Fog-enabled Sensor-Cloud

***Abstract*—In this paper,**

I. Introduction

A. Motivation

B. Contribution

II. Related Work

III. Problem Formulation

A. Problem Scenario

The demand for faster service is inevitable as the computation power and the need for various kinds of applications are increasing exponentially. To cope with this, we introduce the concept of Fog computing to the existing sensor-cloud architecture. A fog layer is placed in between the cloud and the physical sensor node layer. This paper aims to increase end-user satisfaction by handling some time-critical applications in the fog nodes, to render faster services. The time required, by a sensor, to send data back and forth to the cloud, is reduced by introducing the Fog layer since the Physical sensor node layer is closer to the Fog than it is, to the cloud. It also aims to assist the cloud by processing the data which is not essential to analyze in the cloud. Since the fog nodes devices have storage as well as processing power, they themselves process the data received from the sensors.

We divide the user applications into two parts: 1. Time-sensitive applications and 2. Non- time-sensitive applications.

B. Architecture

Figure. shows the structure of the Fog-enabled sensor-cloud architecture, which is based on the architecture used for obtaining Se-aaS. This architecture will help in improving service to the end-users, by using fog nodes and decreasing latency. This section illustrates the architecture of the Fog-enabled Sensor Cloud, having a 4-tier architecture, as shown in the figure.

Physical layer: The physical layer of sensors comprises a heterogeneous pool of sensors. These sensors are registered by their sensor owners. The fog node devices and the cloud resources access the sensor data through virtualization when required by applications.

The set of physical sensors registered to the infrastructure is denoted by S, $S = \{S_1, S_2, S_3, ..., S_x\}$, where any sensor $S_i \in S$ as $(1 \leqslant i \leqslant x)$ where $x \in \mathbb{N}$.

Fog layer: This layer has been introduced into the existing sensor-cloud infrastructure to reduce the latency and induce quick responses for specific applications. It is an extension of the cloud. The fog allows access to the shared resources in the fog layer. In contrast to the cloud, the fog layer is decentralized. Fog nodes, present in the fog layer can communicate among themselves, store data and compute. Thus, storage, computation and network connectivity are three functions of fog nodes. They process data among themselves and minimizes the response time, thus improving the service to the user.

Sensor-Cloud layer: This layer processes most of the data. Most of the end users access this layer through the web portal and is alloted a virtual machine. Virtualization of sensors takes place in this layer and VS groups are alloted to each application, according to their requirements.

End-users:The end-users have one or multiple applications for which they request the use of sensors through the Web portal. They specify all the necessary information. Based on this specification, the request is sent to the fog or the cloud, for processing. For time-critical applications, the fog nodes interact with the users directly, to give faster responses.

The set of end users is denoted by E, $E = \{E_1, E_2, E_3, ..., E_m\}$, where any end user $E_i \in E$ as $(1 \leqslant i \leqslant m)$ where $m \in \mathbb{N}$.

The set of applications for a single end user is denoted by A_i, $A_i = \{A_{i_1}, A_{i_2}, A_{i_3}, ..., A_{i_n}\}$, where $(1 \leqslant i \leqslant m)$ and $m, n \in \mathbb{N}$.

The end-user logs into the web interface to use the services of the FCS. The web interface works as a link between the applications and the FCS infrastructure.

IV. Proposed Model

In order to increase end-user satisfaction, we improve the sensor-cloud architecture by proposing to introduce a fog layer, closer to the physical sensor nodes.

A fog node, when requesting for sensed data from a sensor node, sends a few details about the node at that time instant. The sensor node uses this data to determine the most appropriate node to send the data to. According to this value, the fog nodes are serviced in a priority sequence.

A. Design

When the fog node devices require to sense data, the memory storage, load average, the CPU core temperature and the battery of the device at that instant, are recorded. These data are processed to acquire a single value, which helps in determining its priority factor in case there are multiple fog nodes requesting the same sensor.

We denote a prompt service, represented as $R1(S_n)$. The service can be expressed as

$$R1(S_n) = f_n\{F_{n1}, F_{n2}, ...F_{n\alpha}\} \tag{1}$$

where F_{nj} refers to the j^{th} fog node that requests the service of the node S_n; $f_n()$ is a function that aggregates the parameters of the fog node and maps them for the service of S_n.

Lemma 1: At any time instant T, the mapping R1(.) from sensor nodes to the set of fog nodes is dependent on 5 factors of the fog node device at that instant.

Proof. □

Proposition 1: The mapping R1(.) from the set of available physical sensor nodes S_m at any time instant t_n, to the set of fog nodes that have requested the sensor data, F_α is many-to-one.

The codomain of R1 is F_n and the set of physical sensor nodes is S_n. Since the fog nodes request data from multiple sensors required for application A_i, every sensor node, S, chooses from the F_α^S based on 5 parameters of the fog node device.

B. Mathematical Model

The set of fog nodes is denoted by F, $F = \{F_1, F_2, F_3, ..., F_n\}$, where any fog node $F_i \in F$ as $(1 \leqslant i \leqslant n)$ where $n \in \mathbb{N}$.

There are certain parameters that we consider for each fog node device.

- The total storage space available in the fog node device is denoted by M_{fi}^{ts}. It is expressed in MB.
- The free space available in the storage, at time instant t_n, is denoted by M_{fi}^{s}. It is expressed in MB.

$$M_{fi}^{s} = M_{fi}^{ts} - M_{fi}^{s_{use}} \tag{2}$$

The free memory factor of the secondary storage is given by:

$$MF_{fi}^{s} = \frac{M_{fi}^{s}}{M_{fi}^{ts}} \tag{3}$$

- The primary memory available in the fog node, is denoted by M_{fi}^{tr}. It is expressed in MB.
- The free space available in the primary storage, at time instant t_n,is denoted by M_{fi}^{r}. It is expressed in MB.

$$M_{fi}^{r} = M_{fi}^{tr} - M_{f}^{r_{use}} i \tag{4}$$

The free memory factor of the primary storage is given by:

$$MF_{fi}^{r} = \frac{M_{fi}^{r}}{M_{fi}^{tr}} \tag{5}$$

- The load average, denoted by LA_{fi}, is an array denoting the average number of processes being executed in the fog node for the last t_n amount of time.
- The number of processes running at time instant t_n is denoted by n_p.
- The CPU core temperature at a particular time instant t_n, is denoted by T_{core}.
- The power of the fog node device, at time instant t, is denoted by PB. It is expressed as a percentage.
- The distance between each physical sensor node, j, and fog node, i, that accesses it, is denoted by D_{ij}.
  Let (x_s, y_s) be the longitude and latitude of the physical sensor node and (x_f, y_f) be the longitude and latitude of the fog node. The Euclidean distance between the i^{th} fog node and the j^{th} sensor node is given by:

$$D_{ij} = \sqrt{(x_f^i - x_s^j)^2 + (y_f^i - y_s^j)^2} \tag{6}$$

***Definition* 1.** *The fog layer is responsible for rendering services to the time-critical applications.*

***Definition* 2.** *Total free memory factor (M^f): We compute total free memory factor in terms of the primary and secondary memory storage capacity, which provides the total amount of free space available in the fog node device at time instant t_n.*

$$MF_{fi} = MF_{fi}^{s} + MF_{fi}^{r} \tag{7}$$

***Definition* 3.** *Load factor (L): We compute load factor in terms of the load average of the fog node device till a time instant, t_n, and the number of processes currently being executed (n_p) at that time instant.*

The load average after a time t_n, is given by:

$$LA_{fi} = \frac{\sum_{i=1}^{t} S_i}{t_n} \tag{8}$$

$$L_{fi} = \frac{LA_{fi}}{n_p} \times 100 \tag{9}$$

***Definition* 4.** *Temperature factor (T): We compute temperature factor in terms of the current CPU core temperature (T_{core}) and the range (T_{range}) of safe temperature of the fog node device.*

Let the maximum and the minimum temperatures allowed on the fog node device be T_{max} and T_{min}.

$$T_{range} = T_{max} - Tmin \tag{10}$$

$$T_{fi} = \frac{T_{core} - Tmin}{T_{range}} \tag{11}$$

Lemma 2: The temperature factor of a fog node at a particular time instant t, lies between [0,1].

Proof. We know that,

$$T_{min} \leq T_{core} \leq T_{max} \tag{12}$$

Thus,

$$0 \leq T_{core} - T_{min} \leq T_{max} - T_{min} \tag{13}$$

or,

$$0 \leq T_{core} - T_{min} \leq T_{range} \tag{14}$$

Dividing by T_{range},

$$0 \leq \frac{T_{core} - T_{min}}{T_{range}} \leq 1 \tag{15}$$

Hence, $0 \leq T_{fi} \leq 1$

□

***Definition* 5.** *Power factor (PB_{fi}): It is the remaining battery power of the fog node device, fi, at the time instant t_n.*

C. Fog node device selection by Sensor nodes

In the Fog-enabled Sensor cloud infrastructure, the end users can access the sensor data through the fog nodes for faster processing and quicker responses. Due to multiple fog nodes requiring the same sensor data at time instant, t_n, the physical sensor node, let S_x has to select the most appropriate device from the set of fog nodes, $\{F_1^x, F_2^x, ...F_\alpha^x\}$, requesting data from it.

The values of the parameters of these fog node devices at time instant, t_n are used to calculate the decision factor, β.

$$\beta_{fi} = c \times MF_{fi} \times \frac{1}{L_{fi}} \times \frac{1}{T_{fi}} \times log(PB_{fi}) \times \frac{1}{D_{ij}} \tag{16}$$

where c is a constant.

This factor is calculated in the fog nodes itself and is sent when requesting for the data from the sensor node.

1) Use of Optimal Portfolio Selection strategy: In order to optimize the effectiveness of fog node selection, we use the theory of Optimal Portfolio Selection. This strategy helps to calculate the expected risk of each node, based on the expected energy required and the profit the sensor node at that instance of time.

We assume that the sensor nodes have been in use for D days and each day is divided into t_n time slots.

***Definition* 6.** *Expected energy cost of a sensor node: When data has to be transmitted from sensor to fog node, the total energy used in sensing the data and transmitting it to the fog node device is represented by E_{used}^t.*

Therefore, the expected energy used for sensor node, i, for the next data transmission is expressed as:

$$E_{used}^{i,exp} = \frac{\sum_{t=1}^{T} E_t^i}{T} \tag{17}$$

***Definition* 7.** *Expected service return: Let $S_t^{(i,j)}$ be the number of times a fog node, j has processed data from sensor node, i, till time t. Therefore, expected frequency of service to that particular fog node, j, is expressed as:*

$$S_t^{exp} = \frac{\sum_{t=1}^{T} S_t^{(i,j)}}{T} \tag{18}$$

***Definition* 8.** *Expected Computation delay of fog node: For a particular process, P_n and fog node, j, let the time instant at which it receives the sensor data be t_a and the instant at which all the data has been processed be t_p.*

Thus, the time consumed in processing the data can be expressed as:

$$CD_{P_n}^j = t_{P_n}^p - t_{P_n}^a \tag{19}$$

Therefore, expected computation delay of the fog node is expressed as:

$$CD_j^{exp} = \frac{\sum_{i=1}^{n} CD_{P_i}^j}{n} \tag{20}$$

***Definition* 9.** *The corresponding risk factor is computed by considering the relative computation delay and the service return of two fog node devices, which demonstrates the variation between expected and the actual values of computation delay and service return.*

The expected risk at time slot, t, is expressed as:

$$R_t^{exp} = \sigma_{ij}^{CD} \times \sigma_{ij}^{S} \times E_{used}^{i,exp} \times E_{used}^{j,exp}, i \neq j \tag{21}$$

where σ_{ij}^{CD} is the covariance of the computation delay between the i^{th} and j^{th} time slots. and σ_{ij}^{S} is the covariance of the service return between the i^{th} and j^{th} time slots. Mathematically,

$$\sigma_{ij}^{CD} = \frac{\sum_{t=1}^{T}[(CD_i^t - CD_i^{exp,t})(CD_j^t - CD_j^{exp,t})]}{T} \tag{22}$$

and

$$\sigma_{ij}^{S} = \frac{\sum_{t=1}^{T}[(S_t^i - S_t^{exp,i})(S_t^j - S_t^{exp,j})]}{T} \tag{23}$$

Therefore, the total expected risk is as follows:

$$R_t^{exp} = \sum_{j=1}^{T}\sum_{i=1}^{T} \sigma_{ij}^{CD} \times \sigma_{ij}^{S} \times E_{used}^{i,exp} \times E_{used}^{j,exp}, i \neq j. \tag{24}$$

Now, the fog node is chosen by a sensor node according to the final factor, ϕ.

$$\phi_i = \beta_i - R_t^{exp} \tag{25}$$

We maximize the final factor, ϕ_i i.e.,

$$\begin{aligned} c \times MF_{fi} \times \frac{1}{L_{fi}} \times \frac{1}{T_{fi}} \times log(PB) \times \frac{1}{D_{ij}} \\ - \sum_{j=1}^{T}\sum_{i=1}^{T} \sigma_{ij}^{CD} \times \sigma_{ij}^{S} \times E_{used}^{i,exp} \times E_{used}^{j,exp} \\ , i \neq j. \end{aligned} \tag{26}$$

We apply the Lagrangian function on Equation 26, which is expressed in Equation 27.

Theorem 1. *The time required for the sensor nodes to transmit data to the fog nodes is less than the time required to transmit data to the cloud.*

Proof. Let λ_s be the time instant at which the data packet leaves the sensor node, λ_f be the time instant at which the data packet is received by the fog node and λ_c be the time instant at which the same data packet is received by the cloud. □

V. CASE STUDY

Agriculture

A. Implementation

Hardware

$$\begin{aligned} L_\mu = \frac{c \times MF_{fi} \times log(PB_{fi})}{L_{fi} \times T_c^{fi} \times D_{ij}} - \Big[\sum_{j=1}^{T}\sum_{i=1}^{T} \sigma_{ij}^{CD} \times \sigma_{ij}^{S} \times E_{used}^{i,exp} \times E_{used}^{j,exp}\Big] - \mu_1(MF^{max} - MF_{fi}) - \mu_2(1-L) - \mu_3(T_{max} - Tmin - T_c^{fi}) \\ -\mu_4(PB_{fi} - PB^{thres}) - \mu_5(D_{ij}^{thres}) - \mu_6(E_{used}^{i,exp}) - \mu_7(E_{used}^{i,exp}) \end{aligned} \tag{27}$$

B. Experimental Results

Results - Time and energy

C. Result analysis

Graphs - compare fog-enabled vs cloud - time and energy

VI. Performance Evaluation

compare with theoretical values

***Definition* 10.** *The time efficiency of the fog node is a function of the time instant at which the end user demands its use, λ_{ef}, the time instant at which a request is sent to sensors for data, λ_{fs} and when the data is received by the node λ_{sf}, the time required to process the data, λ_p and the time instant at which data is sent back to the end user, λ_{fe}.*

$$\eta_i = \frac{\lambda_p}{\lambda_{fe} - \lambda_{ef} + \sum_{i=1}^{\gamma}(\lambda_{sf} - \lambda_{fs}) + \lambda_p} \tag{28}$$

VII. Conclusion