall advantages, are prone to faults that have a negative impact on their performance [22] and reliability [23]. Failures or error conditions include hardware failures (power failure, host crash, etc), software errors (memory leak, numerical exceptions, etc) and other types of failures (network congestion) [24]. Fault tolerance in grid systems includes the fault detection/identification and the development of recovery techniques to restore normal operation and enable uninterrupted calculations. The recovery techniques can be divided into task-level and workflow-level techniques [24].

We will focus on the task-level techniques, including checkpoint, grid resources replication, tasks rescheduling and data replication [23]. Checkpointing is the procedure of periodically saving the state of a running process, allowing a failing process to be restarted from its checkpoint [23]. Checkpoint procedures have been utilized in many grid computing applications [25-27]. Grid resource replication involves the simultaneous execution of an identical calculation and the maintenance of an identical state by numerous grid resources [23]. The objective of replication is to guarantee that there is always at least one replica that will complete the calculation in case all the others fail.

Another fault recovery technique is the dynamic rescheduling of a failed task. Huedo et al. [28], proposed an adaptive framework that monitors efficient task execution and adapts resources for better performance. In [29], a low-cost rescheduling policy was presented, providing good performance results. In [30], a novel adaptive rescheduling algorithm was proposed allowing for a collaborative workflow planner. Data replication was also applied as a fault recovery technique in grid systems [31]. In [32], a data replication method based on economic models was proposed, optimizing both the selection of replicas and the dynamic creation of replicas. In [33], a replica catalog algorithm was introduced that automatically detects simultaneous access to the replica by multiple nodes which was implemented as XML-based replica location service.

III. CONCLUSION AND FUTURE WORK

In this paper, we investigated the requirements of the medical applications that are addressed by the grid computing systems and we perform a literature review with grid/cloud computing applications in healthcare sector.

The reviewed literature demonstrates the versatility of grid computing, offering computational efficiency for the processing and analysis of large datasets, facilitating collaborative research endeavors, and enhancing the efficiency of healthcare systems.

In conclusion, this research paper has explored the complex domain of grid computing and its significant implications for medical and healthcare applications. The profound impact of grid computing in this field is clearly demonstrated by its capacity to meet the increasing requirements of data-intensive activities, improve computational effectiveness, and promote collaborative research efforts.

REFERENCES

[1] S.L. van der Storm, M. Jansen, H.A.W. Meijer *et al.*, “Apps in healthcare and medical research; European legislation and practical tips every healthcare provider should know”, *International Journal of Medical Informatics,* vol. 177, pp. 105141, 2023/09/01/, 2023.
[2] V. Breton, K. Dean, and T. Solomonides, "The Healthgrid white paper." pp. 249-318.
[3] S. Shahand, M. Santcroos, A.H.C. van Kampen *et al.*, “A Grid-Enabled Gateway for Biomedical Data Analysis”, *Journal of Grid Computing,* vol. 10, no. 4, pp. 725-742, 2012/12/01, 2012.
[4] R. Ramos-Pollán, M.Á. Guevara-López, and E. Oliveira, “A Software Framework for Building Biomedical Machine Learning Classifiers through Grid Computing Resources”, *Journal of Medical Systems,* vol. 36, no. 4, pp. 2245-2257, 2012/08/01, 2012.
[5] L.M. Dang, M.J. Piran, D. Han *et al.*, “A survey on internet of things and cloud computing for healthcare”, *Electronics,* vol. 8, no. 7, pp. 768, 2019.
[6] H.K. Huang, A. Zhang, B. Liu *et al.*, "Data grid for large-scale medical image archive and analysis." pp. 1005-1013.
[7] R. Lebre, L.B. Silva, and C. Costa, “A Cloud-Ready Architecture for Shared Medical Imaging Repository”, *Journal of Digital Imaging,* vol. 33, no. 6, pp. 1487-1498, 2020/12/01, 2020.
[8] A. Arbona, S. Benkner, G. Engelbrecht *et al.*, “A Service-Oriented Grid Infrastructure for Biomedical Data and Compute Services”, *IEEE Transactions on NanoBioscience,* vol. 6, no. 2, pp. 136-141, 2007.
[9] V.S. Kumar, B. Rutt, T. Kurc *et al.*, “Large-Scale Biomedical Image Analysis in Grid Environments”, *IEEE Transactions on Information Technology in Biomedicine,* vol. 12, no. 2, pp. 154-161, 2008.
[10] R.O. Sinnott, A. Stell, and O. Ajayi, “Supporting grid-based clinical trials in Scotland”, *Health Informatics Journal,* vol. 14, no. 2, pp. 79-93, 2008.
[11] D. Kyriazis, A. Menychtas, D. Dionysiou *et al.*, "Clinical trial simulation in Grid environments." pp. 1-6.

[12] D. Wegener, T. Sengstag, S. Sfakianakis *et al.*, "GridR: An R-Based Grid-Enabled Tool for Data Analysis in ACGT Clinico-Genomics Trials." pp. 228-235.
[13] D. Wegener, T. Sengstag, S. Sfakianakis *et al.*, "Supporting Parallel R Code in Clinical Trials: A Grid-Based Approach." pp. 823-828.
[14] A. Singh, C. Chen, W. Liu *et al.*, "A Hybrid Computational Grid Architecture for Comparative Genomics", *IEEE Transactions on Information Technology in Biomedicine*, vol. 12, no. 2, pp. 218-225, 2008.
[15] V. Hemandezl, M. Robles, and M. Talon, "Blast2GO goes grid: developing a grid-enabled prototype for functional genomics analysis."
[16] T. Nguyen, W. Shi, and D. Ruden, "CloudAligner: a fast and full-featured MapReduce based tool for sequence mapping", *BMC research notes*, vol. 4, no. 1, pp. 1-7, 2011.
[17] B. Langmead, M.C. Schatz, J. Lin *et al.*, "Searching for SNPs with cloud computing", *Genome biology*, vol. 10, pp. 1-10, 2009.
[18] A. Matsunaga, M. Tsugawa, and J. Fortes, "Cloudblast: Combining mapreduce and virtualization on distributed resources for bioinformatics applications." pp. 222-229.
[19] B. Langmead, K.D. Hansen, and J.T. Leek, "Cloud-scale RNA-sequencing differential expression analysis with Myrna", *Genome biology*, vol. 11, pp. 1-11, 2010.
[20] L. Habegger, S. Balasubramanian, D.Z. Chen *et al.*, "VAT: a computational framework to functionally annotate variants in personal genomes within a cloud-computing environment", *Bioinformatics*, vol. 28, no. 17, pp. 2267-2269, 2012.
[21] J.B. Kim, N.D. Nhan, S. Lee *et al.*, "DrugScreener-G: Towards an Integrated Environment for Grid-Enabled Large-Scale Virtual Screening and Drug Discovery." pp. 666-671.
[22] R. Medeiros, W. Cirne, F. Brasileiro *et al.*, "Faults in grids: why are they so bad and what can be done about it?." pp. 18-24.
[23] C. Dabrowski, "Reliability in grid computing systems", *Concurrency and Computation: Practice and Experience*, vol. 21, no. 8, pp. 927-959, 2009.
[24] S. Hwang, and C. Kesselman, "Grid workflow: a flexible failure handling framework for the grid." pp. 126-137.
[25] M. Singh, "Incremental checkpoint based failure-aware scheduling algorithm in grid computing." pp. 772-778.
[26] S. Al-Kiswany, M. Ripeanu, S.S. Vazhkudai *et al.*, "stdchk: A checkpoint storage system for desktop grid computing." pp. 613-624.
[27] S. Malik, B. Nazir, K. Qureshi *et al.*, "A reliable checkpoint storage strategy for grid", *Computing*, vol. 95, pp. 611-632, 2013.
[28] E. Huedo, R.S. Montero, and I.M. Llorente, "A framework for adaptive execution in grids", *Software: Practice and Experience*, vol. 34, no. 7, pp. 631-651, 2004.
[29] H. Zhao, and R. Sakellariou, "A low-cost rescheduling policy for dependent tasks on grid computing systems." pp. 21-31.
[30] Z. Yu, and W. Shi, "An adaptive rescheduling strategy for grid workflow applications." pp. 1-8.
[31] H. Stockinger, A. Samar, K. Holtman *et al.*, "File and object replication in data grids", *Cluster Computing*, vol. 5, pp. 305-314, 2002.
[32] W.H. Bell, D.G. Cameron, R. Carvajal-Schiaffino *et al.*, "Evaluation of an economy-based file replication strategy for a data grid." pp. 661-668.
[33] S.i. Takizawa, Y. Takamiya, H. Nakada *et al.*, "A scalable multi-replication framework for data grid." pp. 310-315.