UNIVERSITY OF CALIFORNIA,
IRVINE

Energy-Efficient Real-Time Heart Monitoring on Internet-of-Medical-Things

THESIS

submitted in partial satisfaction of the requirements
for the degree of

MASTER OF SCIENCE

in Computer Engineering

by

Islam Bayoumy

Thesis Committee:
Associate Professor Mohammad Al Faruque, Chair, Chair
Associate Professor Haithem Taha
Assistant Professor Yasser Shoukry

2021

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I want to thank Professor Mohammad Al Faruque for his continuous support and valuable feedback and guidance from the inception of this work till its completion.

Special thanks and appreciation to my colleagues at ACIPS lab for the fruitful collaboration that made this work possible.

Thanks to my review committee, Professor Haithem Taha and Professor Yasser Shoukry, for their support throughout this work.

ABSTRACT OF THE THESIS

Title of the Thesis

By

Islam Bayoumy

Master of Science in Computer Engineering

University of California, Irvine, 2021

Associate Professor Mohammad Al Faruque, Chair, Chair

The de facto standard in machine learning architecture is to rely on cloud computing to do all the needed data processing from Learning to classification, while edge nodes or IoT (Internet of Things) devices provide the raw data without formatting or filtering. Although this approach shows its efficacy in many areas, it suffers lots of drawbacks. Such as privacy, massive data/storage management, and inefficient power management in limited resources edge nodes, which is our dissertation's concern.

We are listing and testing different architectures that can improve edge node power consumption and efficiency in data handling, enhancing privacy, and data management on cloud computing. We have applied different architectures to the Internet of Medical Things (IoMT) field, which drastically improves real-time monitoring and recording of electrocardiogram (ECG) signals. We have used various communication protocols to show the impact on overall power consumption. Also, we have presented results that can guide architecture and communication selection according to application usage.

Our experiments and profiling were done on actual targets from IoT market. We used Bluetooth Low Energy (BLE) Nordic nRF52 Development Kit and SLWSTK6020B with EFR32BG Bluetooth SoC Starter Kit from Silicon Labs.

The work done in this thesis is also included in the following paper that is currently under review.

J48. B. U. Demirel, I. A. Bayoumy, M. A. Al Faruque, "Energy-Efficient Real-Time Heart Monitoring on Edge-Fog-Cloud Internet-of-Medical-Things", published in the IEEE Internet of Things Journal, 2021

Chapter 1

Introduction and Background

1.1 Introduction and Background Studies

The need for Edge (also can be called IoMT[1] Devices to live long with minimal human intervention dictates that those devices be low power and hence limited processing power. Cloud Computing is considered the natural extension of Edge devices with virtually unlimited processing power to offload data analysis, data storage, and decision making. With the benefits provided by cloud computing, its drawbacks are eminent. In this work, we describe the different architectures for cloud and edge devices and possible architectures variations that show the balance in power and performance based on the application of choice. Our show case algorithms are ECG signal classification running on Edge node, the purpose is to detect and to keep updating Fog layers with normal heart beats, but sends full capture of two minutes' worth of ECG data capture when Arrhythmia is detected. This approach adds intelligence to Edge node instead of just acting as data relay.

[1]IoMT or Internet of Medical Things. The Edge nodes in the realm of Medical Devices that are connected to healthcare IT systems to communicate collected health signals.

1.2 Cloud Computing

Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction [24].

1.2.1 Cloud Characteristics

On-demand self-service: A consumer can unilaterally provision computing capabilities, such as server time and network storage, as needed automatically without requiring human interaction with each service provider.

Broad network access: Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms (e.g., mobile phones, tablets, laptops, and workstations).

Resource pooling: Examples of resources include storage, processing, memory, and network bandwidth.

Rapid elasticity: Capabilities can be elastically provisioned and released, in some cases automatically, to scale rapidly outward and inward commensurate with demand. To the consumer, the capabilities available for provisioning often appear to be unlimited and can be appropriated in any quantity at any time.

Measured service: Resource usage can be monitored, controlled, and reported, providing transparency for both the provider and consumer of the utilized service.

1.3 Edge Computing

With the rise of IoT field in the last decade, cloud computing has emerged as accompanied technology to address IoT devices' limited storage and processing power. With this evolution of cloud technology and the vastly increasing number of IoT devices, achieving low latency in applications has become a challenge. Edge computing was introduced to address some of the drawbacks of cloud computing by moving part or sometimes entire decision-making closer to the data source [15].

Edge devices also can be called IoT devices, are the limited power and processing device that are closer to the data source. These devices have plenty of constraints of course it is dependent on the type of application that those devices serve.

Some can be power plugged so power constraints won't be an issue such as surveillance cameras, some are more power sensitive and hence processing power limited to conserve battery life, such as health monitoring devices, and wearable.

Our interest is geared toward wearable and health monitoring devices and hence the limitation in power and processing.

1.3.1 Edge Characteristics

- Provide a break stage for cloud layer centralized data processing, so calculation becomes decentralized across layers [39].
- Decreases the volume of data transferred across the network as data processing, and filtering are done before transferring, alleviating network congestion.
- Reduces network latency.

With Edge node provided advantageous, more security threats introduced as the real-world attack surface increase. One reason that Edge nodes are more vulnerable to existing security threats is their limited processing power, so some security algorithms are not suited to run on Edge[40][34].

As discussed later, the main contributors is consuming device battery is computation and communication. Edge devices have advantageous of being closer to the data source which helps minimizing the communication power.

1.4 Edge and Cloud Integration

As mentioned earlier, Edge devices are small things with limited storage capacity, energy and processing power and cloud computing have virtually unlimited storage capacity and processing power, so both technologies complement one another.

The result of this integration opened doors to a plethora of applications in different fields, such as healthcare, where continuous monitoring, report and track patient's vital signs are possible. Smart grids where tracking and predicting electricity consumption for better load balancing made easy.

1.5 Edge and Cloud Integration Motivations

Although integration is beneficial in numerous ways, its side effects need novel methodologies to address or mitigate them. A few of the pronounced side effects are listed below:

- **Security:** With this massive number of connected wireless devices (30 Billion by 20201. Security becomes one of the major concerns.

- Examples of recent security vulnerabilities:
  * BlueBorne: Affected 8 billion Bluetooth enabled devices in 2017.
  * KRACK (Key Reinstallation Attacks): Affected all Wi-Fi enabled devices by breaking WPA2.

- **Big Data:** With an estimated number of 30 billion devices that will be connected by 2021, all producing data, storage, and processing of this massive amount of data become a hurdle.

- **Latency introduced by cloud offloading:** offloading processing and decision making to the cloud has its benefits but introduces additional latency, which might be inadequate for real-time sensitive applications.

- **Power consumption:** Although offloading to the cloud extended the limited devices' processing and storage capabilities, power consumption still a significant concern since edge nodes need to communicate with the cloud to send the collected data and back to act upon the taken decision. This communication consumes the most significant portion of the device's power budget.

1.6 Fog and Mist Computing

Fog computing is such an intermediate layer extending the Cloud layer. The Fog Computing layer brings computing, network, and storage services closer to the end-nodes in IoT. Compared to Cloud Computing, this computing layer is highly distributed and introduces additional services to end-devices located in the perception layer. [30]

Some resources omit the mention of Mist Computing altogether and refer to any computational layer between Edge and Cloud as Fog layer [37]. Other resources such as [24] provide more granularity based on the closeness of the extension computational layer; hence Mist Computing layer emerges.

Mist Computing is a Fog layer with similar characteristics, but the only differentiation from Fog is its closer to Edge nodes. So this Mist Layer can be a layer of smartphones that are not as powerful as server farms but are powerful compared to limited Edge nodes[32]. [24]Referred to Mist Computing as a lightweight and rudimentary form of fog computing.

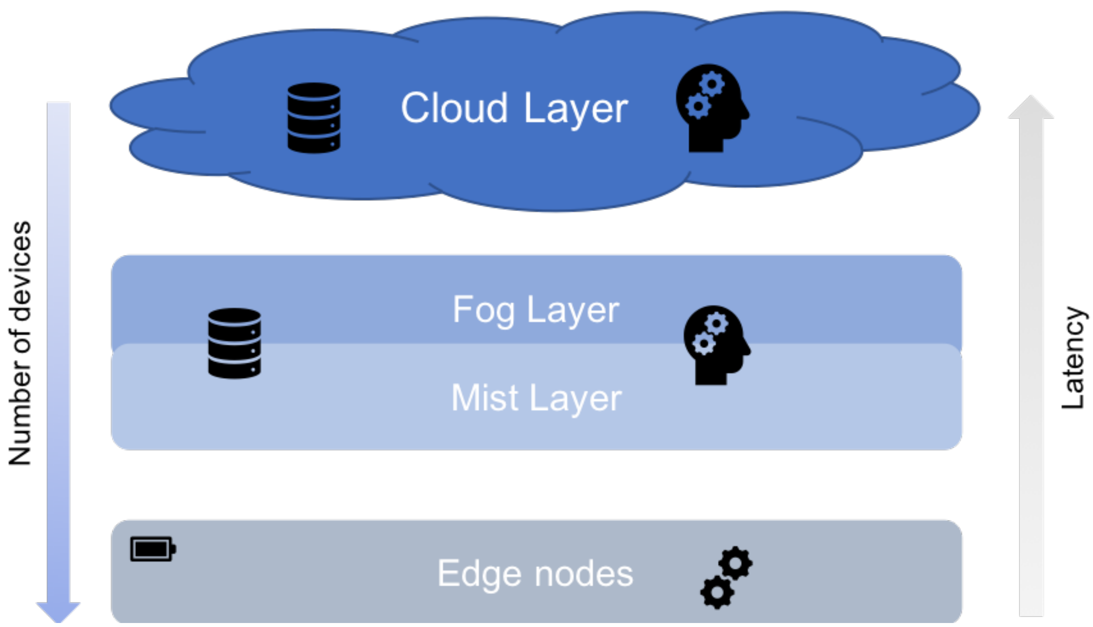


Figure 1.1: Latency and proximity relationship of Cloud, Fog, Mist and Edge

A conceptual relationship between Cloud, Fog, Mist, and Edge layers is depicted in Figure 1.1. We can see that Cloud Layer is furthest from the data source and Edge nodes. Also, Cloud Layer has the highest latency of data exchange but the least number of devices compared to all other layers with the most powerful machines.

Fog and Mist Layers are intermediate between Edge nodes and Cloud Layer; both inherit Cloud Layer characteristic but lower in data exchange latency, and more devices include but with less capability. Mist Layer is considered a lightweight Fog Layer that is closer to Edge nodes. We will use the Fog Layer term for simplicity in this work, but we imply this layer is closer to the Edge node since we use short-range communication protocols.

Edge nodes are the highest number of devices in the network with the least resources available and usually battery-powered. Adding intelligent algorithms for data processing instead of data relay is one of the added values of this work coming in the following few chapters.

1.7 Short-Range Communication Protocols

Amongst wireless data standards families, Wireless Wide Area Network (WWAN), Wireless local area network (WLAN), and Wireless Personal Area Network (WPAN), our focus is on the WPAN family (IEEE 802.15 working group), which is geared toward short-range communication[9].

From WPAN standards we focus on Zigbee, BLE, and UWB. The purpose of choosing a short-range communication protocol is that Fog Layer is closer to the network edge so which permits short-range communication and hence less power consumed. [20]

Zigbee protocol is built on top of IEEE 802.15.4 standard with low power consumption (Lowest power consumption compared to UWB and BLE) and, at the same time, low data rate up to 250 kbps. Zigbee unlicensed Industrial, Scientific, and Medical (ISM) frequency band. This makes Zigbee an ideal communication protocol for low data rate applications, but its support is not widely supported in smartphones[29].

Ultra-wideband (UWB) protocol is also built on top of IEEE 802.15.4 standard but works on unlicensed bands from 3 to 10 GHz, making it immune to interference and no problems with coexistence with other communication protocols. It is not yet widely spread supported in smartphones, but it is gaining ground since 2019 in new smartphones. Current use cases focus on ranging rather than data exchange. UWB power consumption is considered the highest compared to Zigbee and BLE[29], [9],[7].

Bluetooth Low Energy (BLE) protocol is developed and maintained by Bluetooth SIG (Special Interest Group); it uses the ISM frequency band and uses the frequency hopping technique to overcome interference. The supported data rate varies from 125 Kb/s to 2Mb/s starting from Bluetooth v5.0 specification, and it is designed for scratch low power devices [36].

Later in our work, we compared our selected short-range to the long-range communication protocols for architecture selection.

Chapter 2

Real-time Energy Efficient Heart Monitoring on Edge

One of the applications that shows the importance of edge computing is the applications of ELECTROCARDIOGRAM (ECG) signals to detect cardiovascular diseases.

2.1 Importance of ECG Signal

As per the World Health Organization (WHO) reports, cardiovascular heart diseases are major cause of mortality in the total population. WHO evaluated that 30% of world population deaths result from cardiovascular heart diseases and expected 23.6 million will endure to these diseases by 2030 [35]

There are multiple methods to detect cardiovascular diseases; most common methods listed below [3]:

- Electrocardiography.
- Imagining techniques.
- Immuno techniques.

Widely used by clinicians as a routine modality in hospitals, ECG recordings capture the propagation of the electrical signal in the heart from the body surface [1].

Electrocardiography hence ECG is the most suitable for wearable devices with good room of improvement to make it accessible to large number of people[11], [25].

One of the most popular methods for capturing ECG signal is using Holter [19], which continuously records the signal over long periods for 24 to 48 hours or sometimes longer where standard 12-lead 'sometimes 5-lead' ECG provide information on cardiac activity.

Although this method is widely used, it does not provide real-time feedback to users, and recorded results need offline cardiologists analysis, which is expensive and time-consuming.

2.2 Improvements

As we could observe the presented method lacks real-time feedback which when addressed; can reduce response time for intervention in infarct size control or resuscitation of sudden cardiac-death victims [28].

2.2.1 Cloud Based Solution

In this approach, ECG device relies on cloud servers to analyze and classify the acquired signals, but this is done offline. However, it is faster and cheaper than Holter's approach, which requires a cardiologist's analysis. Even though the improvements are significant, but they did not close the gap to achieve real-time analysis to allow fast response time to save patients' lives [13] [5].

2.2.2 Edge Based Solution

As presented in [38] and [14] This solution provides the needed real-time response by performing more calculation on the edge nodes but lacks optimization for edge node limitations of battery life and processing power. Although it works, it will provide a limited device life span as it depletes Edge node battery, which is the most critical resource for any battery-powered Edge node [2] hence introducing new limitations to the edge device.

2.3 Existing Work

An optimized Edge solution requires some novel approaches to overcome Edge nodes' memory, processing power, and battery life limitations. On the other hand, a competitive Edge solution needs to overcome Cloud Layer introduced latency to provide real-time feedback. Some existing work such as [27], [32] explored bringing the computation closer to the Edge nodes, using Fog Layer. This work required high traffic over the network since data was sent in real-time, resulting in rapid battery depletion since the communication consumed most of the power budget.

In addition to being power-hungry, it lacks privacy as data sent in real-time without proper encryption[23], [31]. Although encryption can be added to those solutions, it will require additional memory and processing on already limited capability devices. That also reflect on battery life [21], [17].

2.4 Optimized Edge Solution

The key to providing an optimized Edge Solution can be summarized in the below points:

- Deliver resource-efficient algorithms that can run on the limited Edge nodes[33],[12].
- Bring the extension processing layer closer to the edge using a mix of Cloud, Fog, Mist Computing Layers.
- Minimize data exchange over a wireless network by detecting anomalies and choose that data to send.
- Choose an adequate communication protocol for the application that minimizes the allocated communication power budget.

Chapter 3

Experimental Setup

3.1 Edge Device Selection

In the IoMT field, there are plenty of options to choose Edge devices, so we needed to put criteria that serve our application best to narrow down the selection. Our main criteria are listed below:

3.1.1 Low Power Micro Controller Unit (MCU)

Our top priority in Heart Monitoring devices is to achieve the lowest possible power, so we need to minimize the computation power budget to the lowest possible. This leads to choosing a low-power MCU with enough processing capability and memory. Searching market offering, we landed on selecting ARM Cortex-Mx family or similar since it is designed for low-power versatile embedded applications.[4]

Figure 3.1 provides a pictorials comparison of ARM families, and below are the motivations behind choosing the Cortex-Mx family [18].

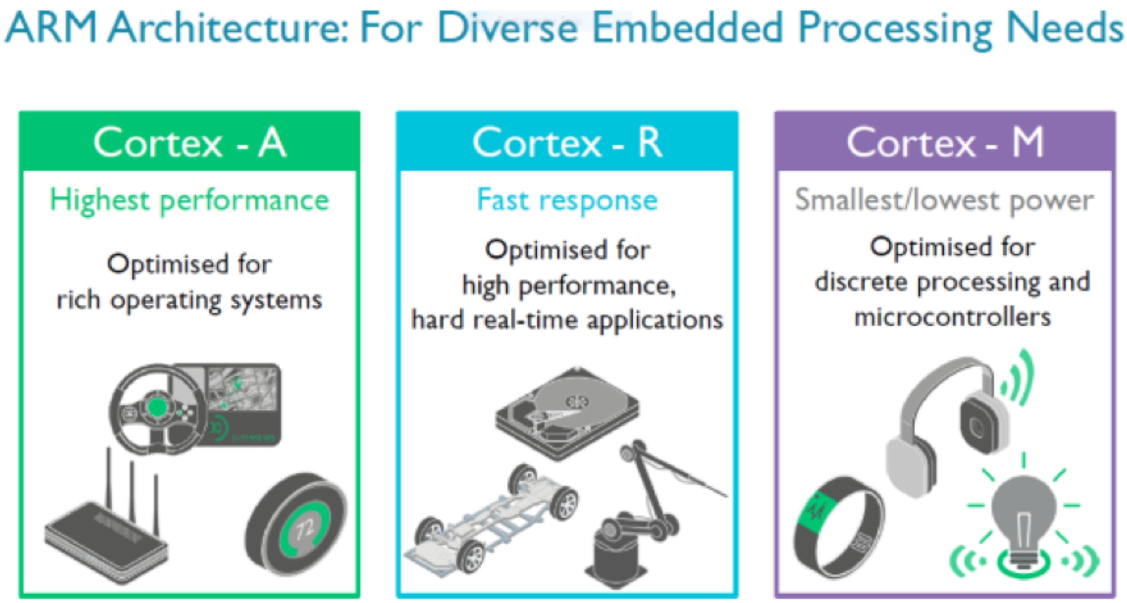


Figure 3.1: Comparison of ARM Processors Families Source:https://www.anandtech.com/show/10049/arm-announces-cortex-r8

- Cortex-A family requires more power than Cortex-M family and is designed for high-end applications that run operating systems like Android or Linux.
- Cortex-R family is designed around real-time applications with hard real time constraints such as automotive or hard disk drive controllers.

3.1.2 Short-Range Communication Protocol:

Since our architecture will utilize Fog Layer features, the processing layer is closer to the Edge node. So long-range data communications that inherently consume more power are not our first choice. 'later, we will compare different data communication protocols to show the impact of correct selection.'

We selected Bluetooth Low Energy (BLE) communication protocol since it is designed from scratch for low-power applications, and almost all smartphones widely support it in the market in contrast to Zigbee, which lacks smartphone support and has a lower data rate [6] and Ultra Wide Band (UWB) that is recently picking-up but its focus is ranging only at this point[10],[29].

3.1.3 Bluetooth Low Energy Support

In Bluetooth version 5 or up, the changes focus only on BLE part and Bluetooth classic remains unchanged. In the new BLE standard increased data rate from 1 Mb/s to 2 Mb/s at PHY level which increased the effective bandwidth by almost 2X. Also message size increased from 31 bytes to 255 bytes.

This should help our application since data exchange will be done faster at almost double the speed with more bytes per packet which decreases the time the device being active, i.e. decreases power consumption. Some other enhancements such as range extended to be 4X and advertising extension those enhancement are not directly affecting our application [36] [8].

BLE specifications differ in the range of supported features; comparing these features, BLE specification version 5.X was the best choice.

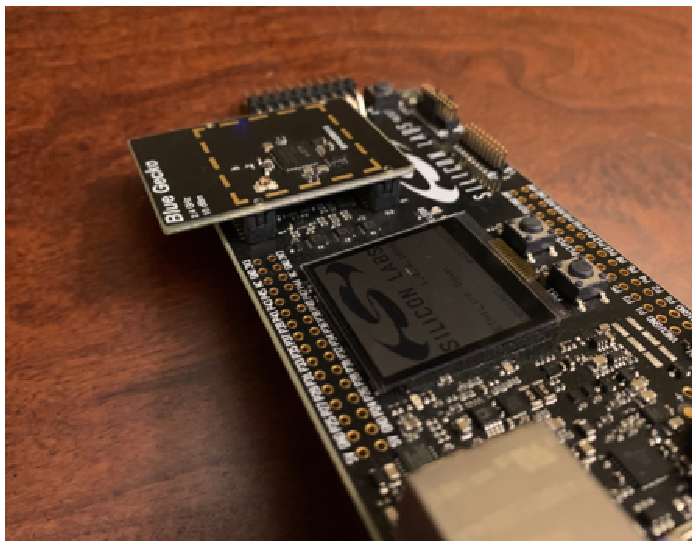
3.1.4 Small Module Area

Some BLE SoC offerings require a two-chip solution with MCU attached to BLE link controller; for example, ATBTLC1000 Xplained Pro from Atmel [22] has superior power consumption compared to other vendors System on Chips (SoCs) but requires additional MCU for the application layer.

Some other offerings have one chip solution with application and full BLE stack running on the same die without external MCU, which provides a more compact solution and less power consumed. We chose a one-chip solution for our case study.

Two platforms met all the above requirements:

EFR32BG13 Blue Gecko which met all of the above requirements, it is a single-die Bluetooth® Low Energy v 5.0 SoC which has 32-bit ARM Cortex-M4 core with 40 MHz maximum operating frequency.

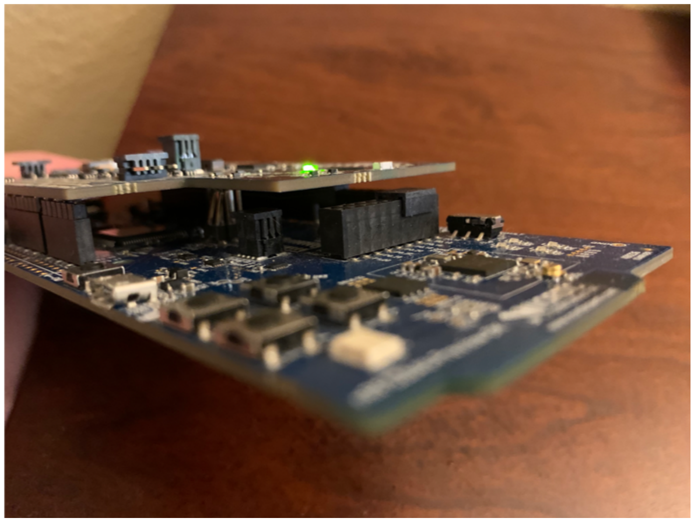


(a)

(b)

Figure 3.2: Blue Gecko Development Kit

nRF52840 from Nordic, it is a single-die Bluetooth ® 5, IEEE 802.15.4-2006, 2.4 GHz transceiver with ARM ® Cortex ® -M4 32-bit processor with FPU, 64 MHz operating frequency.



(a)

(b)

Figure 3.3: nRF52840 Development Kit

3.2 Experimental Setup Architecture

The conceptual architecture for Edge node, Fog Layer, and Cloud is as shown in Figure 3.4. Edge node, which in our case based on previous criteria either nRF52840 or EFR32BG13. The communication protocol between the Edge node and Fog Layer is BLE 5.0. Fog Layer is smartphones with applications collecting, analyzing, and communicating back with Edge node and Cloud Layer through LTE or Wi-Fi.

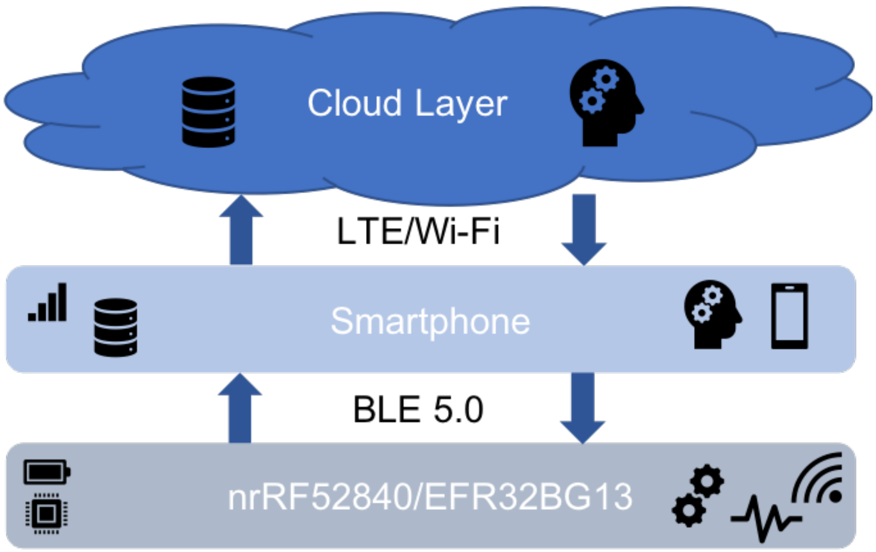


Figure 3.4: Experimental setup architecture

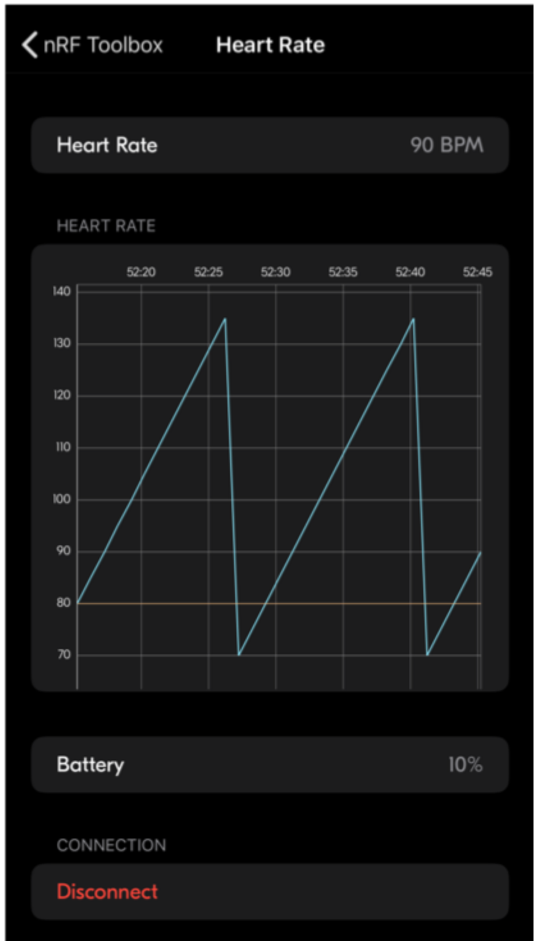
3.2.1 Edge Device Selection

We developed basic applications for both boards, bring-up, and quick evaluation.

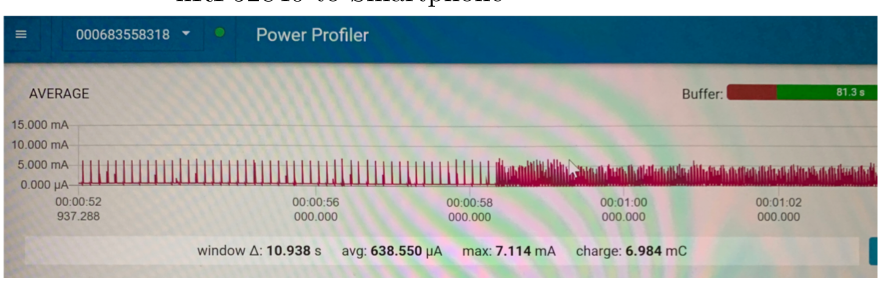
3.2.2 Nordic nRF52840:

We utilized Heart Rate Monitoring Profile to evaluate the setup and to bring up the power profiling setup. Also, we used Nordic smartphone app as a Fog Layer to periodically communicate the heart rate.

This setup and application were to mimic the Figure 3.4 and get used to Nordic profiling. The outcome is shown in Figures 3.5a and 3.5b.



(a) Heart Rate Monitor Profile from nRF52840 to Smartphone



(b) Heart Rate Monitor Profile power profiling at Advertising and Connection

Figure 3.5: nRF52840 Development

3.2.3 EFR32BG13 Blue Gecko

The amount of profiling on nRF52840 for basic profiling and application development was to help in selecting our target platform. We decided to select EFR32BG13 Blue Gecko as our reference platform since its capabilities are more limited than nRF52840. Hence, it will emphasize our showcase to demonstrate the entire operation of ECG classification and communication with the Fog Layer on the most suitable limited feature Edge node.

We developed a framework that accommodates our ECG Algorithms (More detailed on Algorithms in the next Chapter); this Framework, as in Figure 3.6, uses Blue Gecko APIs for BLE transactions that are built on top of Generic ATTribute profile (GATT) and Generic Access Profile (GAP). We used Silicon Labs Simplicity Studio IDE and its Energy Profiler.

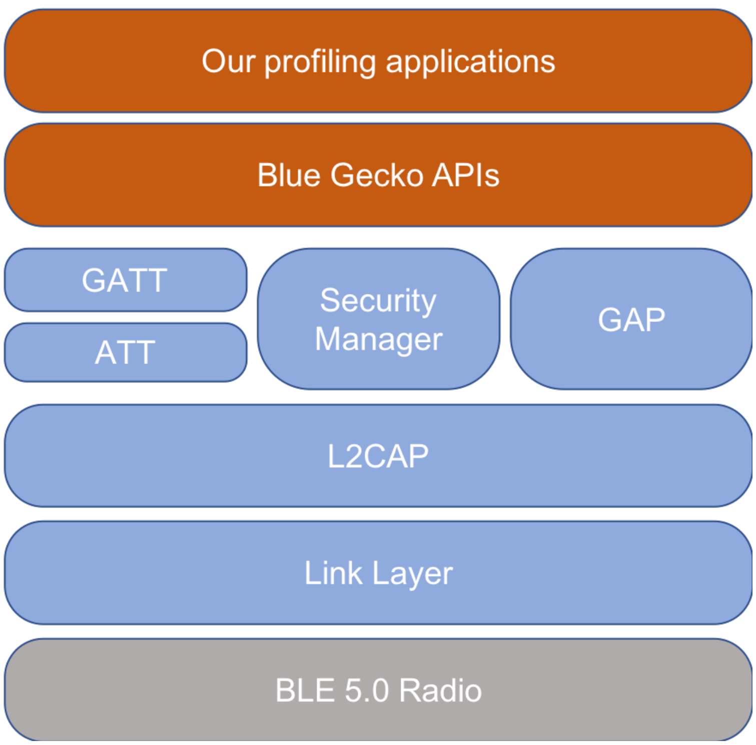


Figure 3.6: Framework architecture

Our Framework performed BLE data exchange as event-based triggered when smartphone (Our Fog Layer) requested data transfer; this is mainly for predictability of data transfer timing; otherwise, Edge node should send the ECG data once it is ready as notifications to Fog Layer.

A snapshot of the data transfer code section is shown in Figures 3.7, 3.8.

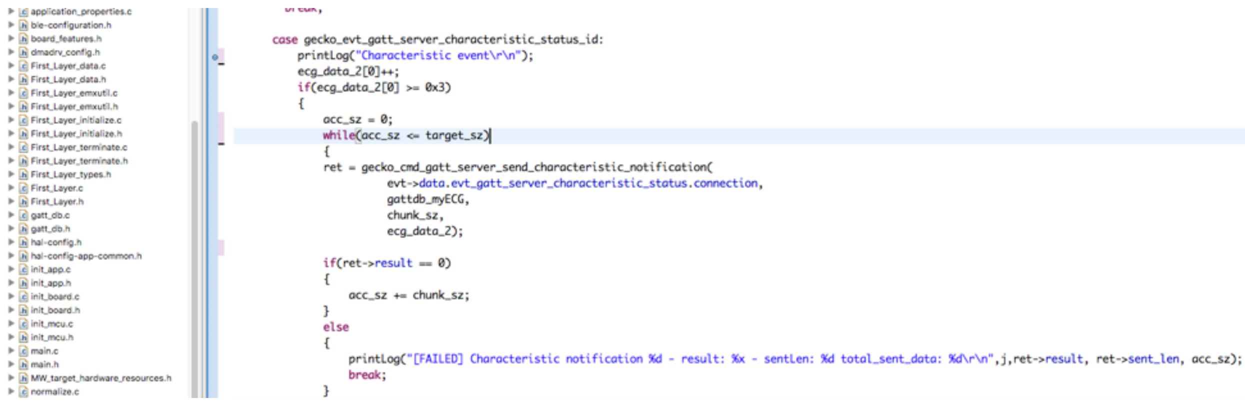


Figure 3.7: IDE Project

```
case gecko_evt_gatt_server_characteristic_status_id:
    printLog("Characteristic event\r\n");
    ecg_data_2[0]++;
    if(ecg_data_2[0] >= 0x3)
    {
        acc_sz = 0;
        while(acc_sz <= target_sz)
        {
        ret = gecko_cmd_gatt_server_send_characteristic_notification(
                  evt->data.evt_gatt_server_characteristic_status.connection,
                  gattdb_myECG,
                  chunk_sz,
                  ecg_data_2);

        if(ret->result == 0)
        {
            acc_sz += chunk_sz;
        }
        else
        {
            printLog("[FAILED] Characteristic notification %d - result: %x - sentLen: 9
            break;
```

Figure 3.8: Framework Code Snippet

Over the Air packets capture is shown in Figure 3.9; in this capture, our Edge node sends BLE packets to the smartphone over the GATT layer. The available sniffer we had was not accurate enough to show the timing or throughput, but we used it for basic packet identification.

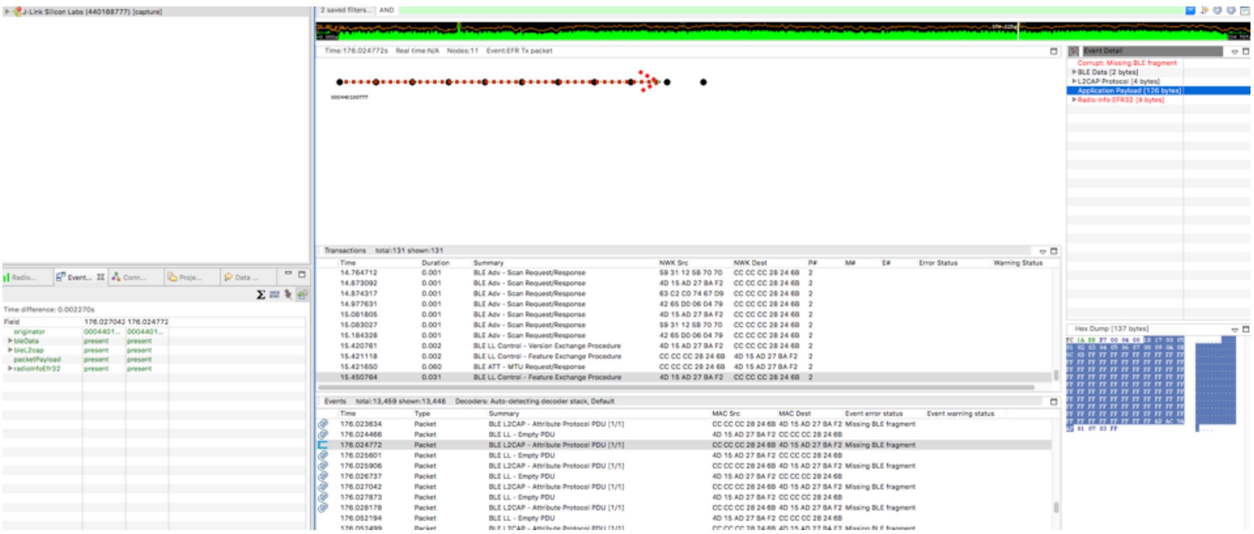


Figure 3.9: Over The Air packets capture

Chapter 4

Communication Power Consumption

As seen from Figure 1.1, there are multiple possible architectures that we can adopt to partition the computation among Edge nodes, Fog, and Cloud servers. The architecture we chose was based on minimizing power consumption on Edge nodes, and maximizing its independence from other servers to minimize latency and to serve the critical ECG real-time feedback better.

As mentioned in previous chapters, we consider two main factors to optimize in Edge nodes, Communication and Computation powers, while considering the target factors discussed above (real-time feedback, i.e., low latency, maximum power optimization for longer device life, and maximum computation accuracy).

4.1 Edge Node Communication

The growing concern in the IoMT era is energy consumption. Even mostly in low power mode, the 14 billion network-enabled devices currently in use worldwide waste 400TWh (Terawatt-hour) of electricity per year. With 50 billion devices in 2023, Power consumption can increase by up to 1400TWh. Data transfer from Edge to Cloud consumes energy, and this increases with the increase of the amount of data to send over to the cloud and how far the node is placed from the cloud base. [30].

[30] shows a conceptual comparison of delay, bandwidth, and power of Fog, Clod, and Mist computing.

The chart below depicts a logarithmic scale comparison for power, delay, bandwidth, and distance between Fog, Cloud, and Mist computing layers to better represent the differences.

From this chart, we can estimate how much optimization we can gain from bringing computation closer to the network's edge. The estimate improvement in power is about 700X and delay is about 600X.

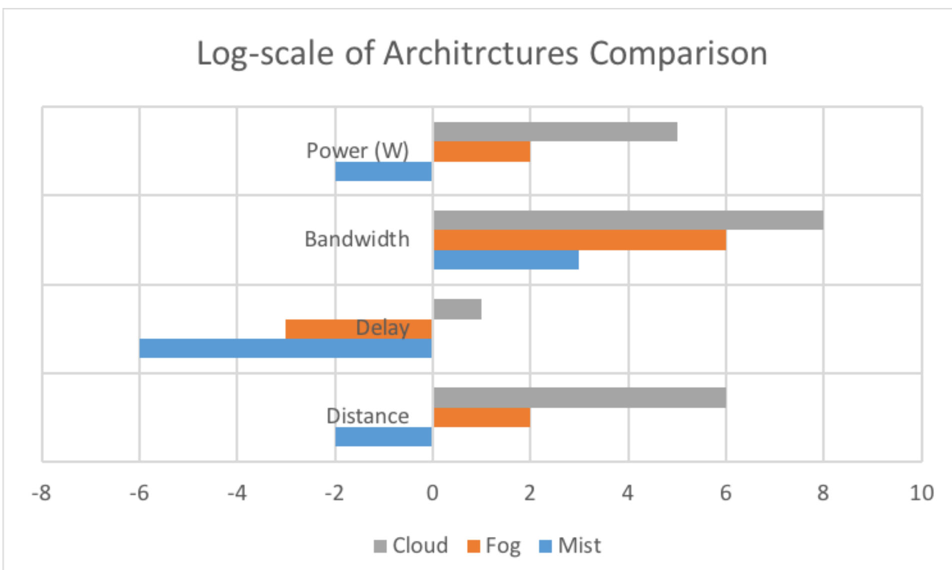


Figure 4.1: Log scale comparison of distance, power, and bandwidth between Cloud, Fog, and Mist layers

The below factors are to consider for edge node communication:

- Amount of transferred data and this should be at minimum.
- Distance from fog nodes that relay data to cloud for further processing and this distance

should be at minimum also, the closer the better. Distance also affects the architecture of choice, hybrid of mist and fog architecture is better suited to our application.

- Communication technology, such as Wi-Fi, LTE, 3G, BLE (Bluetooth Low Energy).

The architecture of choice will dictate the communication consumption contribution from the allocated power budget of the Edge device.

4.2 Communication Technologies Power Profiling

We chose LTE, 3G, Wi-Fi, and BLE communication technologies to consider for our case study. These technologies enable our edge node to work with different architectures from cloud, fog, and mist.

LTE and 3G provide long-range communication with relatively higher bandwidth for cloud-based architecture. In comparison, Wi-Fi provides a medium range of communication for Fog architecture, and finally BLE for a shorter range of communication, mainly for mist architecture for closer computation to the edge of the network.

4.3 BLE Data Exchange Power Profiling

As mentioned earlier, our algorithms need to transfer data to Fog Layer at three different categories as below. In all three cases, we performed the profiling for transmitting the data and the MCU activities and leakage over one second, and from that, we could estimate the consumed power over any period. In the following few sections, we are presenting the consumed power for different ECG algorithms.

4.3.1 Two Minutes Arrhythmia

Edge node collects 800 doubles or 6.25KB (double data type is stored as IEEE 64-bit equivalent to 8-byte of memory) of data worth sending to Mist/Fog nodes. We will call this an Arrhythmia case.

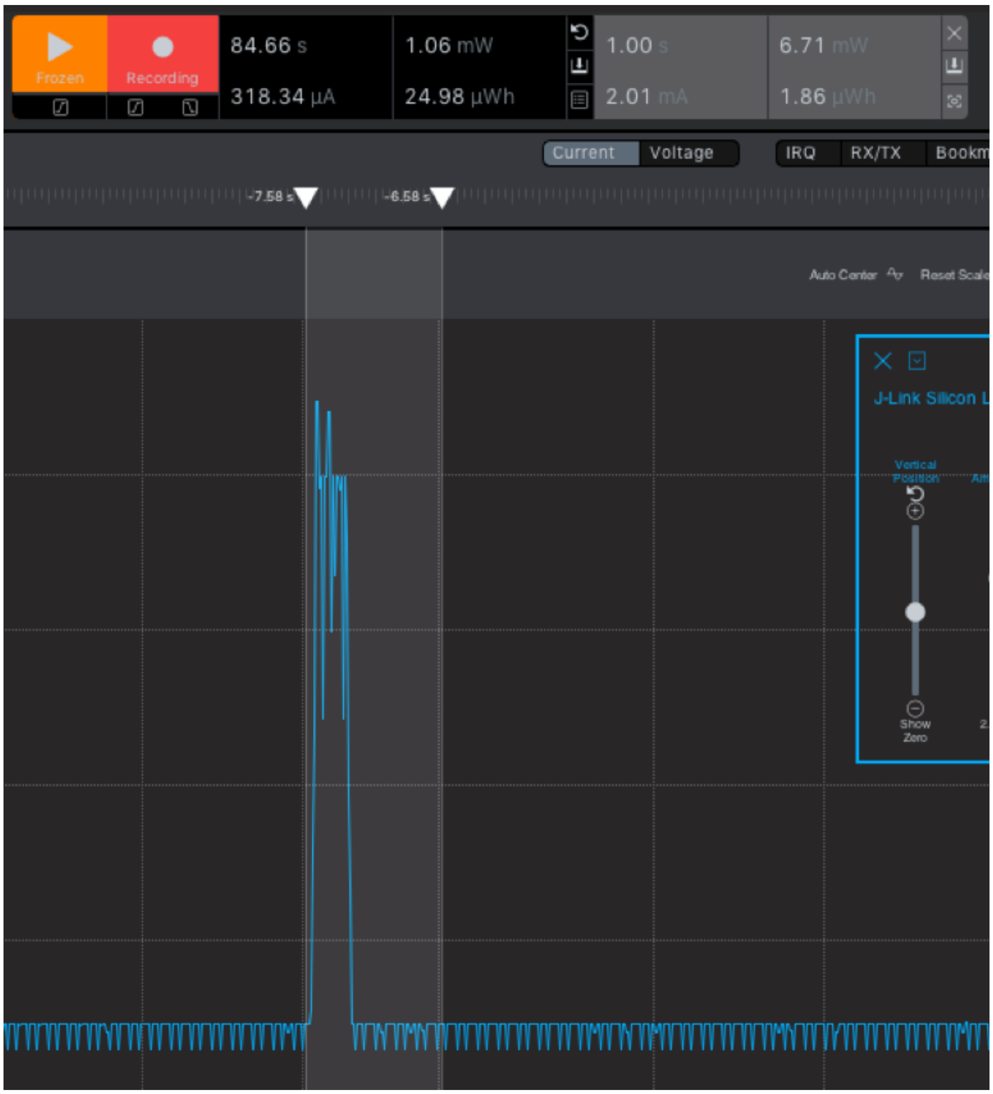


Figure 4.2: Two minutes Arrhythmia

Sending 800 doubles takes around 350ms to transfer on 2M PHY consuming ***5.25 mA***. And average over 1 second is around ***6.7mW*** as shown in Figure 4.2

4.3.2 Normal Beats

In this case, the chosen algorithm sends one integer or 4-byte worth of data transferred to Mist/Fog nodes. The profiling for this case is as below Sending only 4 bytes over to mist/fog

Figure 4.3: Normal beats case

layer average over 1 second takes around ***318uA***, and ***1.06mW*** as shown in Figure 4.3.

4.3.3 Transferring Raw ECG Signal

In this case, the chosen algorithm sends 360 double or 2.8125KB of data to Mist/Fog nodes. We will call it the Raw ECG case. The profiling of this case is as below: In this case, the Edge node acts as a relay for data, where nodes receive raw ECG data and pass it on to the Fog Layer for further processing and classification.

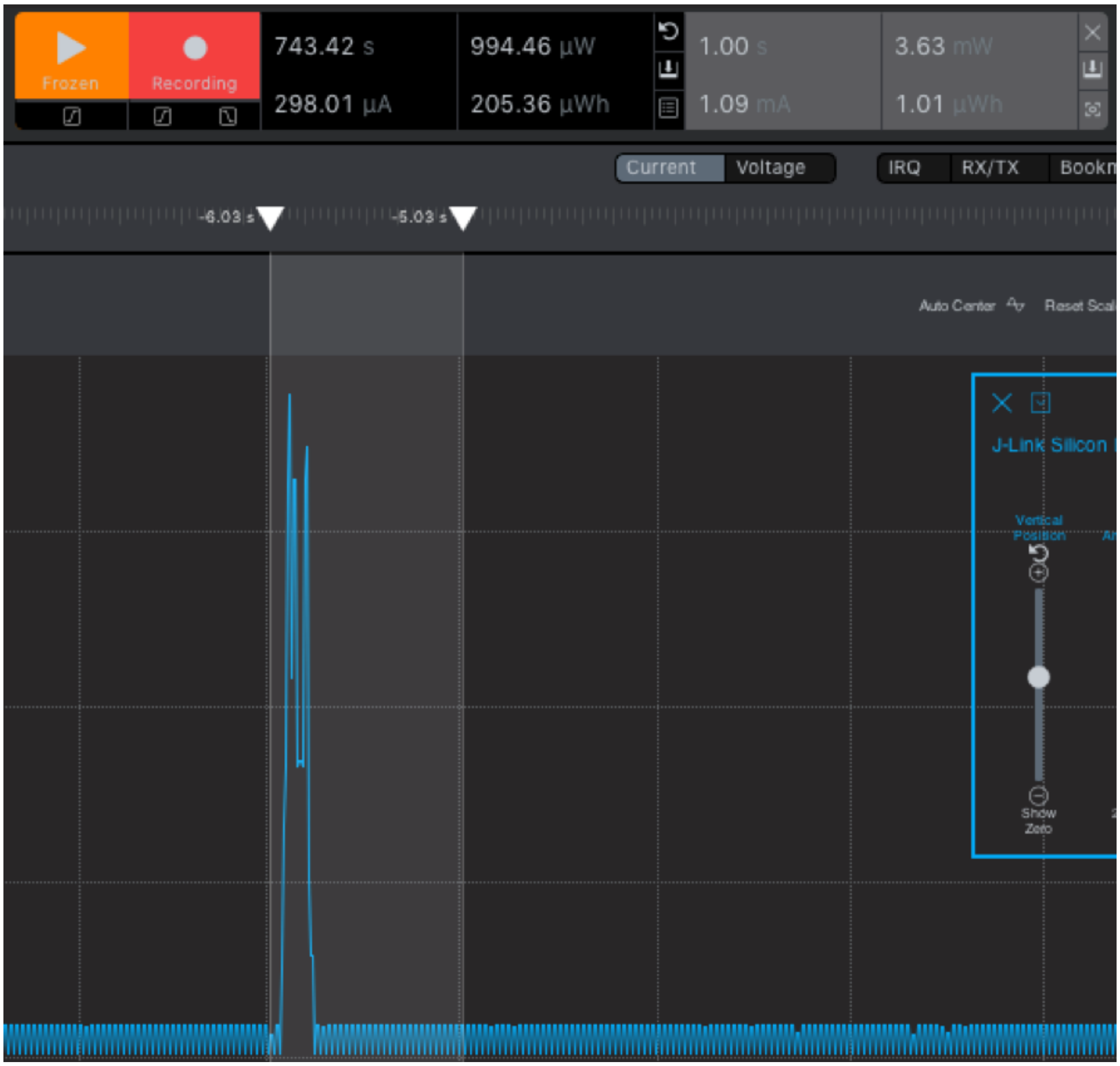


Figure 4.4: Raw ECG Data

Sending 360 double or 2.8125KB over to mist/fog layer average over 1 second takes around ***1.09mA***, and ***3.63mW*** as shown in Figure 4.4

4.4 Communication Power Isolation

The above approach and profiling show the superiority of enabling edge nodes to be smart on choosing which data to send over the network but lacks in breaking down each case's processing and communication contribution.

With the tools we had, we were unable to isolate the communication-only power consumption ' Transceiver and Modem' from overall SoC, including applications running on communication MCU alongside FW protocol stack, leakage, and peripherals. Also, the estimate was done over 1 second. So it is hard to compare with other communication protocols.

The best approach from our perspective is to estimate μJ/bit for BLE communication and compare it to other communication protocols.[16] Followed this method for LTE, 3G, and Wi-Fi, but not for BLE.

4.5 BLE Energy/bit

We needed to extend work from [16] to include BLE communication technology since it covers only LTE, 3G, and Wi-Fi (For cloud and fog approach), and our case study needs to consider BLE for low power and short range of communication for Fog approach.

Our method is to send a bulk of data over BLE and profile the energy consumed to send over this data. The platform we had did not allow separate transceiver and modem power measurement from the whole system, so MCU overhead was included in the profiling number.

This overhead should not be huge since data transmission is usually offloaded to base-band and then modem while MCU is in WFI (Wait For Interrupt) mode where clocks are suspended and MCU in sleep mode waiting for an interrupt to wake up. So this measurement error should not affect our comparison.

We utilized the BLE 5.0 2M Phy feature in our data transfer, which provides double the data rate compared to previous BLE standards (4.0 and 4.2).

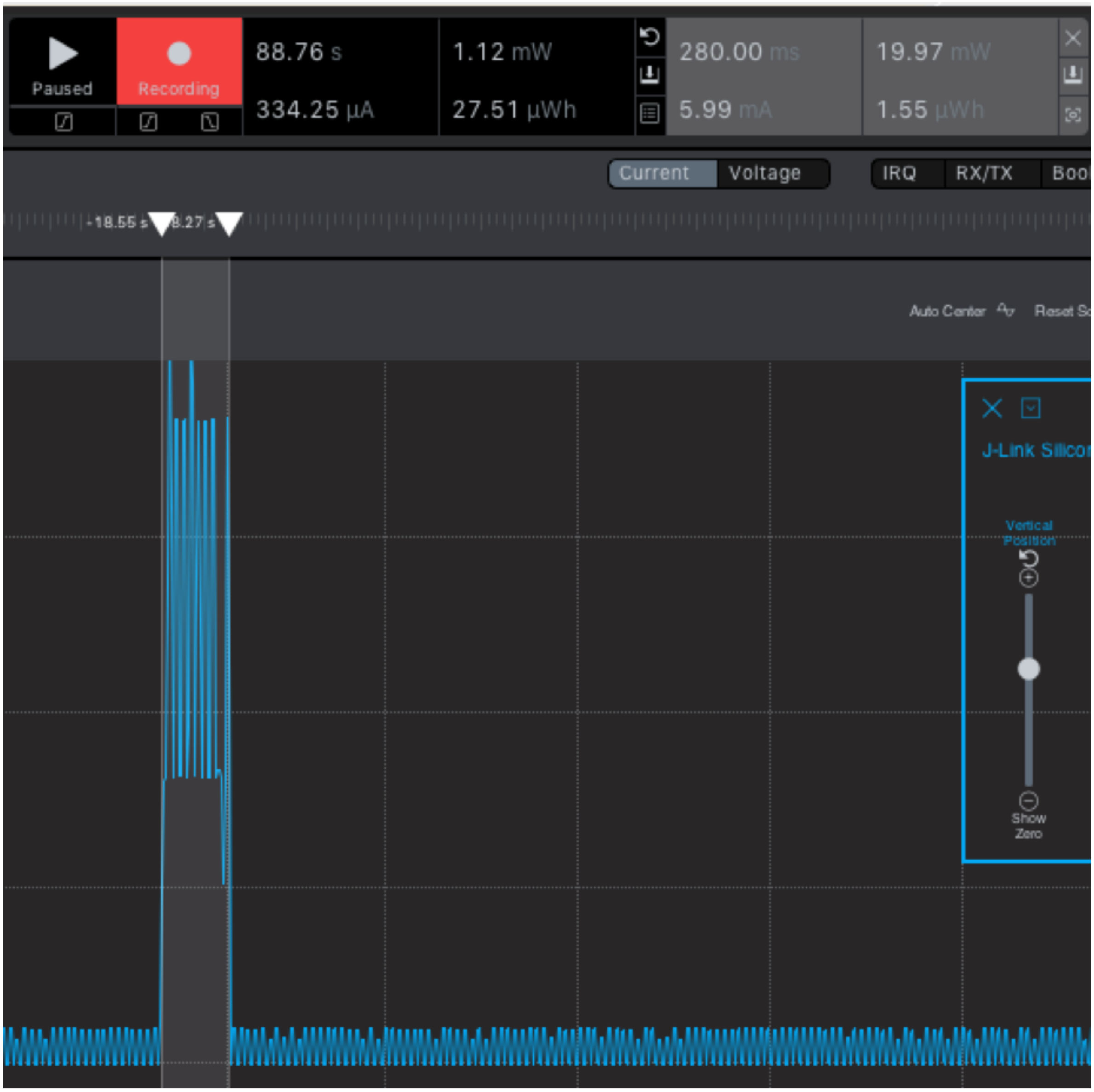


Figure 4.5: BLE energy/bit

From Figure 4.5, the total energy for this transfer is 1.55 μWh for 6480 bytes, then

$$1.55 * 60 * 60 = 5580\mu J$$

Energy per bit will be

$$5580(6480 * 8) = 0.107\mu J/bit$$

After estimating the energy consumed to transmit every bit using the BLE communication protocol, we can compare it with other communication protocols to choose the architecture based on communication range and bandwidth.

4.6 Communication Protocols Comparison

After extending work from [16] we can now illustrate the comparison of the following protocols, Wi-Fi, LTE, 3G, and BLE.

The comparison is made based on sending the estimated amount of data from our algorithm running on the Edge node to the Fog Layer below.

	BLE	$Wi-Fi$	LTE	$3G$
1 bit μJ	0.107	5.0	10.0	13.2

Table 4.1: Energy per bit (1 bit)

BLE is the most efficient communication protocol amongst the four protocols we mentioned in this work. Please refer to Figure 4.6 for energy consumed per bit for each of the protocols.

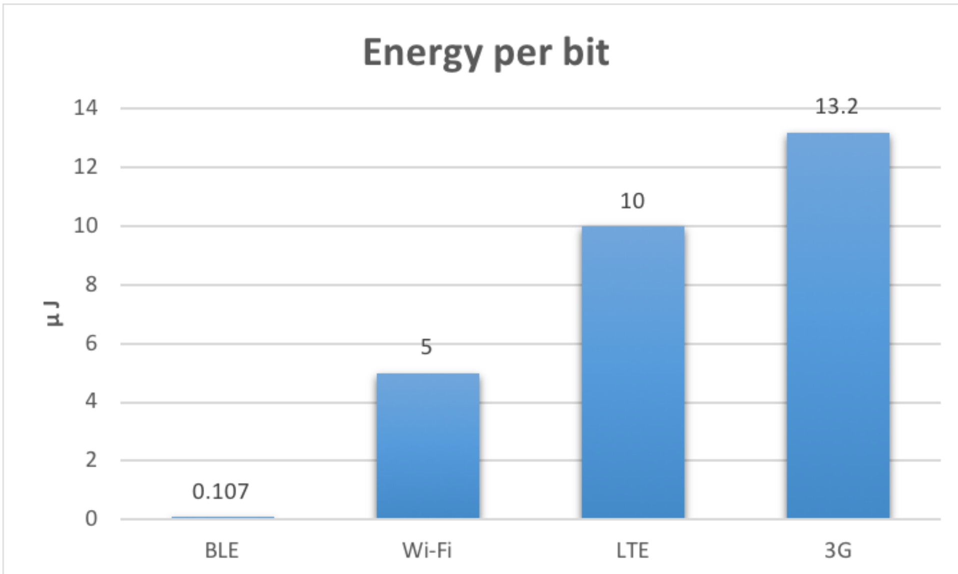


Figure 4.6: Energy per bit for communication protocols

Total energy consumption for each of our four communication protocols is shown in Table 4.2

	BLE	$Wi-Fi$	LTE	$3G$
Two Minutes Arrhythmia μJ	685	32000	64000	84480
Normal Beats μJ	3.42	160	320	422.4
Raw ECG Signal μJ	308	14400	28800	38016

Table 4.2: Total Communication Energy Comparison

For a better illustration of the significant difference that short range communication protocol (BLE) provided for Edge node, refer to Figure 4.7.

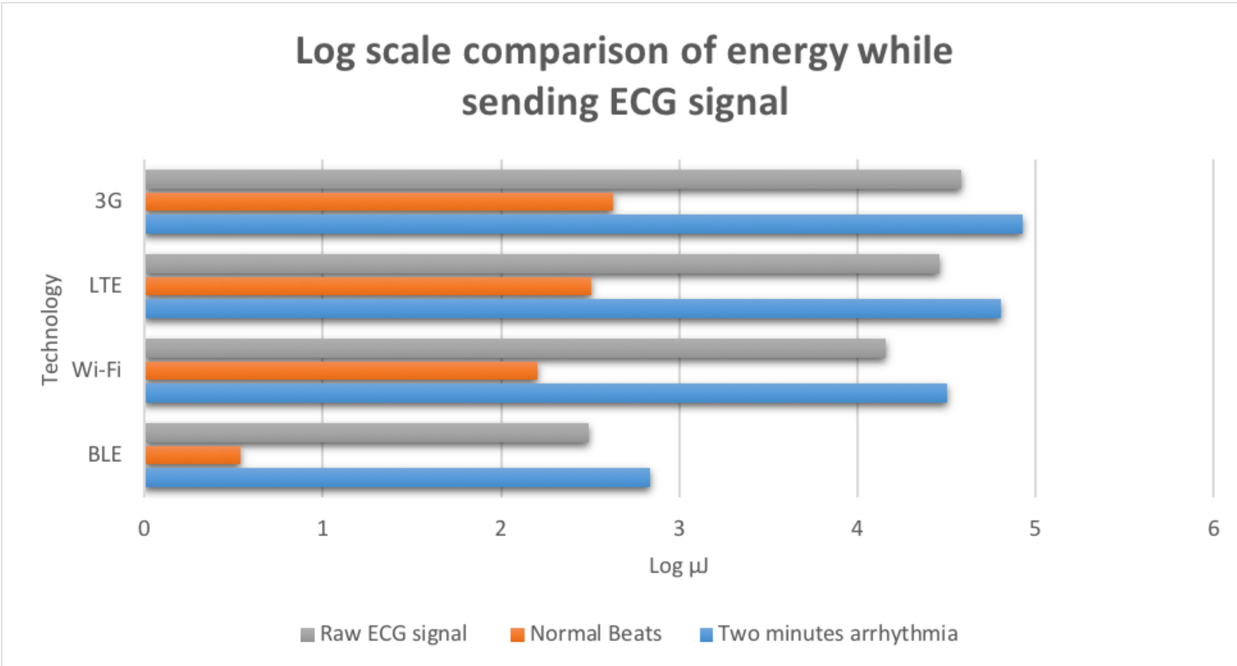


Figure 4.7: Log scale of transmitting ECG signal energy

From this comparison, we can show how choosing the suitable architecture for a particular application can save up to **123X** of the communication power budget; this is on the communication side only.

Also, as [16] discussed, bringing edge nodes closer to its extension layer from cloud to fog or mist can tremendously help in optimizing the used power as shortening the distance improves response time and provide better power optimization, but this assumes using the same communication technology.

Our work shows that in addition to all of the above, choosing the suitable communication technology can drastically contribute to overall power optimization, based on our case study, and this can optimize communication power budget by **123X**.

Chapter 5

Computation Power Profiling

In this chapter, we will demonstrate the computation power contribution of ECG classification algorithms that are running on our Edge node.

The development of ECG classification algorithms was done thoroughly by colleagues AICPS (Autonomous and Intelligent Cyber-Physical Systems) lab 'Thanks for the great contribution to this thesis' while our work focused on the following:

- Port ECG classification algorithms to our Edge node platform.
- Request algorithms modification after identifying memory bottlenecks to fit in our chosen edge node limited memory capacity or processing power.
- Added some optimization to the developed algorithms to run faster on the Edge platform.
- Run memory profiling, CPU utilization, and power profiling.

5.1 ECG Classification Algorithms

The idea in a nutshell of the algorithms is to identify on edge node if the detected ECG signal worth sending over the network in case of Arrhythmia or just it is a Normal Beats. So in the case of Normal Beats, the edge node will send a sample every second to the Fog layer. If Arrhythmia were detected, the algorithm would send two minutes' worth of data to the Fog Layer for further classification and decision (Fog Layer implementation is out of the scope of this work).

Without this intelligence [26] implemented in these algorithms, the Edge node will act as a relay for ECG signal to the Fog Layer. Details for algorithm implementation is explained in details in [12].

5.2 CNN (Convolutional Neural Network) Algorithms Profiling

The algorithm is divided into two parts; the first block is the first layer of the NN (*Neural Network*) and then its output. Both algorithms were ported to *EFR32BG13 Blue Gecko, it has 32-bit ARM Cortex-M4 core with 40 MHz maximum operating frequency and 512 kB of flash and 64 kB of RAM.*

Power profiling and execution time were performed by integrating the algorithms with BLE connection events to mimic the actual behavior of the Edge node (which sends data to Fog Layer after running the algorithm). We then followed cycle counts for the algorithm only independently of BLE connection event.

5.3 BLE Connection Event Profiling

At first, we need to estimate BLE connection event execution time and power consumption to isolate only the computational part for both algorithms.

For maintaining BLE connection, The Edge node needs to wake up every 30ms (configurable connection interval) to respond to the central node and announce its presence if there is no data to exchange.

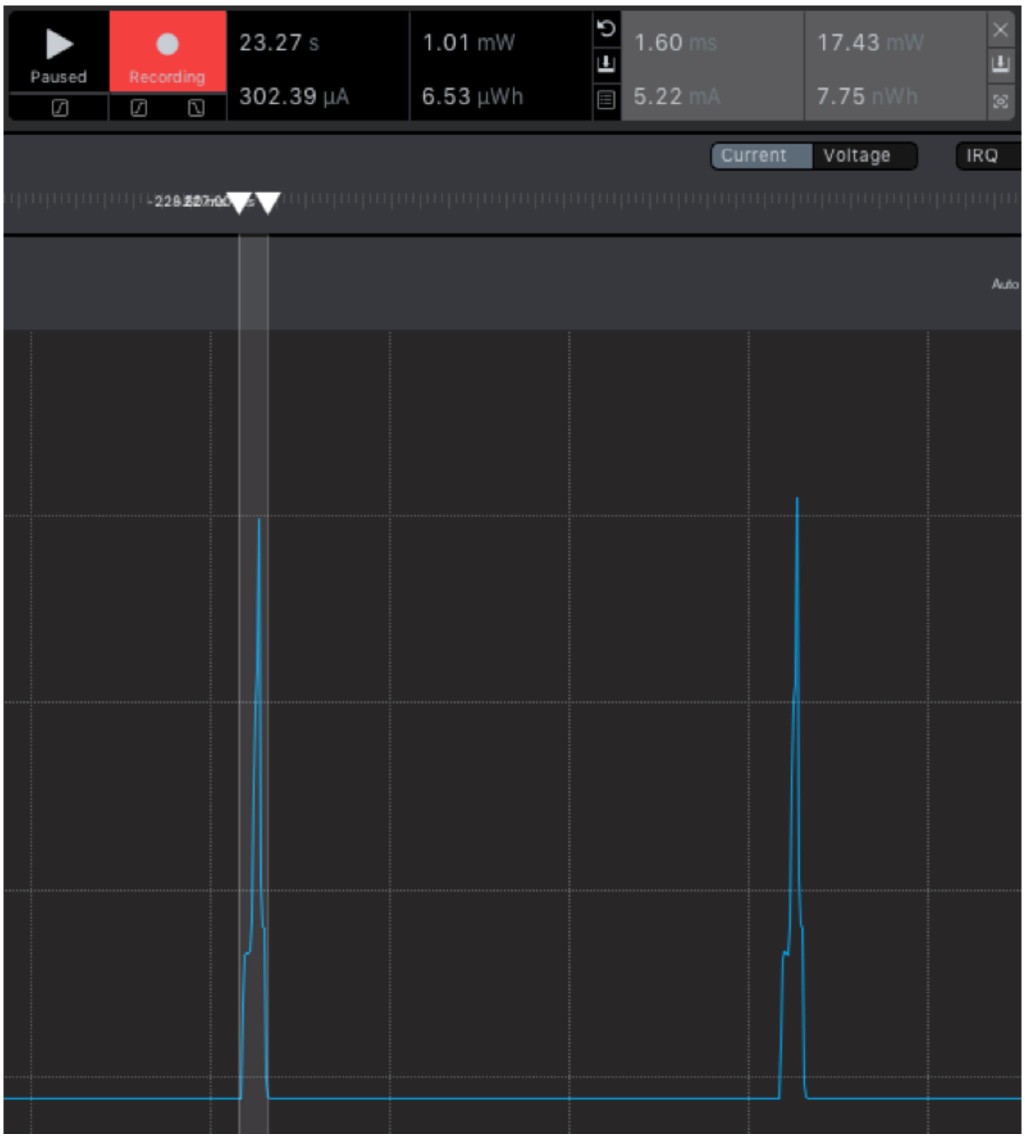


Figure 5.1: BLE connection event

From the above BLE connection event only takes ***1.60ms*** to complete and consumes ***7.75nWh***. We will use these numbers in our coming calculation.

5.4 First Layer Algorithm

This algorithm, along with BLE connection event power profiling, takes ***29.5ms*** and consumes ***119.78nWh*** is as shown in Figure 5.2 [1]

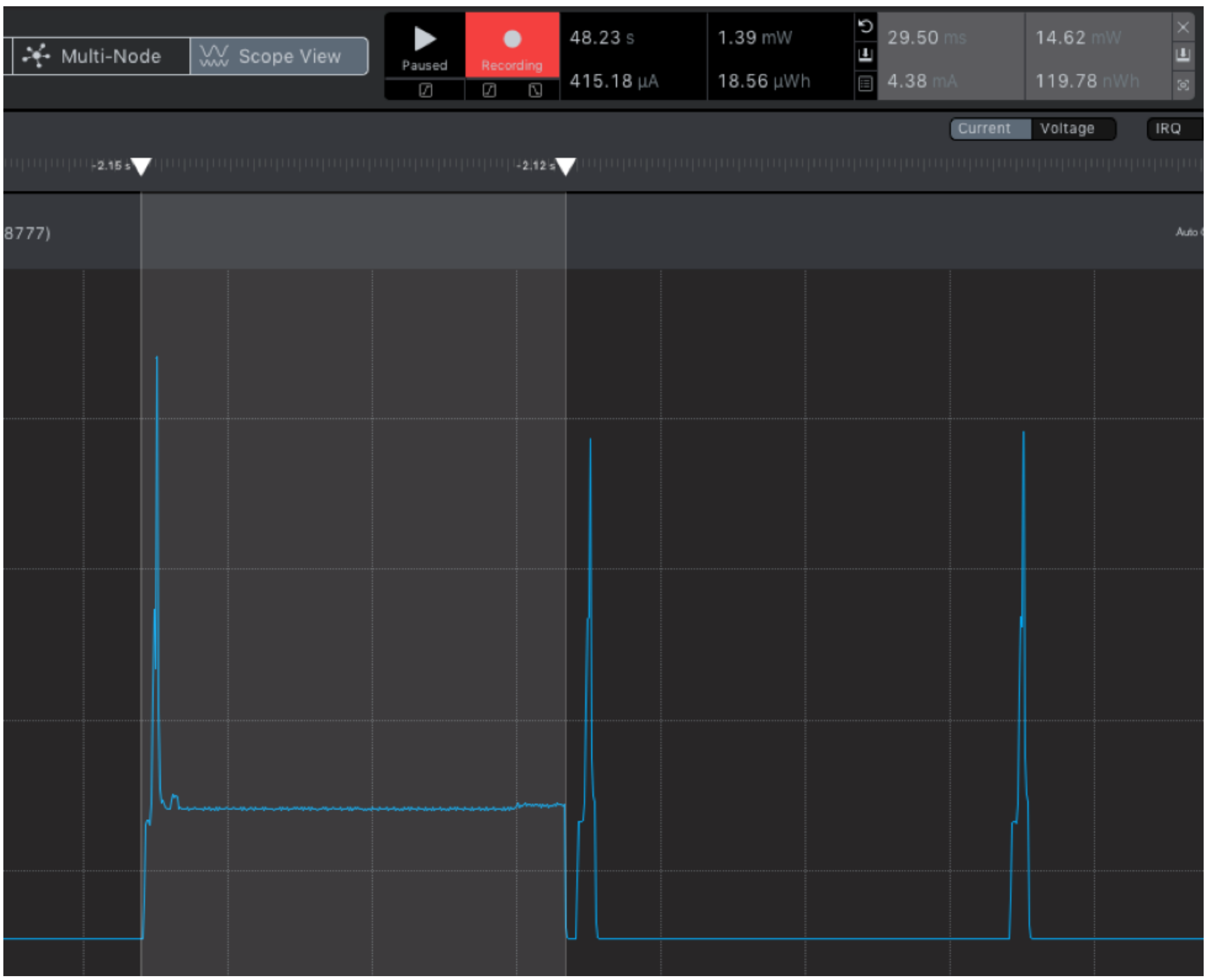


Figure 5.2: Combined BLE connection event and First Layer algorithm power

[1] we chose to combine the algorithm with BLE connection event to as we do not have any trigger to for algorithm execution profiling, so we needed to align running this algorithm with BLE connection event which is periodic and timing is expected in addition to that its closer to the real-time scenario

In order to extract execution time and power consumption for computation part only we will use BLE connection event number from previous section. This leads to execution time to be around ***27.9ms*** and consumes ***112nWh***.

We also confirmed the execution time by code instrumentation that uses DWT (Data Watch and Trace), one of ARM Cortex-M hardware provided features that we use to count execution cycles. The number of cycles we got is ***1059095 cycle***s on a processor running at 40MHz clock; this is equivalent to around ***26.4ms***. This number has 5% measurement error from the manual profiling, which is still an acceptable variation since the manual method to get the power cannot be very accurate.

Our memory usage for this algorithm is as shown in Figure 5.3 with peak RAM usage around ***10KB*** to complete the algorithm.

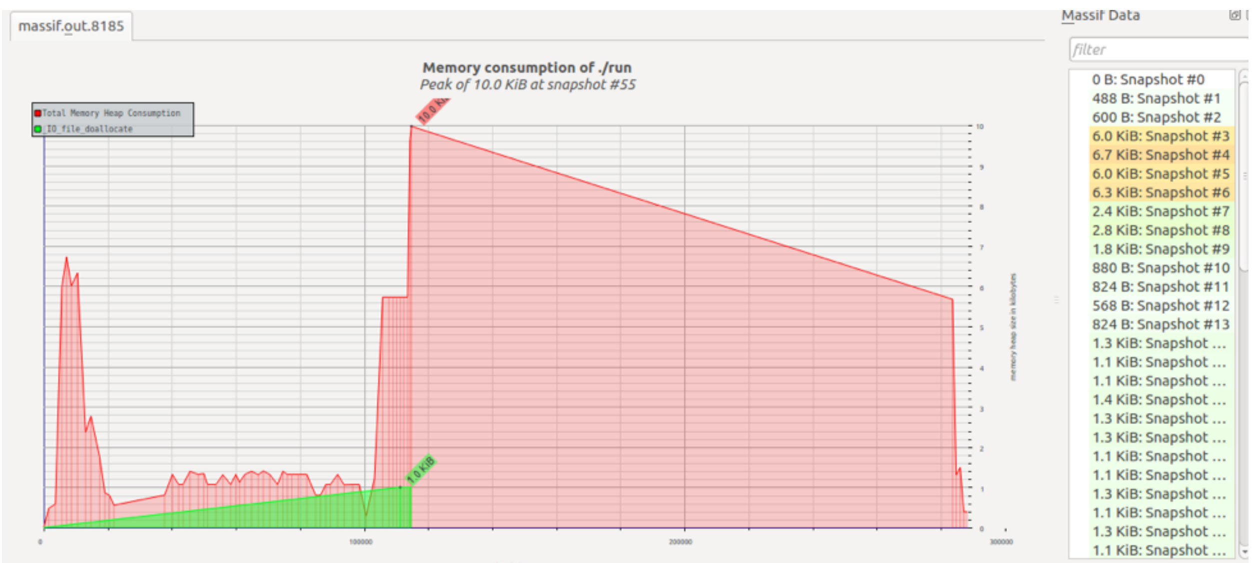


Figure 5.3: First Layer memory profiling

5.5 First Layer Output Algorithm

This algorithm along with BLE connection event power profiling takes ***3.30ms*** and consumes ***15.15nWh*** is as shown in Figure 5.4

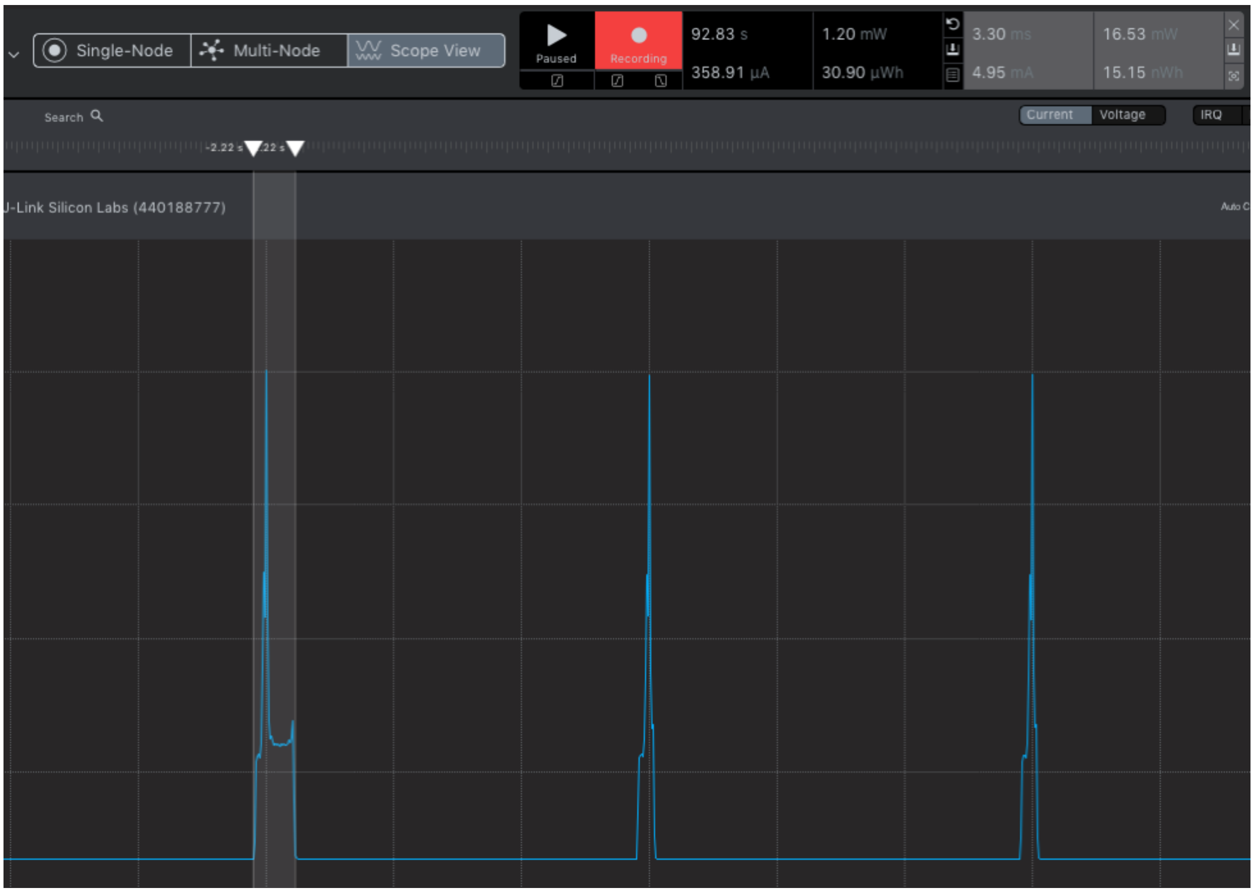


Figure 5.4: Combined BLE connection event and First Layer Output algorithm power

In order to extract execution time and power consumption for computation part only we will use BLE connection event number from previous section. This leads to execution time to be around ***1.7ms*** and consumes ***7.4nWh***.

Our memory usage for this algorithm is as below:

From Figure 5.5; peak RAM usage is around 7KB to complete the algorithm. From the above sections; the designed algorithms are suitable for limited edge nodes as the peak memory

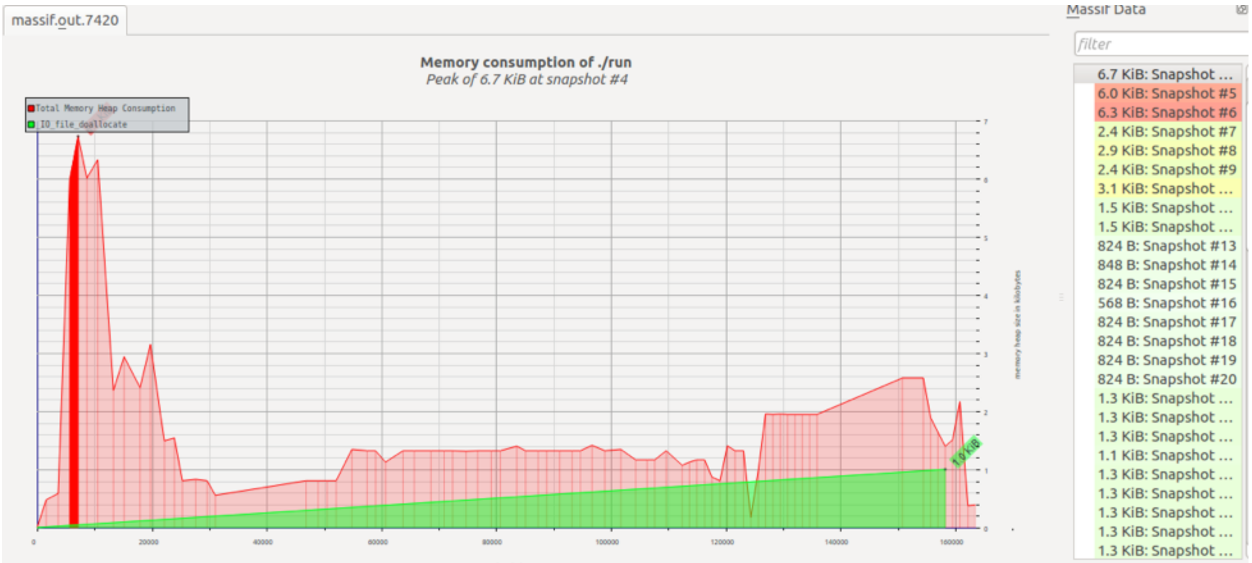


Figure 5.5: First Layer Output memory profiling

usage is around ***10KB*** and, power is kept at the minimum possible in order to make it suitable for coin cell batteries although estimating how long a coin cell battery can power our edge node was not done.

Chapter 6

Conclusion

This work discussed the rules of different layers of machine learning layers from Edge to Fog up to the Cloud layers. The current approach of offloading all the classification and learning work to the Cloud Layer while Edge nodes act as a relay for the data captured from the field comes with a cost.

This approach lacks response time, especially with real-time nature applications, since latency increases with the amount of data to send and waiting for feedback/decision from the Cloud Layer , power consumption, and security.

We focused on power consumption enhancements and latency improvements using Fog architecture, enabling Edge node to be smarter in processing collected data.

This effort was made by choosing optimized algorithms to run on edge nodes to process the data instead of just acting as a relay. Furthermore, by choosing the adequate communication protocol for short-range to take advantage of Fog architecture offering that is close to the data source.

The applied techniques resulted in a solution with overall execution time for a heartbeat takes ***36 ms*** in the edge device with ***55 mW*** power consumption. Also, our proposed methodology is compatible with any devices with a minimum RAM of ***32 KB***.

Bibliography

[1] L. A, M. A, M. JP, L. P, and R. B. Computational techniques for ECG analysis and interpretation in light of their contribution to medical advances. J. R. Soc. Interface 15: 20170821, 2018.

[2] N. Abbas, Y. Zhang, A. Taherkordi, , and T. Skeie. Mobile edge computing: A survey,. IEEE Internet Things Journal, vol. 5, no. 1, pp. 450–465, Feb. 2018., 2018.

[3] Z. Altintas, W. M. Fakanya, and I. E. Tothill. Cardiovascular disease detection using bio-sensing techniques, Talanta, Volume 128, 2014, Pages 177-186. `https://doi.org/10.1016/j.talanta.2014.04.060`, 2014.

[4] L. Arm. Arm — The Architecture for the Digital World. https://www.arm.com/products/silicon-ip-cpu, 2020.

[5] S. B. Baker, W. Xiang, and I. Atkinson. Internet of Things for smart healthcare: Technologies, challenges, and opportunities,. IEEE Access, vol. 5, pp. 26521–26544, 2017., 2017.

[6] D.-V.-S. C, V. LJ, V. R, R.-D. L, and L.-P. J-C. Smart Campus: An Experimental Performance Comparison of Collaborative and Cooperative Schemes for Wireless Sensor Network. . Energies. 2019; 12(16):3135. https://doi.org/10.3390/en12163135., 2019.

[7] R. Cardinali, L. D. Nardis, M. . D. Benedetto, and P. Lombardo. UWB ranging accuracy in high- and low-data-rate applications,. EEE Transactions on Microwave Theory and Techniques, vol. 54, no. 4, pp. 1865-1875, June 2006, doi: 10.1109/TMTT.2006.871993., 2019.

[8] M. Collotta, G. Pau, T. Talty, and O. K. Tonguz. Bluetooth 5: A Concrete Step Forward toward the IoT,. IEEE Communications Magazine, vol. 56, no. 7, pp. 125-131, July 2018, doi: 10.1109/MCOM.2018.1700053. M, 2018.

[9] T. Cooklev. Standards for Wireless Personal Area Networking (WPAN),. Wireless Communication Standards: A Study of IEEE 802.11, 802.15, 802.16 , IEEE, 2004, doi: 10.1109/9781118098837.ch3., 2019.

[10] B. D., K. E., F. V., K. A., and S. I. Comparison of technologies of local patient positioning. . International Journal of Advanced Research and Technology (2021). , 12(1), 362-386., 2021.

[11] M. Dautta, A. Jimenez, K. K. H. Dia, N. Rashid, M. A. A. Faruque, and P. Tseng. Wireless Qi-powered, Multinodal and Multisensory Body Area Network for Mobile Health. IEEE Internet of Things Journal, 2020, 2020.

[12] B. U. Demirel, I. A. Bayoumy, and M. A. A. Faruque. Energy-Efficient Real-Time Heart Monitoring on Edge-Fog-Cloud Internet-of-Medical-Things. IEEE Internet of Things Journal, 2021, 2021.

[13] A. H. R. et al. Automatic diagnosis of the 12-lead ECG using a deep neural network. Nature communications, vol. 11, no. 1, p. 1760, apr 2020., 2020.

[14] J. M. B. et al. Smartwatch algorithm for automated detection of atrial fibrillation,. J. Amer. College Cardiol. ,vol. 71, no. 21, pp. 2381–2388, 2018., 2018.

[15] C. M. Fernández, M. D. Rodríguez, and B. R. Muñoz. An Edge Computing Architecture in the Internet of Things. 2018 IEEE 21st International Symposium on Real-Time Distributed Computing (ISORC), 2018, pp. 99-102, doi: 10.1109/ISORC.2018.00021., 2018.

[16] Huang, F. Qian, A. Gerber, Z. M. Mao, S. Sen, and O. Spatscheck. A Close Examination of Performance and Power Characteristics of 4G LTE Networks,. ACM MobiSys, 2012.

[17] P. Huang, L. Guo, M. Li, and Y. Fang. Practical Privacy-Preserving ECG-Based Authentication for IoT-Based Healthcare,. IEEE Internet of Things Journal, vol. 6, no. 5, pp. 9200-9210, Oct. 2019., 2019.

[18] S. L. Inc. Which-ARM-Cortex-Core-Is-Right-for-Your-Application. https://www.silabs.com/documents/public/white-papers/Which-ARM-Cortex-Core-Is-Right-for-Your-Application.pdf, 2020.

[19] D. Jabaudon, J. Sztajzel, K. Sievert, T. Landis, and R. Sztajzel. Usefulness of ambulatory 7-day ECG monitoring for the detection of atrial fibrillation and flutter after acute stroke and transient ischemic attack,. Stroke, J.Amer. Heart Assoc., vol. 35, pp. 1647–1651, May 2004., 2004.

[20] S. Kahveci and T. Atasoy. From Wireless Personal Area Network (WPAN) to Long Range (LoRa) Technology,. 2019 11th International Conference on Electrical and Electronics Engineering (ELECO), 2019, pp. 679-681, doi: 10.23919/ELECO47770.2019.8990596, 2019.

[21] S. J. Lee, J. Kim, and M. Lee. A real-time ECG data compression and transmission algorithm for an e-health device,. IEEE Trans. Biomed. Eng., vol. 58, no. 9, pp. 2448–2455, Sep. 2011., 2011.

[22] G. Mathias, B. Manuel, B. Cornel, and M. Marcel. Comparing the energy requirements of current bluetooth smart solutions (February 2016) . https://doi.org/10.21256/zhaw-1999, 2016.

[23] F. Meneghello, M. Calore, D. Zucchetto, M. Polese, and A. Zanella. IoT: Internet of Threats? A survey of practical security vulnerabilities in real IoT devices,. IEEE Internet Things Journal, vol. 6, no. 5, pp.8182–8201, Oct. 2019, 2019.

[24] NIST. Computing, The NIST Definition of Cloud. `https://nvlpubs.nist.gov/nistpubs/legacy/sp/nistspecialpublication800-145.pdf`.

[25] M. Odema, N. Rashid, B. U. Demirel, and M. A. A. Faruque. LENS: Layer Distribution Enabled Neural Architecture Search in Edge-Cloud Hierarchies. IEEE/ACM Design Automation Conference (DAC'21), 2021, 2021.

[26] M. Odema, N. Rashid, and M. A. A. Faruque. Energy-Aware Design Methodology for Myocardial Infarction Detection on Low-Power Wearable Devices. IEEE/ACM 26th Asia and South Pacific Design Automation Conference (ASP-DAC'21), Japan, January 2021, 2021.

[27] J. Pan and J. McElhannon. Future edge cloud and edge computing for Internet of Things applications,. IEEE Internet Things J., vol. 5, no. 1, pp. 439–449, Feb. 2018,, 2018.

[28] S. Pavlopoulos, E. Kyriacou, A. Berler, S. Dembeyiotis, and D. Koutsouris. A novel emergency telemedicine system based on wireless communication technology AMBULANCE . IEEE Trans. Inform. Technol. Biomed., vol. 2, pp. 261-267, Dec. 1998., 1998.

[29] M. R., S. M.A.B., Rehman, and A. et al. Comparative performance analysis of short-range wireless protocols for wireless personal area network. . Iran J Comput Sci (2021). https://doi.org/10.1007/s42044-021-00087-1, 2021.

[30] A. M. Rahmani, P. Liljeberg, J.-S. Preden, and A. Jantsch. Fog Computing in the Internet of Things, Intelligence at the Edge, 2018.

[31] S. Ramesh and M. Govindarasu. An efficient framework for privacy preserving computations on encrypted iot data,. IEEE Internet of Things Journal, pp. 1–1, 2020., 2020.

[32] N. Rashid, M. Dautta, P. Tseng, and M. A. A. Faruque. HEAR: Fogenabled Energy Aware Online Human Eating Activity Recognition,. IEEE Internet of Things Journal, doi: 10.1109/JIOT.2020.3008842, 2018.

[33] N. Rashid and M. A. A. Faruque. Energy-efficient Real-time Myocardial Infarction Detection on Wearable Devices. IEEE Engineering in Medicine and Biology Society (EMBC'20), Montréal, Québec, Canada July 20-24, 2020., 2020.

[34] R. Roman, J. Lopez, and M. Mambo. Mobile edge computing, Fog et al.: A survey and analysis of security threats and challenges, . Future Generation Computer Systems,Volume 78, Part 2, 2018, Pages 680-698,, 2018.

[35] S. R. S., S. B. S., and S. R. K. (2018). Detection of coronary artery disease by reduced features and extreme learning machine. Clujul medical (1957), 91(2), 166–175. https://doi.org/10.15386/cjmed-882, 2018.

[36] B. SIG. Bluetooth Core Specification v 5.0. https://www.bluetooth.com/specifications/specs/core-specification-5/.

[37] S. S.P., N. A., and K. R. et al. Fog computing: from architecture to edge computing and big data processing. J Supercomput 75, 2070–2105 (2019). https://doi.org/10.1007/s11227-018-2701-2, 2019.

[38] X. Tang, Z. Ma, Q. Hu, and W. Tang. A Real-Time Arrhythmia Heartbeats Classification Algorithm Using Parallel Delta Modulations and Rotated Linear-Kernel Support Vector Machines,. IEEE Transactions on Biomedical Engineering, vol. 67, no. 4, pp. 978-986, April 2020, doi: 10.1109/TBME.2019.2926104., 2020.

[39] S. Trinks and C. Felden. Edge Computing architecture to support Real Time Analytic applications : A State-of-the-art within the application area of Smart Factory and Industry 4.0. 2018 IEEE International Conference on Big Data (Big Data), 2018, pp. 2930-2939, doi: 10.1109/BigData.2018.8622649., 2018.

[40] Y. Xiao, Y. Jia, C. Liu, X. Cheng, J. Yu, and W. Lv. Edge Computing Security: State of the Art and Challenges, . IEEE, vol. 107, no. 8, pp. 1608-1631, Aug. 2019, doi: 10.1109/JPROC.2019.2918437., 2019.