3. CLIENT AWARE CLOUDLET FRAMEWORK

3.1 INTRODUCTION

The strategic partnership between ubiquitous computing devices, wireless communication infrastructure, and mobile web services with the advancements of cloud computing has laid the foundation for mobile cloud computing. The essential criterion for the enablement of cloud services through mobile is the process of offloading computationally intensive applications from the mobile platform to nearby or remote cloud infrastructure. This offloading process overcomes the limitations to access resource demanding applications involving multimedia, video analysis, 3D modeling etc. Thereby, it improves the performance to work in a near real-time execution, while the computation and storage are handled by the backend process using cloud platform. Offloading becomes the most essential technique especially for the mobile app developers; it gives them the flexibility to program for a mobile device without much concern about its inefficiency. However, certain important client attributes must be taken into consideration while developing a framework. Hence, along with the client aware features that are associated with its major functionalities, a framework to satisfy the requirements of a complete mobile cloud offloading process is proposed in this chapter.

3.2 SURROGATE SYSTEM FOR CLOUD

Mobile devices are intrinsically resource-poor and to augment it, the obvious solution would be to run the mobile application on the cloud platform. Hence the need for offloading to cloud arises. This makes the mobile device to function as a simple user interface device, while the resource-intensive works are done at the remote cloud servers that are always accessible through the internet. This mechanism to augment the mobile works well as long as the mobile network meets the requirements of the offloaded application. The seamlessness in connectivity will be fine for non-interactive and lightweight applications that do not demand heavy network bandwidth, but when it comes to interactive applications with near real-time feedback
and high bandwidth application, will certainly fail to meet the demand and can easily impair the application’s performance.

A simple case of offloading to a commercial cloud server like Amazon EC2 instance over mobile internet data or a Wide Area Network (WAN) will have a reasonably long Round Trip Time (RTT). The end-to-end latency associated with the delay can experience high variability when the user travels. In a study of cloud services by Lewis et al. [65] AWS instances have an average RTT of 74 ms, and this is unlikely to change as most service provider’s aim is to increase the network bandwidth for video streaming.

An alternative to cloud offloading is to have an intermediate server which is connected through the wireless network as depicted in Figure 3.1. The server acts as a surrogate system for the cloud and connects to the distant cloud server using a wired or wireless network. It is one of a kind of hybrid mobile cloud architecture as discussed in section 1.2.1. This server handles all the requests sent from the mobile client and performs the necessary offloading and computation task either partially or fully based on the requirements of the application.

![Figure 3.1: Computation offloading to surrogate system](image)

The choice of having a surrogate system for mobile computation offloading was first proposed by Sathyanarayana et al. [10], and the server was termed as ‘Cloudlet’. This cloudlet functioned as a data-center in a box
whose goal is to bring the cloud closer. This cloudlet uses multiple VMs to serve the mobile devices. This visionary research work gives an insight into the computational offloading concepts targeting common users by placing the cloudlet at various public places like coffee shops, shopping malls, and passenger transit areas. Though some challenges like, network latency, surrogate type and resource allocation have been proposed, other open challenges and real-time implementation issues are not discussed. This pioneering work has led the research communities to develop new and innovative methods to solve the shortcomings of cloudlet offloading.

The proposed research work inspires to adopt this traditional cloudlet for offloading concept and improves upon to augment the user by making the cloudlet intelligent enough to be aware of the client attributes. Further, it facilitates better quality of service through availability and security.

### 3.2.1 Cloudlet functionalities

The cloudlet in general functions as a resource-rich server connected to the internet to serve the nearby mobile devices.

<table>
<thead>
<tr>
<th>Features</th>
<th>Cloud Offloading</th>
<th>Cloudlet Offloading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Server</td>
<td>Distant cloud data center - generally public cloud</td>
<td>Nearby resource rich server - private cloud</td>
</tr>
<tr>
<td>Connectivity</td>
<td>Either 3G or 4G Mobile internet</td>
<td>High speed Wi-Fi connection based on mobile</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>Dependent on public network with multi-path routing - low bandwidth</td>
<td>Private network with one-hop communication - high bandwidth</td>
</tr>
<tr>
<td>Latency</td>
<td>High latency due to inherent internet traffic</td>
<td>Very low latency because of Wi-Fi network</td>
</tr>
<tr>
<td>Cost</td>
<td>High data transfer cost and cloud service cost</td>
<td>Nearly zero cost for data transfer and service</td>
</tr>
</tbody>
</table>
This concept of cloudlet can be envisioned as a local server working through Wi-Fi connectivity that is capable of running the workloads required by a user. Some of the major advantages of using cloudlet based offloading over cloud are listed in Table 3.1. To substantiate it consider a scenario where a student in a university and wished to view a 1 hour video lecture in a particular format. Unfortunately, the video is incapable to run it in his device. Then a cloudlet inside the campus may offer to start the video by offloading it, converting it into the desired format and sending it back. Thereby, saving transfer time, processing time and battery power. To consider this cloudlet offloading it is essential to compare it with other available choices like, what if it is done within the mobile device and at the cloud server.

By considering the same to be done using a commercial cloud server, then the time taken for transfer would be highly dependent on the network capability of the mobile data. Moreover, the cost becomes a major factor that is a variable involving the requirements of the application service. Other drawbacks are the time taken to upload the video and download it after conversion will certainly drain the battery, utilize full computation resource of mobile and even storage. Meanwhile, the user cannot able to do other activities and further, the cost of the internet would make a severe impact on the user. Hence in a limited geographical location, a Wi-Fi linked cloudlet server can easily do the job with one hop communication, high network speed with minimally or no cost involved.

Similarly, when the same video conversion process is done within the mobile (considering that the mobile has a suitable application), then undoubtedly, the resources of mobile would either fail to complete the task or take extremely much time. Thus, offloading is an essential element for heavy duty applications, and further offloading to a surrogate system instead of the cloud is indeed the most preferred choice.

However, the surrogate cloudlet server lacks certain qualities to fulfill the need of a user. Therefore, the proposed approach describes various client aware features as a criterion for offloading decisions and then taking care of all the needs of a client device for a seamless service.
3.3 CLIENT AWARE FEATURES

One of the most important attributes that must be taken into consideration is the client aware feature of the mobile phone. Hence the focus of the research narrow downs to client awareness for dynamic adjusting of offloading to changing operating conditions, by preserving the available sensing and interactivity capabilities of mobile devices. Therefore features like device capabilities, security profile and application availability attributes are important client aware features considered as in Table 3.2 for developing the CAC framework. Deciding on the situations and complexities for offloading the mobile application’s data is the crux of the research. The primary focus is to develop a middleware for the android application, whereby the middleware does the heavy lifting of contents needed for offloading from the mobile. Once the conditions are satisfied it is offloaded to the nearby Resource-rich Middleware (RM) cloud infrastructure, instead of the remote cloud data center. Offloading must satisfy multiple attributes to be checked and prove that the offloading process is worth.

Client aware features are classified into three types as in Figure 3.2.

1. Device attributes
2. Security attributes
3. Application attributes

![Figure 3.2: Client aware attributes](image-url)
**Device Attributes:** The hardware and software attributes related to the mobile devices are considered here. It is well known that the resource poverty of the device is the most important reason for which the very idea of moving to offloading is considered. Hence CPU and memory utilization and its corresponding battery consumption form the most important attributes of the device. Other important features are the network capability, i.e., the type of Wi-Fi connectivity the mobile device is supporting. Based on the present scenario IEEE 802.11 b/g/n and ac are the advanced networks supported in the mobile. Hence, any device that does not match the criteria would perform with poor bandwidth quality, and this directly affects the time to offload. Apart from these, the device MAC address and its hardware features like the battery type and capacity, number of cores and memory contribute to it.

**Application Attributes:** The type of application which are working online and its size helps to judge the performance of the device. Moreover, change in size of the application is one important attribute to identify if there is any malware being injected into the software. Apart from that, the memory utilized by the application, number of threads required to run the application, and the total number of time the applications have offloaded are also considered.

<table>
<thead>
<tr>
<th>Device Attributes</th>
<th>Application Attributes</th>
<th>Security Attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU utilization</td>
<td>Type of Application</td>
<td>Security Certificates</td>
</tr>
<tr>
<td>Memory Utilization</td>
<td>Size of Application</td>
<td>Malware attacks</td>
</tr>
<tr>
<td>Battery Power</td>
<td>Memory Utilized</td>
<td>Trusted Credentials</td>
</tr>
<tr>
<td>MAC address</td>
<td>No. of Threads per App.</td>
<td>Frequency of Access</td>
</tr>
<tr>
<td>Hardware Configurations</td>
<td>No. of times offloaded</td>
<td>Geographical Location</td>
</tr>
<tr>
<td>Network Capability-Bluetooth/Wi-Fi</td>
<td>Other applications used through Wi-Fi</td>
<td>Privacy Constraints</td>
</tr>
</tbody>
</table>
Security Attributes: Any certificate received from the third party would necessarily enhance the security of a device because only for a secure device, a third party would assess and provide it. Similarly, history of malware attacks will state the vulnerability of the device. Trust credentials stored in the trust store would also prove the device trustworthiness. Finally, the location and access patterns criteria helps to judge the client attribute with respect to security.

3.4 PROPOSED CLIENT AWARE CLOUDLET ARCHITECTURE

The Client Aware Cloudlet (CAC) is a three-tier architecture connecting mobile devices and cloud service providers. The CAC is built to function as a micro-private cloud with the Wi-Fi connection to accept the request and serve the front tier mobile clients and a wired/wireless internet to connect the backend cloud servers. The proposed cloudlet (hereafter CAC is interchangeably called as cloudlet) broadcast has different components and services to perform some multiple processes to solve the issues and requirements of mobile users.

3.4.1 Components of CAC

The components of the cloudlet are described below.

- Mobile User Interface Agent (MUIA): The client program that accepts the use case and identifies the type of application, and performs a basic check to locate the access point.

- Computational Analytics Engine-Mobile client (CAE-Mobile): Takes decision to offload based on client aware attributes.

- Computational Analytics Engine - Cloudlet service (CAE-CAC): Identify the client requirements and initiates computation offloading process.

- Client Awareness and Assessment Service (CAAS): Assesses client attributes like location, context and running processes, then, creates a metadata of all the mobile devices for future interaction.
• **Client Aware Security Certification (CASC) Service**: Performs security capability evaluation check to identify the trusted behaviour of client devices and then initiates offloading accordingly. Also recommends services for certified clients.

• **CA Resource Manager (CARM)**: Performs virtual machines management based on the load condition of multiple clients and capability of CAC.

• **Cloud Migration Service (CMS)**: If the requested service is unavailable or if the resources in the CAC are full then the migration service will direct the data to cloud for new services and resources.

These components work in a sequence, triggered by various events while the mobile interacts for offloading process.

### 3.4.2 Architecture of CAC

The architecture of CAC as shown in Figure 3.3 depicts the whole picture of the various processes. To begin with, initially, the cloudlet announces its presence through regular Wi-Fi broadcast signal (1). Any mobile client who is interested to utilize the service of CAC and augment his device performance and save considerable resources can connect by accepting the service offer. Once accepted, Mobile user interface agent (MUIA) gets initialized (2).

The required offloading process must be decided upon through various interactions between the device and the cloudlet. Hence Computational Analytics Engine (CAE) begins to exchange parameters (3) in order to decide the CAE functions at both the end through CAC offloading decision algorithm. Once the decision to offload has been taken, the Client Awareness and Assessment Service (CAAS) collects client attributes for the multiple purposes to support the services offered. These collected attributes are first given to the CA Security Certification (CASC) service for assessing the security and risk factors of the unknown client (5).
Figure 3.3: Client aware cloudlet architecture
Risk index is calculated using CA security certification algorithm and a certificate of security is issued only for those clients who have met the requirements. Then the decision to offload gets the approval to perform the physical offloading of data from the mobile device to the CAC through high-speed WiFi connection (5a). Recommendation for a new service and user rated services are offered as a choice for processing the offloaded data (5b).

Service oriented system like the CAC must have provisions to support the offloading demands of mobile clients. Hence it is essential to have Client Aware Resource Manager (CARM), to manage the incoming resource escalations. The resource demand is raised by the CAE of cloudlet (6a). This demand has to be served with the help of attributes stored at CAAS (6b). The resources, which may either, be a software application or additional VM or storage service request are demanded, then CA resource scalability protocol allocates the resources (7). Since the CAC is designed to function as micro-cloud architecture, the configurations of the system are limited, hence when the need for resource migration arises, the applications, VMs, and storage may be migrated to public cloud servers. Therefore, based on the information given by CARM (8), the Client Aware Migration Service (CAMS) decides to migrate (9). Once the process of computation is completed at the CAC end, then the final results must be sent back to the clients mobile. This process may have done based on the context aware nature of the client. If the client is still waiting for the results within the boundary of WLAN then, the resultant data is automatically pushed (10), else it is stored at clients personal cloud storage and then it can be accessed or pushed when needed (11).

The research, therefore, aims to provide a client aware offloading service for mobile users with various stages and policies. Moreover, the entire computational offloading process is governed through the framework of tasks to improve the QoS in a secured mobile cloud environment.

3.4.3 Functionalities and assumptions of CAC

Based on the client aware features, the research focuses towards building a framework that adopts the basic principles of cloudlet architecture.
Therefore the client attributes strengthen the cloudlet by having best case functionalities that are superior to the traditional cloudlet. Apart from that the research proposes new features for CAC based on the components designed to augments the mobile devices. The functionalities for client aware cloudlet focusing primarily on smartphone requires certain specialized working conditions based on the current trends. Some of the most advanced trends with respect to functionalities that are considered are as follows.

1. Network connectivity
2. Application focus
3. Business intelligence
4. Implementation methodology

### 3.4.3.1 CAC through Wi-Fi connection

New devices populate the mobile market with advanced capabilities and intelligence in different form factor. The connectivity between the cloudlet and mobile devices should be based on the target applications. Since the very purpose of getting into mobile cloud computing is for computationally heavyweight and near real-time applications, the data throughput must have a good bandwidth. Hence, the available choice of connectivity that is dependent on internet protocol is Wireless Wide Area Network (WWAN) and Wireless Local Area Network (WLAN). Current WWAN technology includes 3rd Generation HSPA (High Speed Packet Access) and 4th Generation LTE (Long Term Evolution) technology.

![Table 3.3: Comparison of 4G and Wi-Fi technology](image)

<table>
<thead>
<tr>
<th>Wireless Technology</th>
<th>Download Speed (Mbps)</th>
<th>Upload Speed (Mbps)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Theoretical</td>
<td>Real-world</td>
</tr>
<tr>
<td>WWAN - 4G LTE</td>
<td>&lt; 150</td>
<td>&lt; 20</td>
</tr>
<tr>
<td>WLAN - IEEE 802.11 ac</td>
<td>450 - 1300</td>
<td>150 – 300</td>
</tr>
</tbody>
</table>
Considering the future prospects, the 4G technology is better placed with a real-world Uplink/Downlink speed of 10/20 Mbps. However, considering high bandwidth utilizing applications for the research that is meant to be connected to a nearby surrogate server, 4G is rather slow and moreover costly. Hence adopting the WLAN technology of Wi-Fi IEEE 802.11 ac is the best case based on the speed as shown in Table 3.3, to connect the mobile device with the CAC.

The growth of Wi-Fi based communication is predicted to be the default access mechanism for most of the mobile devices working within a limited range. As per the Cisco’s global mobile data traffic [69], the expected growth of mobile devices connecting Wi-Fi and Wi-Fi only devices is around 57% and 10% Compound Annual Growth Rate (CAGR), shown in Figure 3.4.

Some of the other important features to go with Wi-Fi include the following.

- Wi-Fi access has had widespread acceptance by MNOs globally, and it has evolved as a complementary network for traffic offload purposes—offloading from expensive cellular networks on to lower-cost-per-bit Wi-Fi networks.

- Wireless usage is shifting indoors. Network analytics show that the majority of mobile data usage - close to 80 percent - is indoor and
nomadic, rather than truly mobile. Macro networks were built for voice on the go. Small cell networks are designed to address modern mobile data traffic patterns. This small cell can be achieved through Wi-Fi access points within the indoor location.

- Small cells offer new monetization opportunities by taking advantage of the intelligence inherent in the network, including policy, hyper-location, context, application, and device information. Businesses can use this information to engage with their customers in new ways, including through augmented experiences, location-based content.

- Globally, total public Wi-Fi hotspots (including homespots) will grow six-fold from 2016 to 2021, from 94.0 million in 2016 to 541.6 million by 2021. Total Wi-Fi homespots will grow from 85.1 million in 2016 to 526.2 million by 2021. Homespots or community hotspots are a significant part of the public Wi-Fi strategy.

- Though 5th Generation mobile internet data transfer matches the current speed of Wi-Fi, it is yet to see the light of full-fledged implementation in market space. Also, any type of WWAN connection will have a substantial latency in the range of around 120ms.

Overall, choosing Wi-Fi as the primary communication link between mobile and CAC and linking CAC to cloud with a fixed wired internet connection is the ideal case for the framework.

**3.4.3.2 CAC for video applications**

The choice of application for mobile cloud computing has already been discussed in Chapter 1. Of all the application, the growth of watching video in smartphones has been the trend. Mobile video traffic accounted for 60 percent of total mobile data traffic in 2016 [69]. Mobile video traffic now accounts for more than half of all mobile data traffic. Because mobile video content has much higher bit rates than other mobile content types, mobile video will generate much of the mobile traffic growth through 2021.
Mobile video will grow at a CAGR of 54 percent between 2016 and 2021, higher than the overall average mobile traffic CAGR of 47 percent. Of the 49 exabytes per month crossing the mobile network by 2021, 38 exabytes will be due to video as shown in Figure 3.5.

![Figure 3.5: Mobile data traffic prediction](Source: Cisco's global mobile data traffic)

With this proof of importance towards mobile video, the research mainly focuses towards handling videos.

### 3.4.3.3 CAC as an analytic engine

The most important feature of the proposed framework is its functionality as an analytic engine. It works initially as a decision-making engine when multiple devices enter the server for judging the request and processing it. Then it assesses the behaviour of the client device and hence it can be said as a behaviour analysis engine. Next it performs the task of computing the workload and hence it does the job of a computation engine. Finally, recommending suitable applications for the users is another important function of CAC, so recommendation engine is as well performed. Hence the client aware cloudlet architecture is a typical analytic engine performing various task, which can be further extended to do even complex big data analytics.
3.4.3.4 CAC as a private cloud

The proposed CAC functions mainly as a private cloud deployment, working as a micro-cloud data center. Earlier models have visualized the cloudlet as a data center in a box, working at the edge of the internet. However, the CAC which is developed using open-sourced cloud middleware works in-between mobile clients and cloud servers. This framework makes the model to be highly portable and scalable. It can be easily setup in an ad-hoc manner under crucial circumstances. For example, in the agriculture sector, finding a disease of a plant can be done by capturing the image of the plant using a farmer's low-cost camera mobile phone and sending the image to a local server for processing. This local execution would be the CAC server performing computationally intensive image processing techniques to detect the disease of the plant.

Further being connected to the cloud, the CAC can fetch specific information about the disease by coordinating with other CAC’s, thereby forming a network of Cloudlet’s. Therefore, the CAC in this scenario can be built as a surrogate system to process the offloaded image. Hence a single server private cloud can complete the task of serving multiple users, which can be setup when required during a particular period of harvest and disconnected. For this service, villagers need not pay for any mobile data or for application processing and enabling it as one of the best ways to access government cloud servers for disease analysis.

3.5 SUMMARY

- In this chapter, the need for a surrogate system as the intermediate server for offloading computation is presented with a scenario. Also, a comparison of cloud based offloading and cloudlet based offloading is envisaged.

- A complete architecture of the proposed CAC is elucidated with the flow of events. Various process and components of CAC are
described. The client attributes that are essential for evaluating the CAC framework has been summarized.

- Important functionalities and assumptions of the framework associated with its advantages are given to justify the need for the selection of various components of the CAC framework.

- The growth of various other technologies presented in this chapter like Wi-Fi, real-time mobile applications, and analytics of mobile data suggest for advancements and future scope of mobile cloud computing.

- The next chapter presents the offloading decision process based on the client aware features discussed in this chapter.