Geo-Cloudlet: Time and Power Efficient Geospatial Query Resolution using Cloudlet

Jaydeep Das*, Anwesha Mukherjee†, Soumya K. Ghosh‡, Rajkumar Buyya§

*Advanced Technology Development Centre
††Department of Computer Science and Engineering
Indian Institute of Technology Kharagpur, West Bengal, 721302, India
§Cloud Computing and Distributed Systems (CLOUDS) Lab
School of Computing and Information Systems
The University of Melbourne, VIC 3010, Australia
E-mail: *jaydeep@iitkgp.ac.in, †anweshamukherjee2011@gmail.com, ‡skg@cse.iitkgp.ac.in, §rbuyya@unimelb.edu.au

Abstract—Geospatial data analysis is an emerging area of research today due to the potential to enable varied location-aware services. The existing centralized cloud-based analysis becomes time and computing-intensive for huge amount of geospatial data processing. This paper addresses the challenge of time and power-efficiency in QoS-aware geospatial query resolution. We propose a cloudlet based hierarchical paradigm, namely Geo-Cloudlet, where the cloudlets contain the geospatial data of the districts. The state and national level geospatial data are stored inside the state cloud and country cloud respectively. The query resolution is performed by either the cloudlet or by the state cloud or country cloud depending upon the geographical region related to the query. The experimental analysis illustrates that the proposed architecture Geo-Cloudlet reduces the latency up to 61.3% and power consumption up to 61.1% over the use of only remote cloud servers for geospatial query resolution.

Index Terms—Geospatial Query; Geo-Cloudlet; Quality of Service; Time-efficient; Power-efficient

I. INTRODUCTION

Geospatial data storage and analysis is a major challenge due to its large volume. With the advancement in location acquisition systems, sensor networks, and mobile computing, a huge volume of geospatial data is collected [1]. For resolving geospatial queries, the huge volume of data has to be processed, which is computationally intensive and requires ubiquitous network access. Cloud computing offers a platform for geospatial data processing because of its ability to provide ubiquitous network access, on-demand self-service, resource pooling, rapid elasticity, and measured services [2]–[4]. Geospatial cloud computing is a cloud computing paradigm that is driven by geospatial sciences and optimized by spatio-temporal principles for enabling geospatial science discoveries within a distributed computing environment [1]. However, storing and processing of data completely inside the long distant cloud servers increases latency in query resolution. As a solution towards this problem, fog computing has been used for geospatial data processing in [5]. However, the intermediate devices like switch, routers which act as fog devices, partially offload data or computation. In our work, we will use cloudlets for geospatial data processing. Cloudlet is a computer or cluster of computers, which stores the frequently accessed data and acts as an agent between the client device and cloud [6]. Cloudlet has reduced the energy and latency over remote cloud servers in computation offloading [7]. Our objective is to provide a time and power-efficient paradigm for geospatial query resolution. In this paper, we have introduced a cloudlet based hierarchical architecture for geospatial information storage and processing, and the mathematical model of latency and power consumption for the proposed paradigm is developed. In our paradigm, mobile devices are the clients, which generate geospatial queries.

A. Related Work

Geospatial information refers to the data with respect to a geographical place, in terms of geographic coordinates. Geospatial data collection takes place by Geographic Information System (GIS) [8]. Using multi-dimensional data set, a method has been proposed for geospatial query resolution in [9]. The use of cloud data centres for geospatial data storage and analysis has been demonstrated in [1], [2], [3]. For massive geospatial data processing, a GIS querying framework has been proposed in [10]. In Geospatial Cloud computing, the application tier is used for geospatial services. There are various categories of geospatial services: Web Map Service (WMS) [11], Web Coverage Service (WCS), Web Feature Service (WFS) [12], Catalog Service for the Web (CSW) and Web Processing Service (WPS) [13]. Open Geospatial Consortium (OGC) constraints geospatial service chain based geospatial query resolution on the cloud has been discussed in [4]. For VM allocation with geospatial service chain learning a method has been discussed in [14]. However, access to long distant cloud servers for geospatial data processing increases the latency. As a solution, fog computing has been used for geospatial data analysis in [5], [15], [16]. Fog devices are the intermediate devices between the user and the cloud servers, which participate in data processing [17]. Edge devices are also nowadays used for offloading computation and storage [18]. Geospatial query processing in edge devices has been discussed in [16]. However, the processing of voluminous
geospatial data requires high-end processing, which is beyond the capability of the edge or fog devices. Cloudlet being a computer or cluster of computers, has the ability to store and process a large volume of data [6]. In our work, we use cloudlets and propose a hierarchical paradigm for geospatial query resolution.

B. Motivation and Contribution

Geospatial data are sensitive and of huge amounts. The fog computing based model for different applications such as health care uses fog devices for preliminary data processing. But in geospatial query resolution, a huge amount of geospatial data has to be processed. Hence the use of a traditional fog or cloud-based model is not enough for geospatial application. A distributed hierarchical model with good Quality of Service (QoS) is required for geospatial query resolution. Our objective is to propose a hierarchical paradigm for geospatial data storage and analysis, which will improve the QoS in terms of the power consumption of the user device and latency. The key contributions of this paper are:

- A geospatial data hierarchy is maintained based on country, state, and district. The national level geospatial data is maintained inside the country cloud, the state level geospatial data is maintained inside the state cloud and the district level geospatial data is maintained inside the cloudlets. The proposed cloudlet based architecture for geospatial query resolution is referred as Geo-Cloudlet.
- As instead of keeping all geospatial data inside the cloud servers, geospatial data are maintained inside the cloud and cloudlet in a distributed manner, the load on the cloud data centre is reduced.
- By storing geospatial data in such a hierarchical manner an extra layer of data security is provided. Only country cloud can solve the national level geospatial queries, only state cloud can solve the state level geospatial queries and only cloudlets can solve the district level geospatial queries.
- The latency in geospatial query resolution and power consumption of the user device during that period are calculated for the proposed architecture and experimental analysis for geospatial query resolution using the proposed model is performed. The experimental results demonstrate that the use of cloudlet provides better QoS in terms of power consumption and latency than the remote cloud-based geospatial query resolution system.

The rest of this article is organized as follows. Section 2 discusses the proposed hierarchical paradigm Geo-Cloudlet. Geospatial query analysis has been illustrated in section 3. In section 4, the latency in the proposed paradigm is calculated. Performance analysis has been done in section 5. In the last section, we conclude the paper.

II. Hierarchical Architecture of Geo-Cloudlet

The cloudlet-based system architecture of Geo-Cloudlet is shown in Fig. 1. In a country like India, there are a number of states, and each state contains a number of districts. In our system, we are considering a hierarchical architecture (Fig. 2), where level 1 is the country i.e. national region. The national region is divided into the state regions in level 2. Each state region is decomposed into the district regions in level 3. Two categories of the geospatial cloud are considered- state cloud and country cloud. National level geospatial data is maintained inside the country cloud, and state level geospatial data is maintained by the state cloud. The district level geospatial data is maintained inside the cloudlets. As instead of keeping all geospatial data inside the cloud servers, geospatial data are maintained inside the cloud and cloudlets in a distributed manner, the load on the cloud data centre is reduced. Only country cloud can resolve the national level geospatial query, only state cloud can resolve the state level geospatial query and only cloudlets can resolve the district level geospatial query. By storing geospatial data in such a hierarchical manner an extra layer of data security is provided.

In our system, cloudlet contains the geospatial information of the district and resolves geospatial queries related to the district level geospatial data. Otherwise, if the geospatial query is related to the state level geospatial data, the state cloud resolves the geospatial query. In other respect, if the geospatial query is related to the national level geospatial data,
The country cloud resolves the geospatial query. Thus, three following cases appear in our Geo-Cloudlet paradigm:

- **Case 1- Cloudlet resolves district level geospatial query:**
  When a mobile client generates a query related to the geospatial data of district level, the cloudlet with which the mobile device is connected, resolves the geospatial query. The sequence diagram of this type of geospatial query resolution is depicted in Fig. 3.

- **Case 2- State cloud resolves state level geospatial query:**
  When a mobile client generates a geospatial query related to the state level geospatial data, the cloudlet is unable to resolve because it is not containing that geospatial data. In that case, the cloudlet forwards the geospatial query to the state cloud, which further forwards the same to the country cloud. After analyzing the country level geospatial data, the country cloud sends back the result to the state cloud, which forwards the result to the cloudlet. The cloudlet forwards the result to the mobile device. This type of geospatial query resolution is depicted through a sequence diagram in Fig. 5.

### III. Geospatial Query Resolution

Geospatial data analysis is performed for resolving different types of geospatial queries, generated by the users. In our system, the cloudlet, state cloud, and country cloud perform geospatial data analysis to resolve the queries related to the district, state, and national level geospatial data respectively. In this section, we have discussed different types of geospatial queries along with the geospatial operations [19] to be performed on the geospatial data to resolve them.

*Geospatial Query 1 (District level query):* For preparing bridges over the drainage system where the roads are crossed over it in P district.

```sql
SELECT Rd.Lat, Rd.Long
FROM P_road Rd, P_drain Dr
WHERE Cross(Rd.Shape, Dr.Shape) = 1
AND district_name = 'P';
```

To resolve the above geospatial query, find the latitude (Lat) and longitude (Long) of the crossing point of the road...
and drainage table. Here topological geospatial operation ‘cross’ is performed over spatial join query on the Road and Drainage table of P district.

**Geospatial Query 2 (State level query):** Find out the One-Way roads and create 2.74 meters buffer to make it Two-Way extension in Q state

```sql
SELECT road_name
FROM Q_road
WHERE road_type = 'One_Way'
AND Buffer(area.shape, 2.74)
AND state_name='Q';
```

For resolving this geospatial query, find out the one-way road names from the road table of Q state and perform geospatial analysis operation ‘Buffer’ of fixed length 2.74 meters.

**Geospatial Query 3 (National level query):** List out the towns of Country U where S River supplies water within 150 kilometer distance.

```sql
SELECT T.Name
FROM Town T, River R
WHERE Overlap(T.Shape, Buffer(R.Shape, 150))=1 AND R.Name = 'S'
AND country_name='U';
```

To resolve this type of geospatial query, find out the town names where S river overlaps towns of country U. Here geospatial analysis operation ‘Buffer’ of fixed length 150 kilometers and topological geospatial operation ‘Overlap’ have been used.

**IV. LATENCY IN GEOSPATIAL QUERY RESOLUTION**

The parameters used in latency calculation are defined in Table I. The sum of propagation, communication, processing, and queuing latency is the total latency consumption while resolving the geospatial query for a mobile device. The propagation latency for case 1 is given as,

$$L_{p1} = d_{ml}/S_p$$

The communication latency in case 1 is given as,

$$L_{com1} = (D_{mt}/U_{ptml}) \ast (1 + f_{umt}) + (D_{tm}/D_{wmt}) \ast (1 + f_{dmt})$$

The processing latency in case 1 is given as,

$$L_{pr1} = D_{q1}/S_t$$

The total latency for case 1 is given as,

$$L_{case1} = L_{p1} + L_{com1} + L_{pr1}$$

The power consumption of the user device (mobile device) during this period is given as,

$$P_{case1} = L_{p1} \ast P_i + L_{com1} \ast P_c + L_{pr1} \ast P_i$$

The propagation latency for case 2 is given as,

$$L_{p2} = (d_{ml} + d_{ls})/S_p$$

The communication latency in case 2 is given as,

$$L_{com2} = L_{com1} + (D_{ls}/U_{ptls}) \ast (1 + f_{uls}) + (D_{st}/D_{wts}) \ast (1 + f_{dst})$$

The processing latency in case 2 is given as,

$$L_{pr2} = D_{q2}/S_s$$

The total latency for case 2 is given as,

$$L_{case2} = L_{p2} + L_{com2} + L_{pr2}$$

The power consumption of the user device (mobile device) during this period is given as,

$$P_{case2} = L_{p2} \ast P_i + L_{com2} \ast P_c + L_{pr2} \ast P_i$$

The propagation latency for case 3 is given as,

$$L_{p3} = (d_{ml} + d_{ls} + d_{sc})/S_p$$

The communication latency in case 3 is given as,

$$L_{com3} = L_{com2} + (D_{sc}/U_{ptsc}) \ast (1 + f_{usc}) + (D_{cs}/D_{wsc}) \ast (1 + f_{dsc})$$

The processing latency in case 3 is given as,

$$L_{pr3} = D_{q3}/S_c$$

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d_{ml}$</td>
<td>Distance between mobile device and cloudlet</td>
</tr>
<tr>
<td>$d_{ls}$</td>
<td>Distance between cloudlet and state cloud</td>
</tr>
<tr>
<td>$d_{sc}$</td>
<td>Distance between state cloud and country cloud</td>
</tr>
<tr>
<td>$S_p$</td>
<td>Propagation speed</td>
</tr>
<tr>
<td>$D_{q1}$</td>
<td>Data amount processed for resolving query in case 1</td>
</tr>
<tr>
<td>$S_t$</td>
<td>Data processing speed of cloudlet</td>
</tr>
<tr>
<td>$D_{q2}$</td>
<td>Data amount processed for resolving query in case 2</td>
</tr>
<tr>
<td>$S_s$</td>
<td>Data processing speed of state cloud</td>
</tr>
<tr>
<td>$D_{q3}$</td>
<td>Data amount processed for resolving query in case 3</td>
</tr>
<tr>
<td>$S_c$</td>
<td>Data processing speed of country cloud</td>
</tr>
<tr>
<td>$D_{ml}$</td>
<td>Data amount transmission from mobile device to cloudlet</td>
</tr>
<tr>
<td>$D_{im}$</td>
<td>Data amount transmission from cloudlet to mobile device</td>
</tr>
<tr>
<td>$D_{ls}$</td>
<td>Data amount transmission from cloudlet to state cloud</td>
</tr>
<tr>
<td>$D_{st}$</td>
<td>Data amount transmission from state cloud to country cloud</td>
</tr>
<tr>
<td>$U_{ptml}$</td>
<td>Data transmission rate from mobile device to cloudlet</td>
</tr>
<tr>
<td>$D_{wmt}$</td>
<td>Data transmission rate from cloudlet to mobile device</td>
</tr>
<tr>
<td>$U_{ptls}$</td>
<td>Data transmission rate from cloudlet to state cloud</td>
</tr>
<tr>
<td>$D_{wts}$</td>
<td>Data transmission rate from state cloud to country cloud</td>
</tr>
<tr>
<td>$f_{dmt}$</td>
<td>Uplink failure rate between mobile device and cloudlet</td>
</tr>
<tr>
<td>$f_{dls}$</td>
<td>Uplink failure rate between cloudlet and state cloud</td>
</tr>
<tr>
<td>$D_{q1}$</td>
<td>Queuing latency for case 1</td>
</tr>
<tr>
<td>$P_i$</td>
<td>Power consumption of mobile device in idle mode</td>
</tr>
<tr>
<td>$P_c$</td>
<td>Power consumption of mobile device in active mode</td>
</tr>
</tbody>
</table>
**Fig. 6.** Result of Geospatial Query 1

**Fig. 7.** Result of Geospatial Query 2

**Fig. 8.** Result of Geospatial Query 3
The total latency for case 3 is given as,

\[ L_{case3} = L_{p3} + L_{com3} + L_{pr3} \]  \hspace{1cm} (14)

The power consumption of the user device (mobile device) during this period is given as,

\[ P_{case3} = L_{p3} \times P_i + L_{com3} \times P_c + L_{pr3} \times P_i \]  \hspace{1cm} (15)

In the next section, the performance of the proposed system will be evaluated through experimental analysis, and the latency and user device’s power consumption during query resolution will be determined.

V. PERFORMANCE EVALUATION

We have performed the experimental analysis using the private cloud of the institution, Google Cloud Platform (GCP), mobile devices and cloudlet. Geospatial data are of two types: vector data and raster data. In this experiment, we have used vector data. We have used the geospatial reference system, EPSG:32645 and the geospatial data related to the road, rail, and land area of W and X districts of Y state of country Z.

A. Devices used in experiment

Two mobile phones are used in our experiment, which generate geospatial queries. We have used a laptop as a cloudlet, which is connected with mobile devices through the hotspot. The cloudlet is storing the geospatial data of the districts W and X. We have used the private OpenStack cloud of our institute, where a Virtual Machine (VM) instance is taken. In our experiment, the private cloud is used to analyse state level data. We have used Google Cloud Platform (GCP) where we have created one VM instance. In our experiment, the GCP is used to analyse country level data. The laptop which is used as a cloudlet is connected with the network through a Wi-Fi access point (AP). The AP is connected with Access Switch (AS), which is connected with Distribution Switch (DS). DS is connected with the Core Switch (CS), which is connected with the router. The used protocols are TCP, UDP, and ICMP. The data transmission rate is 40 Mbps. The configurations of the mobile phones, laptop (used as cloudlet), and VM instances are presented in Table II. QGIS (version 2.18.28) [20] is used for processing and analyzing geospatial queries.

B. Experimental results of query resolution

Four geospatial queries are generated from the mobile devices related to districts W, X, state Y, and country Z:

- One query is related to district W.
- One query is related to district X.
- One query is related to state Y.
- One query is related to country Z.

The laptop (cloudlet) is holding the geospatial data related to districts W and X. Therefore the laptop resolves the first two geospatial queries using QGIS and sends the result to the mobile devices. The private OpenStack cloud VM is containing the data of state Y. Hence for the third query, the private cloud VM is accessed. The GCP VM is containing the data of country Z. Thus for the fourth query, the VM of GCP is accessed. The results of geospatial query resolution after geospatial data analysis are presented in Fig. 6, Fig. 7, Fig. 8, and Fig. 9. Total latency in geospatial query resolution considering the propagation, communication, and processing (as discussed in section IV), are presented in Table III along with the power consumption of the mobile device during that period. The latency and power consumption are measured in seconds (sec) and milli-Watt (mW) respectively.

In Fig. 6, the result of geospatial query 1 is displayed. W district land use land cover (LULC) map image is considered. The user of mobile phone 1 queries for the land area over \( 100 m^2 \) in the W district. The geospatial data of W district is analyzed, and the ‘filter’ operation is used. The cloudlet (laptop) resolves the query and sends the result to the mobile phone 1. The latency for resolving this query and the power consumption of the mobile device during this period are presented in Table III and compared with the results if GCP VM resolves the query and sends the result to the mobile device. This is observed from case study 1 that using our system the latency and mobile device’s power consumption has been reduced to 61.3% and 61.1% respectively.
TABLE II
CONFIGURATIONS OF DEVICES USED IN EXPERIMENT

<table>
<thead>
<tr>
<th>Device</th>
<th>RAM</th>
<th>HDD</th>
<th>Processor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mobile phone 1</td>
<td>3 GB</td>
<td>32 GB</td>
<td>Qualcomm MSM8940 Octa Core</td>
</tr>
<tr>
<td>Mobile phone 2</td>
<td>2 GB</td>
<td>8 GB</td>
<td>64-bit 1.2 GHz Qualcomm Snapdragon 410 quad-core processor</td>
</tr>
<tr>
<td>Cloudlet</td>
<td>4 GB</td>
<td>250 GB</td>
<td>Intel Core i5</td>
</tr>
<tr>
<td>GCP VM instance</td>
<td>3.75 GB</td>
<td>250 GB</td>
<td>1 vCPU</td>
</tr>
<tr>
<td>OpenStack VM</td>
<td>4 GB</td>
<td>45 GB</td>
<td>2 vCPU</td>
</tr>
</tbody>
</table>

TABLE III
LATENCY AND POWER CONSUMPTION IN QUERY RESOLUTION

<table>
<thead>
<tr>
<th>Case Study</th>
<th>Geospatial Query</th>
<th>Transmitted data between consecutive nodes (MB)</th>
<th>Latency in Geo-Cloudlet (sec)</th>
<th>Latency in GCP VM (sec)</th>
<th>Power consumption in Geo-Cloudlet (mW)</th>
<th>Power consumption in GCP VM (mW)</th>
<th>Reduction in latency in Geo-Cloudlet than GCP</th>
<th>Reduction in power in Geo-Cloudlet than GCP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SELECT area_name FROM District_W WHERE area &gt; 100;</td>
<td>1.1</td>
<td>1.14</td>
<td>2.9464</td>
<td>1.41</td>
<td>3.63</td>
<td>61.3%</td>
<td>61.1%</td>
</tr>
<tr>
<td>2</td>
<td>SELECT Rd.Lat, Rd.Long FROM X.road Rd, X.drain Dr WHERE Cross(Rd.Shape,Dr.Shape)=1;</td>
<td>1.6</td>
<td>4.87</td>
<td>6.068</td>
<td>6.004</td>
<td>7.48</td>
<td>19.74%</td>
<td>19.78%</td>
</tr>
<tr>
<td>3</td>
<td>SELECT road_name FROM State_Y WHERE road_type = 'One Way' and Buffer(area.shape, 2.74);</td>
<td>3.75</td>
<td>6.35</td>
<td>9.286</td>
<td>7.83</td>
<td>11.45</td>
<td>31.61%</td>
<td>31.6%</td>
</tr>
<tr>
<td>4</td>
<td>SELECT road_name FROM Country_Z WHERE road_type = 'One Way';</td>
<td>5.7</td>
<td>13.8</td>
<td>13.8</td>
<td>16.97</td>
<td>16.97</td>
<td>Nil</td>
<td>Nil</td>
</tr>
</tbody>
</table>

In Fig. 7, the result of geospatial query 2 is displayed. The user of mobile phone 2 queries for the crossing points of drainage and roads for preparing bridges in X district. The geospatial data of X district is analyzed, and ‘cross’ operation is used. The cloudlet (laptop) after resolving the geospatial query sends the result to the mobile phone 2. The latency for resolving this query and the power consumption of the mobile device during this period are presented in Table III and compared with the results if GCP VM resolves the query and sends the result to the mobile device. The experimental results of case study 2 show that using our system the latency and mobile device’s power consumption has been reduced to 19.74% and 19.78% respectively.

In Fig. 8, the result of geospatial query 3 is displayed. The user of mobile phone 1 queries for one-way streets of Y state of width 2.74 meters (9 ft), to double the road width to make it two-way lane. For this purpose, the Y state geospatial data is analyzed, and a buffer of 2.74 meters over the one-way.
road is created. After resolving this geospatial query using private OpenStack cloud VM, the result is returned back to the mobile phone 1. The latency for resolving this query and the power consumption of the mobile device during this period are presented in Table III and compared with the results if GCP VM resolves the query and sends the result to the mobile device. From the experimental results of case study 3, it is observed that using our system the latency and mobile device’s power consumption has been reduced to 31.61% and 31.6% respectively.

In Fig. 9, the result of geospatial query 4 is displayed. The user of mobile phone 2 queries for one-way roads from Z country road map. The geospatial data of country Z is analyzed, and ‘filter’ operation is used. After resolving the geospatial query using GCP VM, the result is sent back to the mobile phone 1. The latency for resolving this query and the power consumption of the mobile device during this period are presented in Table III. As GCP VM is used for resolving query 4, the latency and power consumption of the mobile device is the same for the proposed model and if GCP resolves the query.

In Fig.10 the latency in query resolution and power consumption of user devices during that period are presented with respect to four experimental studies. From the experimental results of the four case studies of query analysis (see Table III and Fig.10), we observe that using the cloudlet up to 61.3% reduction in latency and 61.1% reduction in power consumption of the mobile device can be achieved than only remote cloud-based system, and the QoS is improved.

VI. CONCLUSIONS AND FUTURE WORK

In this paper, a cloudlet based hierarchical paradigm for geospatial query resolution, namely Geo-Cloudlet, has been proposed. The cloudlets contain geospatial data of the district regions. For the state and national level geospatial data storage and analysis, state cloud and country cloud are used respectively. When a geospatial query is received from a mobile device regarding the district region, the cloudlet resolves the geospatial query after analyzing the geospatial data and responds to the mobile device. Otherwise, if the geospatial query is regarding the state or national level geospatial data, the cloudlet responds using the state cloud or the country cloud. The experimental results illustrate that our proposed system improves the QoS by reducing the latency up to 61.3% and power consumption of the user device up to 61.1% than only remote cloud-based query resolution. Thus we can conclude that the proposed framework, Geo-Cloudlet, is a time-efficient paradigm as well as provides low power consumption of the user device during the query resolution period. As we are using a multi-tier framework, partial data processing and dealing with the inter-dependency among different task segments during geospatial data processing is a challenging future scope.

REFERENCES