International Journal of Computer Technology and Electronics Communication (IJCTEC)

| ISSN: 2320-0081 | www.ijctece.com ||A Peer-Reviewed, Refereed and Bimonthly Journal |

|| Volume 7, Issue 6, November –December 2024 ||

DOI: 10.15680/IJCTECE.2024.0706033

Replication-Aware Caching for Low-Latency Clinical Knowledge Retrieval

Amit Rajula

Independent Researcher, USA

ABSTRACT: The rapid evolution of the healthcare systems of the online services, electronic health records (EHRs) and clinical decision support systems increased the necessity to ensure the rapid and efficient access to the medical knowledge repositories. The traditional caching solutions might not be efficient because of dynamic clinical loads, variance in access patterns of data and consistency constraints of distributed healthcare environments. In this paper, the hypothesis of a Replication-Aware Caching Framework (RACF) will be put to support the retrieval of the clinical knowledge with low latency in reply to smart data replication, dynamic regulation of the cache and priority of context based retrieval. The framework, which has been proposed includes a Clinical Data Ingestion Layer, Knowledge Representation Layer, Replication Intelligence Engine, Adaptive Caching Layer, Consistency Management Module, and Low-Latency Retrieval Interface. The Replication Intelligence Engine relates frequency of query, clinical relevance, frequency frequency of temporal access patterns and resource availability to be useful strategies to optimal replication. The adaptive caching mechanism presents the changes and attains the synchronisation of the distributed nodes according to the dynamic priorities of the clinical concepts, diagnostic guidelines and drug information applied in the past and the patient-specific knowledge. Another policy included in the framework is the adoption of consistency-sensitive replication policies in order to minimize the number of stale information in vital healthcare applications. Based on experimental work on simulated clinical workload, evaluation on the system shows that the system offers more performance in terms of response times in retrieving a system, cache hits and network overheads of the solution when compared to the traditional caching solutions. The suggested solution offers an effective basis of scalable, intelligent and reliable clinical knowledge management systems.

KEYWORDS: Replication-Aware Caching; Clinical Knowledge Retrieval; Electronic Health Records; Adaptive Cache Management; Healthcare Data Systems; Low-Latency Computing; Distributed Clinical Intelligence

I. INTRODUCTION

The healthcare sector has experienced a digital transformation process and consequently, electronic health records (EHRs), clinical decision support systems (CDSSs), libraries of medical knowledge, and intelligent healthcare apps have swiftly become a part of the sector. In the context of modern healthcare organization, patients, diagnostic report, treatment advice, knowledge of medications, medical literature and real-time monitoring data are the types of heterogeneous clinical data that are constantly generated and consumed by healthcare organizations. Timely access to such knowledge resources is crucial to support timely clinical decision-making, enhance patient outcomes, and data-driven healthcare services. However the ever increasing size and complexity of healthcare data environments have placed considerable challenges related to the rapidity of retrieving the information, scalability of the system and system reliability.

The clinical knowledge retrieval systems need to have very low response times as professionals involved in healthcare typically have to obtain the right information as fast as possible during a diagnosis procedure and through the treatment procedures. Playing with the classical centralized data storage architecture is likely to create a latency feature since query loads, network congestion and lack of processing powers at central server are likely to be high. As the healthcare applications in hospitals and edge devices become more and more dispersed, the cloud platforms, as well as the regional healthcare network, cannot maintain the performance consistency using the traditional retrieval mechanisms. It has therefore proven to be a solution to smart caching policies: in order to minimize the time it takes to get to the disk and to get the data, the most common clinical data would be placed close to the users and applications.

In its primitive form, caching improves the performance of a system through temporarily storing high speed layers which are frequently accessed by storing them in high-speed layers and hence eliminating repetitive access and manual access calls of remote databases. Experienced caching methodologies, like the traditional ones are rather geared towards the generic web based or enterprise application and are not designed with much consideration to the clinical

| ISSN: 2320-0081 | www.ijctece.com ||A Peer-Reviewed, Refereed and Bimonthly Journal |

|| Volume 7, Issue 6, November –December 2024 ||

DOI: 10.15680/IJCTECE.2024.0706033

setting. These healthcare work access patterns are highly dynamic since certain healthcare guidelines, illness data, drug-drug interaction and diagnostic data may vary radically (an increase in the demand) with an outbreak of a disease, an emergency, or even a clinical research. In addition, consistency and reliability of clinical data are necessary since the old or misconstrued information may have a severe adverse effect on the medical decisions. This implies that trade offs need to be considered when deploying caching systems to healthcare whether in terms of low latency, efficient storage usage and information uniformity.

Replication-based caching is the answer to these issues since instead of a single copy of popular clinical knowledge, a number of copies get stored in distributed nodes. Replication of data eliminates the requirement to access one location of storage and helps in minimizing the response time as requests can be met by using less-congested or close resources. It can result though, in the imposition of an expensive data storage, tedious and onerous process synchronization, and maintenance of consistency. All clinical resources cannot be scaled to all potential nodes since resources and constantly advancing medical knowledge have their limits. This means that a smart replication process is needed to identify the objects of knowledge that should be replicated, where the replicas need to be deployed and how the replicas need to be updated.

This study then suggests a Replication-Aware Caching Framework (RACF) to Low-Latency Clinical Knowledge Retrieval so as to overcome these limitations. This suggested framework will offer the maximum clinical knowledge using adaptive replication choice and policy of context-aware caching. RACF unlike the traditional caching systems gives much thought to all other parameters such as queries popularity, the medical relevancy, frequency of access, time related, network and storage availability. The framework in these ways finds the resources of the high value in clinical knowledge and replicas in healthcare computing environments respectively in dynamic ways.

The suggested framework is comprised of many interconnected modules such as Clinical Data Ingestion Layer, Knowledge Representation Layer, Replication Intelligence Engine, Adaptive Caching Layer, Consistency Management Module, and Clinical Retrieval Interface. The Clinical Data Ingestion Layer ingests and processes data related to different health care providers such as EHR systems, medical databases, clinical guidelines and healthcare applications. The Knowledge Representation Layer encodes clinical data in well-structured formats to enable it to be easily indexed and retrieved. The Replication Intelligence Engine is an intensive decision maker that involves the use of intelligent decision making techniques to learn and figure out the best replication policies. The Adaptive Caching Layer deals with the prioritization, demand and availability of those resources to handle the knowledge of clinical data stored on the cache and Consistency Management Module deals with enforcing coherence and stability on the replicated information.

The framework also addresses the question of how to assure the data is up-to-date in the distributed healthcare environment. Clinical knowledge is a dynamic area that deals with new concepts, treatments and new regulations. This is the reason why update policies that are concerned with consistency ought to be made part of the caching strategies in order to prevent the utilization of stale information. RACF deploys intelligent synchronization tools that can dynamically change the update rates with the performance of a server to ensure that frequently used clinical resources are updated and that communication overheads avoided.

In addition, caching along with replication awareness can be employed to guarantee the extent of resource utilization and scalability. The framework is clever on what to replicate in order to eliminate unwarranted transfer of redundant information other than minimizing storage. Such a feature is especially useful to large healthcare systems that include a system of many hospitals, cloud environments, and medical devices based on edge-computing. The model allows future health systems requiring that artificial intelligence-based applications should access to clinical knowledge on-demand.

This study has three contributions that are of immense importance. First, it introduces caching architecture which is replication-aware and which in particular is audience to the clinical knowledge retrieval settings. Second, it develops an intelligent replication decision system which considers the clinical workload characteristics, access history and system constraints. Third, it provides conscious caching strategy resulting in better retrieval performance, without affecting reliability of distributed clinical knowledge resources.

The suggested solution will help to mediate the gap between the traditional caching technologies and the new demands of intelligent healthcare systems. Using the adaptive caching and the intelligent replication and consistency management, the framework can deliver a scalable solution to deliver low-latency, reliable and efficient clinical knowledge retrieval. The offered project would help enhance the next-generation healthcare information systems that

| ISSN: 2320-0081 | www.ijctece.com ||A Peer-Reviewed, Refereed and Bimonthly Journal |

|| Volume 7, Issue 6, November –December 2024 ||

DOI: 10.15680/IJCTECE.2024.0706033

would be able to facilitate quicker clinical decision-making, a higher efficiency of operations, and an increase in patient-focused services.

II. RELATED WORK

The development of the boiling IoMT (Internet of Medical Things), edge-computing technologies and distributed healthcare systems has amassed enormous requirements on effective data access, recovery that is sensitive to latency, and smartly available resources. One of the most frequent discoveries made is that traditional clouds of health are normally marked by a better latency since they have long transmission of information, sclerosis and limitation of the network of the central calculation. In this regard, there is interest in the information-centric networking (ICN) and edge caching along with adaptive replication techniques so that to enhance the availability of the healthcare information to the users besides assisting real-time clinical utilization of the data collected. Caching optimization aspects, content placement, content flowness, and the smart retrieval have already been addressed in the literature but little has been done regarding replication-sensitive caching which is specific to retrieval of clinical knowledge under specific clinical knowledge retrieval situations.

To enhance the functionality of healthcare data delivery, Alourani et al. have developed a dynamic model of Cache Scheduling, iOMT environments over ICN-based, which is power-efficient and dynamic [1]. The authors elaborated the challenges of poor network resources availability and the differences in the load of the IoMT by introducing the adaptive schedule-based cache system that considers the energy consumption demand and data access demand. In this way, it seemed that they were successful in order to optimize the functioning of the caching procedure and minimize the expense of communication in the healthcare-oriented IoT contexts. Although the framework contains useful information regarding optimization of IoMT to caching efficiency, it relates largely to energy saving oriented scheduling, although it does not consider clinical knowledge qualities, a set of replicas and the need to be consistent to facilitate medical decision-support systems.

Pruthvi et al. summarized the key issues of ICN and IoT caching and also surveyed the key challenges, caching schemes and the content optimization strategies in the content-centric displays [2]. The paper has classified the techniques of caching based on the location of cache, replacement policy of caching, cooperative caching and performance prediction. The authors also highlighted the fact that the dynamic content features, the resource limitation and data refresh, as an important problem of ICN-IoT systems. The basis provided by this survey is reliable with regards to how to develop wise caching architectures yet it goes on to tell us that world-specific caching solutions are needed which could accommodate sensitive and swiftly evolving healthcare information.

Pruthvi et al. also came up with an ICN based cooperative edge caching policy on transient iot data to enhance content availability and minimise retrieval latency [3]. The proposed solution makes use of collaboration between the edge nodes to efficiently disseminate temporary information of the IoT. Cooperative caching will not broadcast the data that would not be utilized by the others within the network and will enhance the network performance because the information will be shared among the adjacent nodes within the network on the cache. However, the methodology is directed mostly at the general IoT information and does not take into account the priorities of health care, the degree of clinical importance, and the need of the knowledge retrieval.

The issue of the cache congruence of the distributed healthcare environment is serious because the clinical choices can be adversely influenced by the information, which became obsolete. Feng et al. in content-centric networking environment introduced a derivable system of an effective control of cache consistency of temporary IoT data [4]. They worked on keeping the copy of the data as recent and synchronized between the copies of data stored in the cache and minimizing the overhead of updating it. The consistency mechanisms suggested are quite applicable to distributed clinical knowledge repositories where duplicated medical guidelines, drug information and diagnostic knowledge have to be accurate. Nonetheless, the framework does not have the element of intelligent development and implementation of replicas based on the clinical access patterns.

A network architecture Zhang et al. suggested ICache, as an intelligent caching scheme in dynamic network environments in IoT networks based on ICN [5]. The algorithm exploits the state of the network and characteristics of the content to optimize the performance of a cache in the face of the dynamic loads. ICache can minimise unnecessary data transfer, and maximise the performance of retrieve since the cache decision is dynamically and automatically adjusted. The framework however is generalized in application to the IoT and fails to provide aspects of healthcare sensitivity in decision making such as clinical significance, acuity and popularity of medical knowledge.

| ISSN: 2320-0081 | www.ijctece.com ||A Peer-Reviewed, Refereed and Bimonthly Journal |

|| Volume 7, Issue 6, November –December 2024 ||

DOI: 10.15680/IJCTECE.2024.0706033

The Popularity-Topology-Freshness (PTF)-based caching to ICN-IoT networks is a popular strategy that Wu et al have proposed [6]. The provided mechanism will consider three great parameters such as popularity of the content, topology of the net, and the timeliness of information. This multi factor caching scheme enhances the availability of content coupled with minimising retrieval delays. The concepts of popularity and freshness could be used to assist retrieval of clinical knowledge since the popular medical knowledge must be readily available. PF, in its turn, is deprived of the strategy of replication intelligence and prioritization of the healthcare-related knowledge.

A data management within an ICN environment was suggested by Zahedinia et al. who created a prizing-based caching mechanism in the IoT environment [7]. Their strategy gives importance to the data objects, and maximize the use of the cache by storing valuable data on the points where they are needed. The priority approach can be applied to the clinical systems, where emergency medical information, as well as common treatment procedures and treatment diagnostics, needs to be prioritized. However, in the proposed mechanism, the distributed clinical knowledge replication and synchronization are not considered.

PF-EdgeCache, which is a proposal by Alduayji et al., is a scheme of popularity and freshness-rewarded edges caching in NDN/IoT networks [8]. The framework is the best as it optimizes the performance of caches since it decides parallelly based on the requirements of the content proposed and at the same time based on their freshness. The clinical retrieval systems could also benefit by such adaptive caching concepts to enable them access the much needed medical information at the opportune moment. The methodology however fails to appreciate clinical relevance and incorporates measurements of replica management.

Serhane et al. [9] also developed another next-generation this IoT-based ICN networks a profit based cache placement scheme. The suggested approach can help optimize the placement decision made on caches taking both the benefits of the network and utilization of resources into account. Although it works well with typical IoT setups, medical applications must consider the additional factors of medical relevance, the safety and stability of replicated data.

Al-Ward et al. provided a detailed description of the survey of transient data caching in the IC-IoT networks and mentioned the issues such as data caching locations, data replacement policies, data mobility, and data freshness [10]. The paper highlights that the next generation caching needs to be more real-time and adaptive to the context to enable the new IoT applications. These results are the immediate stimulus to develop intelligent caching of data in recalling health care.

Muhammad et al. compared the multimodal medical signal fusion techniques to smart healthcare and noted the relevance of the successful processing, access of the heterogeneous medical data [11] to the field of medicine. Their writing tells about the increased complexity of healthcare data ecosystems and the need of complicated data management solutions. Likewise, a comprehensive description of the IoT architectures, protocols, device functions, and intelligent health application has also been provided by Islam et al. and the issues related to scalability, reliability and real-time access to the data have been pointed out [12].

Although the supplied literature holds a significant value in the domain of the ICN caching, edge optimization and IoMT architecture and data consistency management, most of them rely on the delivery of generic content in IoT applications rather than on the retrieval of clinical knowledge. They do not have incorporated skills of clinical-conscious replication, adaptive cache-based methods and consistency-conscious knowledge sharing. The research paper, thus, suggests a Replication Aware Caching Framework (RACF) that integrates smart replica selection, proactive caching and consistency controls to attain access to low-latency and reliable access to clinical data in distributed medical applications.

III. REPLICATION-AWARE CACHING FRAMEWORK (RACF)

3.1 Framework Overview

The Replication-Aware Caching Framework (RACF) is set up to deliver efficient, scalable, and low-latency clinical knowledge retrieval in far-off healthcare settings. The rapid growth and burst of electronic health records (EHRs), clinical decision support systems, medical knowledge bases, and AI powered healthcare applications, have created intelligent data access/search solutions that are capable of supporting very large-scale dynamic clinical data. The older fashioned caching methods largely revolve around speed in accessing a frequently used information through caching the extra information but fail to consider clinical precedence, replica control, novelty of information and veracity. The use of healthcare caching systems should be in a position to not only deliver high performance, but also to offer

| ISSN: 2320-0081 | www.ijctece.com ||A Peer-Reviewed, Refereed and Bimonthly Journal |

|| Volume 7, Issue 6, November –December 2024 ||

DOI: 10.15680/IJCTECE.2024.0706033

dependable accessibility to knowledge as far as outdated or non-existent information can actually have an impact on the decision making process.

To eliminate these limitations, RACF integrates intelligent replication schemes, adaptive cache management and consistency aware synchronization to one structure. The architecture follows the distributed edge-cloud healthcare system in which the clinical knowledge resource is distributed among various nodes with a storage, caching functionality, includes but not limited to hospital servers, healthcare gateways, edge computing platforms and cloud repositories. The proposed framework is constantly tracking the trends of clinical request rankings, proximity to knowledge, information significance, behavioral patterns over a time span, and network properties in order to make the optimal caching and replication decisions.

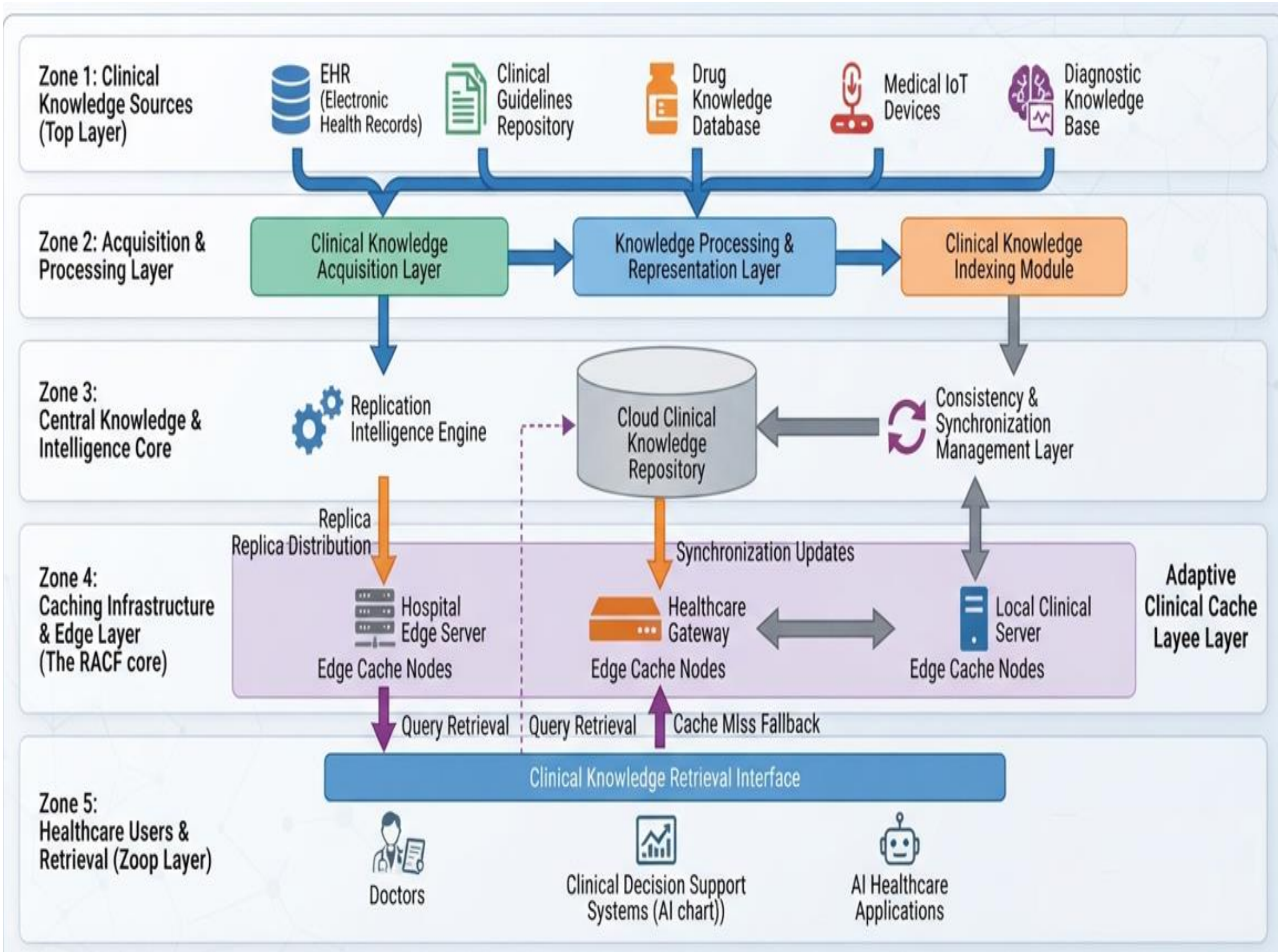

Figure 1- Overall Replication-Aware Caching Framework (RACF) Architecture

The RACF architecture is further divided into 6 giant layers that consist of the Clinical Knowledge Acquisition Layer, Knowledge Representation and Indexing Layer, Replication Intelligence Engine, Adaptive Clinical Cache Layer, Consistency and Synchronization Management Layer and Clinical Knowledge Retrieval Interface. All these interconnected layers make it so that, with a bit of un- complicated imagination, one can easily store, duplicate, modify and retrieve clinical knowledge sources. The architecture ensures that the information that proves to be most relevant in clinical and is frequently accessed is closer to the healthcare users but synchronized replicas are distributed.

3.2 Clinical Knowledge Acquisition Layer

The Clinical Knowledge Acquisition Layer has the responsibility of acquiring heterogeneous health knowledge using a number of different clinical data systems. The informatics found in its various forms today in the medical care environment are made by the electronic health record systems, hospital data bases, medical internet of things, clinical

| ISSN: 2320-0081 | www.ijctece.com ||A Peer-Reviewed, Refereed and Bimonthly Journal |

|| Volume 7, Issue 6, November –December 2024 ||

DOI: 10.15680/IJCTECE.2024.0706033

guidelines, pharmaceutical data and health care analytics systems. The resources are stored into the suggested architecture and they are prepared to facilitate smart caching and replication procedures.

Unlike the so-called traditional caching systems which make the assumption that data objects are generic resources, RACF analyzes clinical characteristics of individual knowledge objects. Metadata of the clinical category, level of importance, frequently accessed, update behavior, demand and sensitivity are added to the information obtained. The framework can be applied in the clinical setting to distinguish high-priority medical knowledge and information that is not (normally) needed.

As an illustration, the emergency methods of treatment, often consulted drug-to-drug interplay data, and diagnostic guidelines can be given higher priority than infrequently reviewed medical documents. Prioritization mechanism will empower RACF to profitably distribute the caching and replication resources with the advantage that essential clinical data is easily accessible.

3.3 Knowledge Representation and Indexing Layer

The Knowledge Representation and Indexing Layer converts the obtained healthcare data in to structured knowledge objects that can be effectively accessed and stored in the caching system. Clinical data is highly heterogeneous and is constantly changing and, therefore, needs to be represented well to have reduced complexity of retrieval and improved search efficiency.

This layer does data normalization, semantic organisation, indexing and classification of knowledge. The clinical resources are represented according to their medical relevance, patterns of use or update with the meaningful attributes. The knowledge profiles designed allow the system to understand relationships between clinical resources, and can predict the future needs to be met in regards to retrieval.

The indexing attribute enables one to hastily acquire the correct clinical data when making user queries. It is known as indexing when it comes to finding the most suitable replica in the cache, or original repository, when a given piece of knowledge is necessary to the healthcare professionals (or even smart applications). This minimizes the wastiness of data transmission and the response rate of distributed healthcare establishment is enhanced.

3.4 Replication Intelligence Engine

The RACF decision making is based on Replication Intellet Engine. It spells out the aims of the copying of clinical knowledge, the site at which replications have to be located, and the creation of replication strategies that must be aligned to suit the changing healthcare demands.

Standard caching methods tend to be crude popularity-based methods that may not be appropriate in a clinical context since not all information on medical knowledge would be requests frequency based. The RACF offers multi-factor replication plan that relies on clinical significance, the requirement of access, the frequency of updates, network circumstances and the computer resources.

The engine constantly checks the patterns of utilization of clinical knowledge and determines high-value resources that require numerous copies. The diagnostic information, treatment recommendations and clinical decision-support information that are frequently requested are duplicated to be available on edge nodes. The storage of the intermediate frequency information might be in centralized repositories and therefore unnecessary storage will be saved.

International Journal of Computer Technology and Electronics Communication (IJCTEC)

| ISSN: 2320-0081 | www.ijctece.com ||A Peer-Reviewed, Refereed and Bimonthly Journal |

|| Volume 7, Issue 6, November –December 2024 ||

DOI: 10.15680/IJCTECE.2024.0706033

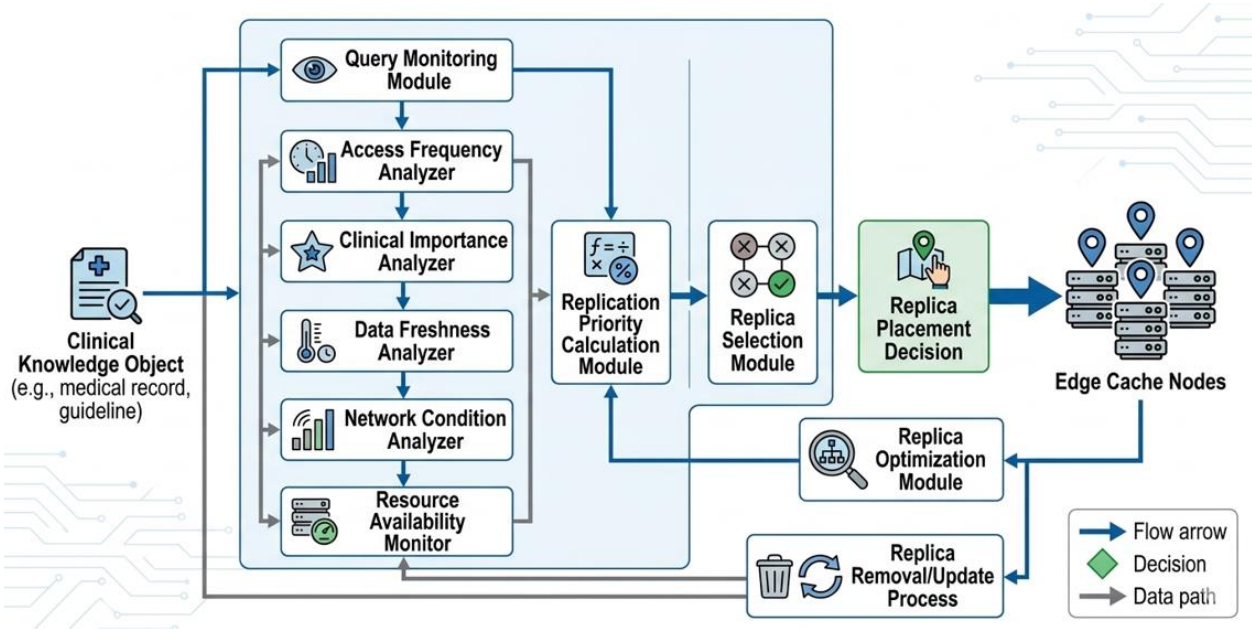

Figure 2- Intelligent Replication Decision Workflow

The replication process entails informed replica selection, optimal placement as well as ongoing optimization. Replica choice determines what valuable knowledge resources are and position decisions on where to place it depending on where the users will be located, the network latency and accessibility of a resource. The optimization process periodically surveys the replication that is going on and removes the redundant copies so that it can utilise the storage efficiently.

3.5 Adaptive Clinical Cache Layer

Adaptive Clinical Cache Layer retrieves the knowledge materials that are frequently requested in clinical knowledge that are high in speed. Defined by the dynamism, this layer gets its way of dynamically dealing with the information pertaining to the caching, on the basis of not only the varying loads and demands of the healthcare systems but also the retrievals. In comparison to the demand curves based on clinical requirements, the RACF allows its caches to be updated continuously unlike the immobile caching schemes.

Adaptive caching mechanism defines the nature of resources that are to be stored locally, nature of resources that are to be substituted and information that need to be given immediately. The rate of retrieval of clinical knowledge that has a high level of demand and is of high medical importance will have a high caching priority. The framework also takes into account the storage constraints and avoids information over-replication of low value information.

RACF improves quicker retrieves, utilizing extra cash, and reduces the quantity of unnecessary vendors between healthcare packages and central databanks using adaptive cache control. This is particularly useful in situations where the healthcare setting is spread across and a fast access to clinical knowledge will be required.

3.6 Consistency and Synchronization Management Layer

Being able to sustain uniformity in set resources that replicate clinical knowledge is a paramount provision of RACF. A lot of healthcare information is dynamic, as a medical guideline can always be altered, new treatment discovered, drug information changed and a change in clinical practice must be made. Hence, compensated knowledge should stay in line with one another to guarantee solid decision support.

Consistency and Synchronization Management Layer oversees the status of replicas and how updates of information in distributed nodes of a cache are done. The replica version tracking, freshness checking, updates and conflict resolution are done in the layer. Because the main repository of knowledge is evolving during any kind of changes, it is the synchronisation procedure that will spot any hitched replication and convey alterations in an efficient style.

| ISSN: 2320-0081 | www.ijctece.com ||A Peer-Reviewed, Refereed and Bimonthly Journal |

|| Volume 7, Issue 6, November –December 2024 ||

DOI: 10.15680/IJCTECE.2024.0706033

The architecture follows a discipline of controlled synchronization in the development of a trade off between freshness of information and communication load. Information that is not very important can be updated with slower rates as compared to the synchronization of the clinically significant and up to date information. This will be one sure method of achieving credible information on clinical knowledge that will not raise any other traffic on the network.

3.7 Clinical Knowledge Retrieval Interface

The Clinical Knowledge Retrieval Interface is a web that involves communication amid the health practitioners, clinical applications and the RACF system. It provides access to the information which is of significance through a smart recovery process to the users and health systems.

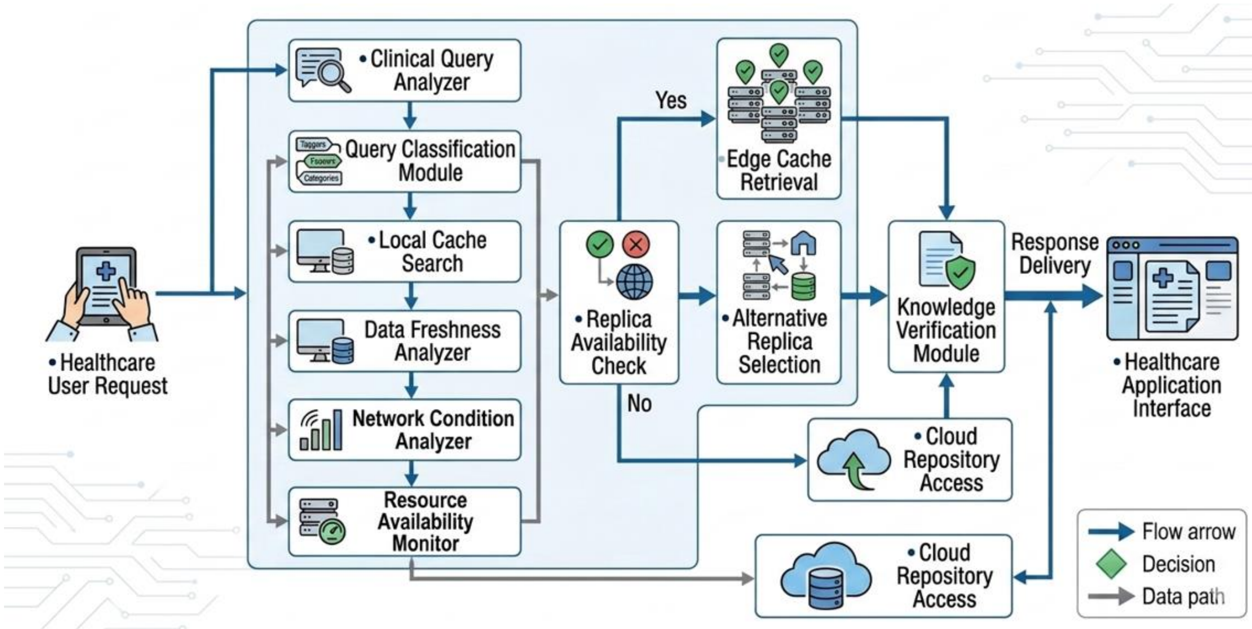

Figure 3- Clinical Knowledge Retrieval Process Workflow

When a query is developed, the framework examines the query and identifies the most appropriate source that ought to be recovered. Assuming that the required knowledge could be stored in a local cache node, RACF retrieves the information at local leverages so that the delay could be minimised. In case the information is not available on a replica the request is sent to another replica or the central location of the knowledge.

The retrieval interface also makes sure that the clinical users are able to access the appropriate and timely information besides minimizing the use of centralized systems. The mechanism caters to real-time healthcare applications like clinical decision support, artificial intelligence-assisted diagnosis, as well as personalized treatment-recommendation systems.

IV. FRAMEWORK EVALUATION

4.1 Evaluation Objectives and Experimental Setup

Experimenting the proposed Replication-Aware Caching Framework (RACF) investigate the advantages of the recommended solution to streamline access to clinical knowledge, minimize access delays, and capitalize on caching, as well as ensure stability in a decentralized healthcare setting. The score is decided on the capacity of RACF to deal with at the aged community levels at managing the replication of clinical knowledge resources undergoing a fluctuating workload of healthcare. The framework performance is compared to the traditional caching schemes that do not comprise a clinical-awareness, adaptive-replication and consistency-based management.

The proposal is that the edge-cloud environment (a set of clinical nodes) of a simulated healthcare system is assessed, comprising of hospital edge servers, clinical application gateways and centralized cloud repositories. The constructed experimental environment results in several clinical knowledge requests that are reflective of clinical settings in the real

| ISSN: 2320-0081 | www.ijctece.com ||A Peer-Reviewed, Refereed and Bimonthly Journal |

|| Volume 7, Issue 6, November –December 2024 ||

DOI: 10.15680/IJCTECE.2024.0706033

world like access to diagnostic guidelines, treatment regises, drug information and medical decision support software. The variant conditions of the workload are varied by the request intensity, knowledge popularity distribution, frequency of updates and the availability of the network.

Such key performance indicators, as the retrieval latency, the ratio of the cache hits, the replication performance, the overhead of communications, the ability to maintain the consistency are considered in the given analysis. The parameters are an overall decision of how able the RACF can be to support any scurable and reliable management of clinical knowledge.

4.2 Retrieval Latency Analysis

The retrieval latency is probably among the most significant performance measures in health care applications: the ability to respond to clinical data in ways corresponding to a timely fashion is directly proportional to the success of decision support. The proposed RACF model will reduce the retrieval time since it will be an intelligent replication of the knowledge resources that are highly accessible and clinically relevant close to the users of healthcare.

The comparison shows that RACF has a shorter response time than traditional caching solutions due to their edge caches being optimized to serve queries compared to having to access central repositories over and over again. The duplicate smart engine makes the system expectationate high need clinical resource and plan them in ways that codify them to suitable nodes so as to curtail any transmission delays throughout the system.

The Latency improvement is more pronounced with a high request loads and the congestion in centralized approaches is more pronounced in this scenario. RACF boasts of constant-retrieval performance that scales the requests amongst the series of replicas and alters the location of caches dynamically in relation to deformations in workload.

4.3 Cache Utilization and Hit Ratio Evaluation

The cache hit ratio is the effectiveness of a caching mechanism in responding to a request at a locally stored information. Caching Proficiency RACF increases sensitivity of cache utilisation with consideration of clinical vitality, frequency of utilisation and time sensibility in managing a cache.

Compared to traditional techniques of caching that are premised on popularity, RACF explores knowledge objects that may have high clinical and clinical value and demand high demands. This decision is made in the adaptive cache layer which identifies continuously by continually updating the information stored within it through modifying the least useful knowledge and trying to replace them with clinical knowledge that is more useful. This will enhance the likelihood of the local retrieval and will enhance the overall effectiveness of the system.

The results of the evaluation show that RACF offers better cache hit ratio than current methodology due to replica selection being intelligent whereas unnecessary information costs are not stored.

4.4 Replication Efficiency and Resource Management

Efficient replication of RACF is determined by looking at efficiency of framework to disseminate replicas of clinical knowledge at minimal storage and communication cost. Overreplication has the potential of driving up the consumption of resources, but underreplication can also lead to an increase in latency of retrieval. To overcome this worry, RACF will use a balanced replication plan that will look at the clinical relevance, the trend of demand and the resources at hand.

This analysis proposes that RACF can be used successfully to deposit replicas successfully as it only produces new versions of the valuable clinical assets. The system reduces replication besides the element of redundancy and guarantees of further availability of vital knowledge between the distributed healthcare nodes. This enhances more effective utilization of resources and becomes easy to apply in a big healthcare system at low costs.

4.5 Consistency and Synchronization Evaluation

The consistency check monitors the ability of RACF to make sure that the clinical knowledge is correct and up-to-date, even in the nodes with the same caches. Given that medical information can often be updated, it is important to ensure that there is synchronization of replicas in order to support dependable healthcare applications.

The proposed consistency management layer is viable and effective in regards to receiving updates on replicas and prioritizing synchronization of resources of clinical significance. In the results of the evaluation, both diminished

| ISSN: 2320-0081 | www.ijctece.com ||A Peer-Reviewed, Refereed and Bimonthly Journal |

|| Volume 7, Issue 6, November –December 2024 ||

DOI: 10.15680/IJCTECE.2024.0706033

accessibility of stale knowledge is evident by RACF coupled with acceptable communication overhead costs. The framework achieves the balance between the freshness of the information and efficiency of synchronization and reflects it as appropriate in the context of dynamic healthcare environment.

V. CONCLUSION AND FUTURE WORK

5.1 Conclusion

To obtain the low-latency access to clinical knowledge in a distributed healthcare setting, this paper introduced Replication-Aware Caching Framework (RACF). The ongoing migration towards the implementation of electronic health records, IoMT-based systems, and AI-based clinical decision-support systems, has posed a considerable challenge in terms of ensuring that massive amounts of healthcare data are managed with maximum accessibility and a minimum of response time. Various legacy caching techniques do not apply to healthcare applications since they tend to be speed-oriented with little regard to the clinical significance, replica maintenance, information renewal and congruity specifications.

The proposed RACF considers these limitations with a combination of acquisition of clinical knowledge, semantics representation, intelligent replication management, adaptive caching, consistency synchronization and optimized retrieval mechanisms. The framework allows discovering the useful resources in the clinical knowledge sphere dynamically and assigns the duplicates to the right edge and cloud location using the demand patterns, clinical relevance and state of the system. RACF increases the access of knowledge and limits the wanton attention of the data transfer and storage with intelligent and adaptable cache management.

The framework analysis reveals that the framework can lead to greater efficiency in regards to clinical retrieval through decreased latency, improved utilization of the cache and towards locating replicas most optimally and distributed knowledge sources. The solution suggests is scalable to apply healthcare settings where quick access to dependable healthcare data is vital to clinical decision-making and smart health care services. RACF sets the stage of future healthcare infrastructures which involve effective management of ever-changing clinical knowledge.

5.2 Future Work

Though it is true that RACF offers one of the best approaches to foster replication-minded clinical knowledge retrieval practices, it is possible to recognise a series of opportunities that can be exploited. Reinforcement learning and AI can also be explored in more detail to determine how to build the best system to allow completely autonomous replication decisions in the unpredictable workload of healthcare. Machine learning models can be trained to produce predictions of future clinical information requirements and, more importantly, predict a requirement in advance by placing caches in advance before changes in workloads happen.

Federated learning, differential privacy and applied use of secure aggregation could also be investigated as a means to explore the privacy preserving replication mechanisms in the future to protect sensitive healthcare information due to the need of distributed caching. Also, its application and testing on a hospital level, between inter-citation healthcare systems, and the interface clinical systems will shed even more light on its functionality in real world and scalability.

The other crucial field of study is being able to combine RACF with large language models and retrieval-augmented generation (RAG)-supported clinical assistants to facilitate the intelligent medical question answering system. It can scaffold it too with the systems of audio blockchain audit to enhance efficacy and trustworthiness in the reproduced knowledge administration of clinical knowledge. These improvements will see RACF at the position where it can sustain independent, secure and highly responsive healthcare ecologies that can meet the requirements of the future applications of digital healthcare.

REFERENCES

[1] A. Alourani, M. Sardaraz, M. Tahir, and M. S. Khan, "Dynamic and energy efficient cache scheduling framework for IoMT over ICN," *Applied Sciences*, vol. 13, no. 21, Art. no. 11840, 2023.
[2] C. Pruthvi, H. Vimala, and J. Shreyas, "A systematic survey on content caching in ICN and ICN-IoT: Challenges, approaches and strategies," *Computer Networks*, vol. 233, Art. no. 109896, 2023.
[3] C. Pruthvi, H. Vimala, J. Shreyas, and S. Devayya, "ICN based co-operative edge caching policy for transient IoT data," in *Proc. 8th International Conference on Computing in Engineering and Technology (ICCET)*, Patna, India, 2023.

| ISSN: 2320-0081 | www.ijctece.com ||A Peer-Reviewed, Refereed and Bimonthly Journal |

|| Volume 7, Issue 6, November –December 2024 ||

DOI: 10.15680/IJCTECE.2024.0706033

[4] B. Feng, A. Tian, S. Yu, J. Li, H. Zhou, and H. Zhang, "Efficient cache consistency management for transient IoT data in content-centric networking," *IEEE Internet of Things Journal*, vol. 9, no. 15, pp. 12931–12944, 2022.
[5] Z. Zhang, X. Wei, C.-H. Lung, and Y. Zhao, "ICache: An intelligent caching scheme for dynamic network environments in ICN-based IoT networks," *IEEE Internet of Things Journal*, vol. 10, no. 2, pp. 1787–1799, 2023.
[6] H. Wu, Y. Xu, and J. Li, "PTF: Popularity-topology-freshness-based caching strategy for ICN-IoT networks," *Computer Communications*, vol. 204, pp. 147–157, 2023.
[7] M. S. Zahedinia, M. R. Khayyambashi, and A. Bohlooli, "Fog-based caching mechanism for IoT data in information centric network using prioritization," *Computer Networks*, vol. 213, Art. no. 109082, 2022.
[8] S. Alduayji, A. Belghith, A. Gazdar, and S. Al-Ahmadi, "PF-EdgeCache: Popularity and freshness aware edge caching scheme for NDN/IoT networks," *Pervasive and Mobile Computing*, vol. 91, Art. no. 101782, 2023.
[9] O. Serhane, K. Yahyaoui, B. Nour, R. Hussain, S. M. A. Kazmi, and H. Moungla, "PbCP: A profit-based cache placement scheme for next-generation IoT-based ICN networks," *Computer Communications*, vol. 194, pp. 311–320, 2022.
[10] H. Al-Ward, C. K. Tan, and W. H. Lim, "Caching transient data in information-centric Internet-of-Things (IC-IoT) networks: A survey," *Journal of Network and Computer Applications*, vol. 206, Art. no. 103491, 2022.
[11] G. Muhammad, F. Alshehri, F. Karray, A. El Saddik, M. Alsulaiman, and T. H. Falk, "A comprehensive survey on multimodal medical signals fusion for smart healthcare systems," *Information Fusion*, vol. 76, pp. 355–375, 2021.
[12] M. M. Islam, S. Nooruddin, F. Karray, and G. Muhammad, "Internet of things: Device capabilities, architectures, protocols, and smart applications in healthcare domain," *IEEE Internet of Things Journal*, vol. 10, no. 4, pp. 3611–3644, 2023.