Challenges and Opportunities in Mobile Cloud Computing for Big Data Analytics

Gurjit Singh Bhathal

Department of Computer Science and Engineering, Punjabi University, Rajpura Road, Patiala-147002, India
Email: *gurjit.bhathal[at]gmail.com*

Abstract: ***Large-scale distributed systems, such as cloud and mobile cloud deployments, offer valuable services that enhance people's quality of life and increase organizational efficiency. Cloud computing addresses the challenges of peer-to-peer (P2P) computing and introduces P2P cloud systems as an extension of federated cloud. These decentralized systems, composed of independent nodes and resources without central control, can efficiently provision resources at a low cost. As a result, there is a wide range of mobile applications and services designed to effortlessly scale to billions of mobile devices. Data-driven applications are particularly prominent in this context. In this research paper, we begin by examining current layered cloud architectures and offering a solution for managing large amounts of data. Next, we investigate the potential of using a P2P Cloud System (P2PCS) for data processing and analysis on a large scale. We then suggest an effective hybrid mobile cloud computing model based on the concept of cloudlets. Finally, we strengthen our recommendations by introducing and evaluating security and privacy measures to protect against potential attacks.***

Keywords: Cloud Computing, Mobile Cloud Computing, Big Data, P2P Cloud System, Data Analytics, Security, Cloudlet

1. Introduction

The rapid advancements in cloud and mobile computing have led to the emergence of mobile cloud computing, which has revolutionized the way applications and services are developed, ultimately enhancing the quality of life and efficiency in organizations [1]. The widespread use of mobile devices has resulted in a proliferation of products across various industries such as social networks, education, healthcare, and government, generating large amounts of data commonly referred to as big data. While these technological advancements have brought about numerous benefits, they have also given rise to new challenges, including the need for effective storage and processing of big data, ensuring user privacy, and safeguarding sensitive information [2].

Big data refers to a large amount of structured, semi structured, and unstructured data that includes implicit information and is gathered from various devices like smartphones, personal computers, traffic cameras, and sensors [3]. The term 'big' not only highlights the vast size of the data (often measured in terabytes, petabytes, or zettabytes) but also encompasses different data types and the speed at which data is generated or collected. Big data processing can be initiated on-demand or on a scheduled basis depending on the nature of the task at hand. Typically, this processing, also known as big data analytics, involves analyzing and examining large datasets in order to gain insights and make informed decisions. [4]

One important example is the extensive patient record databases found in hospitals. Hospitals must securely store and safeguard their patients' medical information from unauthorized access for a variety of reasons [5]. This information should be readily accessible for medical analysis to help make informed decisions. These decisions could range from deciding whether or not to proceed with a medical procedure, to determining if a patient should be discharged or readmitted. Such decisions not only impact the patient's well-being, but also help to reduce costs (such as avoiding unnecessary hospital stays) and promote a more efficient healthcare system.

Governments possess extensive datasets that they must manage to serve the public at various levels. This includes nations, states, and cities. Similarly, sports teams rely on vast amounts of fan data to forecast ticket sales and analyze team tactics. Additionally, social media platforms play a significant role in today's society, impacting marketing, advertising, and manufacturing practices. With a large user base, social media companies must analyze massive sets of data to provide superior service to both individual users and business clients seeking to reach this audience.

For many real-world problems, cloud infrastructures are seen as the ideal solution to handle large amounts of data. Processing in this context doesn't just involve storing data, but also analyzing it. Most cloud providers offer data analytics services, sometimes referred to as Analytics as a Service (AaaS), integrated into their infrastructure. AaaS offerings are comprehensive, allowing for analytics on large and diverse datasets. With efficient connection layers, the cloud can gather data from reputable sources and prioritize them. After analysis, the results can be visualized to uncover valuable insights.

There are three primary cloud architectures: peer to peer (P2P), federated, and centralized. Centralization is best suited for applications that require low communication delays, commonly used in computing clusters and datacenters of many cloud providers. Clients are connected to the closest datacenter to minimize communication latency due to geographical partitioning of cloud resources over wide distances. Federated cloud is utilized to create larger clouds by combining multiple smaller clouds. This architecture is advantageous when clients need to maintain a high level of confidentiality while distributing data geographically.P2P cloud technology involves expanding the idea of federated systems by creating a cloud network that does not rely on centralized components for control and oversight. Instead, it is built on independent peers and resources, allowing for resource provisioning at a low cost due to limited management requirements. The benefits of big data analysis become clear when insights are provided to users after analyzing the data. Cloud computing is often seen as the most economical choice for big data analysis. In addition to storing data, the cloud computing platform offers users access to various sophisticated tools in areas such as machine learning and artificial intelligence. By utilizing these tools, users can examine and interpret a wide range of data formats, including videos, images, and text. When working with large amounts of data, traditional database methods may not be the best option. Database queries can become complex and expensive when dealing with many different attributes. Additionally, there are scalability issues to consider, along with concerns about data security and ownership. Storing and accessing data in the cloud can also be risky, as data owners must relinquish some control to the cloud service providers.

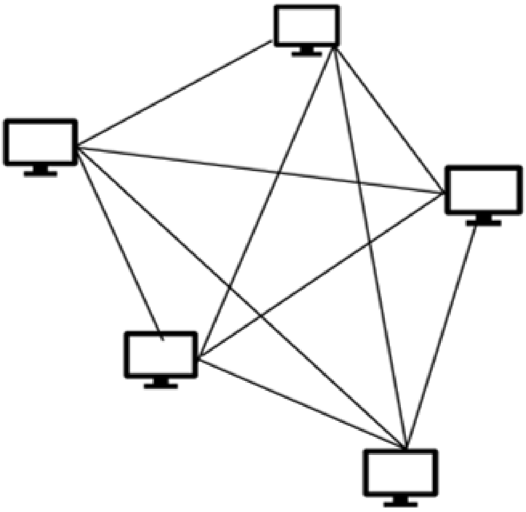

Figure 1: P2P Architecture of Distributed Computers

2. Related Work

Recently, there has been a growing interest in Peer-to-Peer Cloud Systems (P2PCS). Various studies have been conducted to explore different architectures for creating cloud computing systems [6]. More recent research by [7] focuses on developing a P2P-cloud system using affordable cloud resources to offer a wide range of services. Additionally, [8] introduced a mobile cloud computing model that can support various practical applications, including storing and analyzing data from sensors like fire and motion sensors, as well as IoT devices. The gathered data is then transmitted to the mobile cloud model.

However, cloud and mobile cloud systems face numerous threats and potential attacks that could jeopardize the privacy and integrity of important user data [9]. Some studies have integrated cloud infrastructures with tailored services for different industries. This means that the cloud is designed to cater to specific services, like cloud computing for manufacturing or healthcare.

3. Cloud Architecture and Big Data

The foundation of efficient big data management lies in robust cloud architectures personalized to handle large-scale data processing and storage. This section explores into the details of cloud infrastructure and its essential role in simplifying big data analytics. Subsequently, the discussion transitions to discovering innovative paradigms such as P2P cloud systems and mobile cloud computing, elucidating their integration with big data analytics to harness distributed computing capabilities and enhance data accessibility and processing efficiency.

3.1 P2P Cloud System

The P2P Cloud System is made up of a group of hosts or nodes, known as peers, that all run the same software processes. These processes are organized and executed based on the architecture layers provided. The first layer, also known as the Peer Sampling Service (PSS), uses a simple gossip protocol [10] where each node receives a list of neighboring nodes that it can communicate with. Each neighboring node in the local view has an ID (such as an IP address) and a timestamp. Neighboring nodes are added to the local view based on the time of their first interaction, as indicated by the timestamp. Neighbors regularly exchange and combine their perspectives; they remove the oldest entries to maintain a fixed list size determined by each node. Because the list of nodes can change with each message, the local view is a dynamic list. PSS is seen as a highly efficient solution for decentralized environments were nodes control resources independently.

In the next layer, the Slicing Service (SS) ranks nodes based on user requests using specific criteria. When a user asks for a particular set of nodes, all matching nodes are grouped to create a slice or sub cloud. For example, a user could request the top 5% fastest nodes to form a slice.

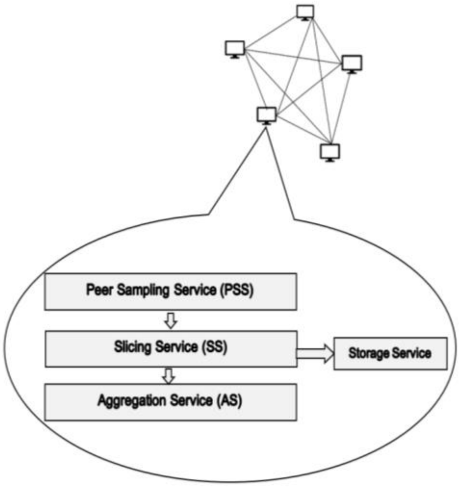

Figure 2: Layered Architecture of P2P Cloud System

The third layer of the cloud system is known as the Aggregation Service (AS). In this layer, cloud-wide parameters are provided to any node upon request without needing to access the global cloud registry. These parameters include information about the status of the cloud system, such as the total number of nodes, average load, and utilization.

Instead of relying on a central unit, decentralized aggregation methods are used to generate these values. The data collected in this way can be accessed through a Monitoring Service (MS) using APIs that run on top of AS. These APIs allow users to monitor the state of nodes and view the network topology. Some other important parts of the P2PCS include the Storage system, which is set up as a distributed service. The authentication system in P2PCS plays a key role in controlling access and granting the appropriate permissions to authorized users.

3.2 Mobile Cloud Computing and Big Data

The progress of big data innovation is ongoing, fueled by sophisticated analytics supported by cloud and mobile cloud computing (Fig. 2). The emergence of data-driven applications is simplifying various aspects of our lives. However, the true significance of big data lies not in its sheer volume, but in its capacity to be transformed, refined, and linked in significant ways. This shift towards "data with intelligence" underscores the necessity of a strong partnership between mobility and cloud computing [11].

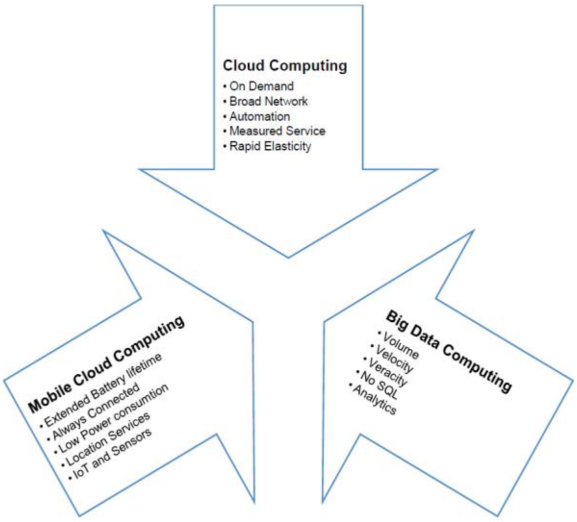

Figure 3: Big Data, Cloud Computing and Mobile Computing together for Analytics

In order to achieve this goal, companies will need to upgrade their mobile apps to include analytic capabilities. This will allow for data to be processed and evaluated before being collected. To enable mobility, the analytic apps are developed and stored using cloud computing technology. Users can access these apps through web servers on their mobile browsers, regardless of the specific mobile operating system, memory, and capacity requirements.

The latest offering in the world of technology is providing detailed analysis of complex large-scale data through mobile cloud computing to meet the demands of businesses. This is achieved by combining Infrastructure as a Service (IaaS) and Software as a Service (SaaS) to create Big Data as a Service (BDaaS). Security is a top priority in this field, with measures in place to ensure the safety of data at both the system and application levels [12].

When discussing application security, it is crucial for runtime applications catering to big data analytics on mobile devices to have built-in self-protection mechanisms. Traditional firewalls and perimeters are no longer sufficient to provide the necessary high-level security. Developers of such applications must implement adaptive access control measures. Additionally, at the system level, a combination of technologies including machine learning and text mining can be utilized to create programs for predicting and preventing threats. By addressing security concerns at both the application and system levels, advanced protection standards can be ensured against potential threats.

3.3 Cloudlet for Mobile Cloud Computing

A "cloudlet" refers to a small-scale cloud data center or server located at the edge of the network, closer to mobile devices. These cloudlets provide computing resources and services to mobile devices in proximity, reducing latency and bandwidth usage by offloading computation and storage tasks from remote cloud servers. They enable mobile devices to access cloud services and resources with lower latency and higher bandwidth, improving the performance and responsiveness of mobile applications. Cloudlets are particularly useful for latency-sensitive applications, real-time processing, and offloading resource-intensive tasks from mobile devices to nearby servers. Overall, cloudlets bridge the gap between mobile devices and remote cloud servers, enhancing the capabilities and performance of mobile cloud computing environments.

3.4 Secure Crypto-Processors (SCPs) for Mobile Devices

Secure Crypto-Processors (SCPs) are integral to Mobile Cloud Computing (MCC) for Big Data Analytics, ensuring data security across mobile devices and cloud servers. SCPs provide dedicated hardware for cryptographic operations, safeguarding data confidentiality and integrity. They address security concerns like data breaches by enabling robust encryption and secure key management. SCPs enhance MCC's security posture, crucial for handling sensitive data in big data analytics. With SCPs, MCC environments can leverage cryptographic techniques confidently, ensuring data protection and integrity throughout the analytics process, thus enhancing trust and reliability in data-driven decision-making

4. Big Data Analytics with MCC

Mobile cloud computing (MCC) offers a powerful yet challenging environment for big data analytics. While network limitations and device constraints can hinder processing, MCC unlocks real-time analytics, personalized experiences, and large-scale data collection through smartphones. The key lies in balancing these challenges with the exciting opportunities for innovation and data-driven decision making.

4.1 Challenges in MCC for big data analytics

The challenges in mobile cloud computing for big data analytics are multifaceted, reflecting the complexity of integrating these technologies to harness the potential of data-driven insights while navigating various obstacles. Here are some key challenges to consider:

- Data Security and Privacy: With the proliferation of mobile devices accessing cloud services, ensuring the security and privacy of sensitive data becomes paramount. Mobile devices are more vulnerable to security breaches, raising concerns about data protection and compliance with regulations like GDPR.

- Resource Constraints: Mobile devices have limited processing power, memory, and battery life compared to traditional computing systems. This limits the complexity and scale of data analytics tasks that can be performed directly on the device.
- Data Integration and Management: Integrating heterogeneous data sources from various mobile devices and cloud platforms can be complex. Ensuring data quality, consistency, and interoperability across these sources is a significant challenge.
- Scalability and Performance: As the volume, velocity, and variety of data generated by mobile devices continue to grow, scalability and performance become critical considerations. Ensuring that mobile cloud computing infrastructure can handle large-scale data analytics tasks efficiently is a key challenge.

4.2 Opportunities in MCC for big data analytics

As above discuss some challenges in mobile cloud computing similarly Opportunities open the new door for big data analytics to harness the potential of data-driven insights while navigating various obstacles. Here are some key points to opportunities:

- Real-Time Insights: Mobile devices generate vast amounts of data in real-time, providing opportunities for real-time analytics and decision-making. Mobile cloud computing enables the processing of this data in the cloud, allowing organizations to derive actionable insights rapidly.
- Cost-Effective Scalability: Mobile cloud computing offers cost-effective scalability by providing on-demand access to cloud resources. Organizations can scale their big data analytics infrastructure dynamically based on workload demands, without the need for significant upfront investment in hardware.
- Enhanced Collaboration: Mobile cloud computing enables collaboration and data sharing among geographically dispersed users. By leveraging big data analytics in the cloud, organizations can facilitate collaborative decision-making and knowledge sharing among employees, partners, and customers.
- Innovation and Competitive Advantage: Organizations that effectively leverage mobile cloud computing for big data analytics gain a competitive advantage by harnessing insights from data to drive innovation, optimize processes, and improve customer experiences.

5. Security Issues and Counter Measures

In mobile cloud computing for big data analytics, potential security threats include data breaches, where unauthorized access to sensitive information occurs during data transmission or storage in the cloud. Man-in-the-middle attacks can intercept data exchanges between mobile devices and cloud servers, compromising confidentiality. Denial-of-service (DoS) attacks may disrupt cloud services, leading to data unavailability and system downtime. Additionally, malware targeting mobile devices can compromise data integrity and privacy, posing significant risks to big data analytics in mobile cloud computing environments.

To mitigate security risks in mobile cloud computing for big data analytics, solutions include implementing robust encryption protocols to secure data in transit and at rest. Employing multi-factor authentication enhances access controls, while regular security audits and updates help to identify and address vulnerabilities, ensuring the integrity and confidentiality of data.

During mobile cloud computing for big data analytics, the system is vulnerable to various security threats. Data breaches, unauthorized access, and data tampering become significant risks, jeopardizing the confidentiality, integrity, and authenticity of sensitive data, leading to compromised security and potential loss of trust in the system.

To overcome from this problem Secure Crypto-Processors (CPs) play a crucial role in mobile cloud computing for big data analytics by ensuring the confidentiality, integrity, and authenticity of sensitive data and cryptographic operations. Integrated into mobile devices and cloud servers, SCPs provide a trusted execution environment for cryptographic processes, safeguarding cryptographic keys and data from unauthorized access and malicious attacks. This secure infrastructure enables organizations to harness the power of big data analytics while maintaining robust security measures, essential for protecting privacy and ensuring data integrity in a mobile cloud computing environment.

Cloudlets serve as decentralized data centers at the network edge, reducing latency and bandwidth constraints inherent in mobile cloud computing for big data analytics. By offloading computational tasks closer to mobile devices, cloudlets facilitate real-time processing of large datasets, enabling faster insights. This proximity also enhances data privacy and security by minimizing data transmission to remote servers, making cloudlets invaluable for handling sensitive big data analytics tasks in mobile environments.

6. Conclusions and future work

There are various technological advancements that are benefiting both individuals and organizations by enhancing efficiency. Some of these trends are cloud computing and mobile cloud computing. Despite their usefulness, these technologies come with their own set of challenges, such as privacy concerns and data security issues. In our study, we focused on recent emerging technologies like cloud systems, mobile cloud computing, P2P cloud systems, big data, and storage solutions. Additionally, we examined the security and privacy issues related to these technologies and discussed the significant attacks that pose a threat to cloud computing systems. We also looked at the traditional countermeasures used to combat these attacks.

Furthermore, our research delved into the layered P2PCS architecture and its significance in the realm of big data analysis. We then explored potential new strategies to combat security threats, presenting four innovative encryption techniques and assessing their practicality in safeguarding big data in cloud environments. These techniques include format preserving encryption, homomorphic encryption, verifiable computation, and secure multi-party computations. Additionally, we introduced a hybrid mobile cloud model and performed simulations to demonstrate the viability of utilizing mobile cloud computing models in practical big data applications, such as the healthcare sector. We also conducted a comprehensive analysis comparing performance parameters.

In conclusion, it is determined that the cloud and mobile cloud computing environments are well-suited for hosting and analyzing big data. These environments face various security threats from both developed and developing attacks. New and efficient countermeasures are necessary to protect big data in these settings. The next phase would involve exploring the implementation of advanced cryptographic techniques like homomorphic encryption and Format Preserving Encryption in real-world cloud environments to ensure the security of big data.

References

[1] J. Liu, E. Ahmed, M. Shiraz, A. Gani, R. Buyya and A. Qureshi, "Application partitioning algorithms in mobile cloud computing: Taxonomy, review and future directions," *Journal of Network and Computer Applications*, pp. 99-117, February 2015.

[2] I. Yaqoob, E. Ahmed, A. Gani, S. Mokhtar, M. Imran and S. Guizan, "Mobile ad hoc cloud: A survey," *Wireless Communications and Mobile Computing*, vol. 16, pp. 2572-2589, 2016.

[3] A. Oussous, F. Z. Benjelloun, A. A. Lahcen and S. Belfkih, "Big Data technologies: A survey," *Journal of King Saud University - Computer and Information Sciences*, vol. 30, no. 4, pp. 431-448, 2018.

[4] G. S. Bhathal and A. Singh, "Big data: Hadoop framework vulnerabilities, security issues and attacks," *Array*, vol. 1, no. 1, pp. 1-8, 2019.

[5] H. A. Kurdi, "HonestPeer: An enhanced EigenTrust algorithm for reputation management in P2P systems," *Journal of King Saud University Computer and Information Sciences*, vol. 27, no. 3, pp. 315-322, 2015.

[6] J. Wang, W. Zhang, Y. Shi and S. Duan, "Industrial Big Data Analytics: Challenges, Methodologies, and Applications," *IEEE Transactions on Automation Science and Engineering*, pp. 1-13, 2018.

[7] P. Kumari and P. Kaur, "A survey of fault tolerance in cloud computing," *Journal of King Saud University - Computer and Information Sciences*, vol. 33, no. 10, pp. 1159-1176, 2018.

[8] L. A. Tawalbeh, W. Bakhader, R. Mehmood and H. Song, "Cloudlet-Based Mobile Cloud Computing for Healthcare Applications," in *IEEE Global Communications Conference*, Washington, DC, USA, 2016.

[9] M. B. Mollah, M. A. K. Azad and A. Vasilakos, "Security and privacy challenges in mobile cloud computing: Survey and way ahead," *Journal of Network and Computer Applications*, vol. 84, pp. 38-54, 2017.

[10] M. Jelasity, S. Voulgaris, R. Guerraoui, A. Kermarrec and M. V. Steen, "Gossip-based peer sampling," *ACM Transactions on Computer Systems* , vol. 25, no. 3, pp. 1-8, 2007.

[11] -M. Liroz-Gistau, R. Akbarinia, E. Pacitti, F. Porto and P. Valduriez, "Dynamic Workload-Based Partitioning Algorithms for Continuously Growing Databases," *Transactions on Large-Scale Data- and Knowledge-Centered Systems XII. Lecture Notes in Computer Science*, vol. 2083, pp. 105-128, 2013.

[12] Gupta, Chakraborty and Rajput, "Cloud security using encryption techniques.," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 5, p. 425–429, 201

Author Profile

Dr. Gurjit Singh Bhathal is currently working as an Assistant Professor in Department of Computer Science and Engineering, Punjabi University, Patiala (Pb). He has received a Ph.D. in Faculty of Engineering and Technology. He has more than 20 years of experience in teaching and industry in India and abroad. He has supervised more than 35 M.Tech. dissertations and more than 72 publications. His research interests include Big Data, Cloud Computing, Information Security, Cyber Security, and Data Analytics. Dr. Bhathal was also awarded an Outstanding Scientist in Computer Science and Engineering at 4th Annual Research Meet – 2018.