Mobile Cloud Computing: Offloading Mobile Processing to the Cloud

Sanjay P. Ahuja[1] & Jesus Zambrano[1]

[1] University of North Florida, USA

Correspondence: Sanjay P. Ahuja, School of Computing, University of North Florida, Jacksonville, Florida, USA.
E-mail: sahuja@unf.edu

Abstract

The current proliferation of mobile systems, such as smart phones and tablets, has let to their adoption as the primary computing platforms for many users. This trend suggests that designers will continue to aim towards the convergence of functionality on a single mobile device (such as phone + mp3 player + camera + Web browser + GPS + mobile apps + sensors). However, this conjunction penalizes the mobile system both with respect to computational resources such as processor speed, memory consumption, disk capacity, and in weight, size, ergonomics and the component most important to users, battery life. Therefore, energy consumption and response time are major concerns when executing complex algorithms on mobile devices because they require significant resources to solve intricate problems.

Offloading mobile processing is an excellent solution to augment mobile capabilities by migrating computation to powerful infrastructures. Current cloud computing environments for performing complex and data intensive computation remotely are likely to be an excellent solution for offloading computation and data processing from mobile devices restricted by reduced resources. This research uses cloud computing as processing platform for intensive-computation workloads while measuring energy consumption and response times on a Samsung Galaxy S5 Android mobile phone running Android 4.1OS.

Keywords: mobile cloud computing, computation offloading, energy consumption, performance, cloud computing, Amazon EC2, windows azure, android, android memory heap size

1. Introduction

Lately, recent trends aim towards the integration of huge variety of functionality and applications within a mobile device, which means mobile devices are directly affected by two main factors: battery life limited by energy consumption and processing time limited by poor mobile resources. Additionally, the limited hardware capability in current mobile devices is an obstacle to supporting the increasing high-processing demands of the latest applications and of future developments.

Energy and response times are two key design considerations across a spectrum of computing solutions, from supercomputers and data centers to handheld phones and other mobile devices (Ranganathan, 2010). Low power consumption, energy efficiency and performance have been critical concerns for developing electronic devices such as personal computers and mobile devices, and this importance seems to be increasing (Miettinen, 2010).

While the most energy-efficient setup for many current mobile applications is local computing, there clearly are workloads that can benefit from moving to remote infrastructures (Miettinen, 2010) due to these workloads demand higher computing resources. An obvious solution to the resource constraint on mobile devices is to offload computation. Therefore, offloading mobile processing could mean an effective solution to overcome the limited resources on mobile phones. However, when using this approach, the critical aspect for mobile clients is the trade-off between energy consumed by computation and the energy consumed by communication (Miettinen, 2010). In addition, there are many concerns when offloading data to remote infrastructures such as data delivery, standard interface, trust, security, privacy (Mandeep, 2015) and reliability on the wireless network as well as on remote infrastructure service availability.

Currently, there is a new set of services being widely implemented in the computer field that are able to perform computing-intensive tasks remotely with a lower cost of energy consumption on the mobile system and a shorter response time, namely, Cloud Computing Services.

1.1 Android Mobile OS

Android, one of the most popular mobile operative systems powers hundreds of millions of mobile devices around the world. The mobile OS utilizes a Linux based kernel; it implements its permission based security model and most of its code is under Apache License. Unlike on other mobile operating systems like Apple's iOS, Palm's WebOS or Symbian, Android applications are written in Java and run in virtual machines (Brahler, 2010). Therefore, Android has a broad Java library support (Java.io, Java.security, Java.net, Java.sql). In addition, Android uses its own VM; the Dalvik VM, see Figure 1. The Dalvik VM manages garbage collection, memory in a per-app cap size and allows sharing core libraries through a methodology called The Zygote. The Zygote is an approach that enables both sharing of code across VM instances and providing fast startup time of new VM instances. The Zygote design assumes that there are a significant number of core library classes and corresponding heap structures used across many applications. Generally, these core library classes are read-only and are therefore a good candidate for preloading and sharing across processes. In other words, this is data and classes that most applications use but never modify. These characteristics are exploited to optimize sharing of this memory across processes (Ehringer, 2010).
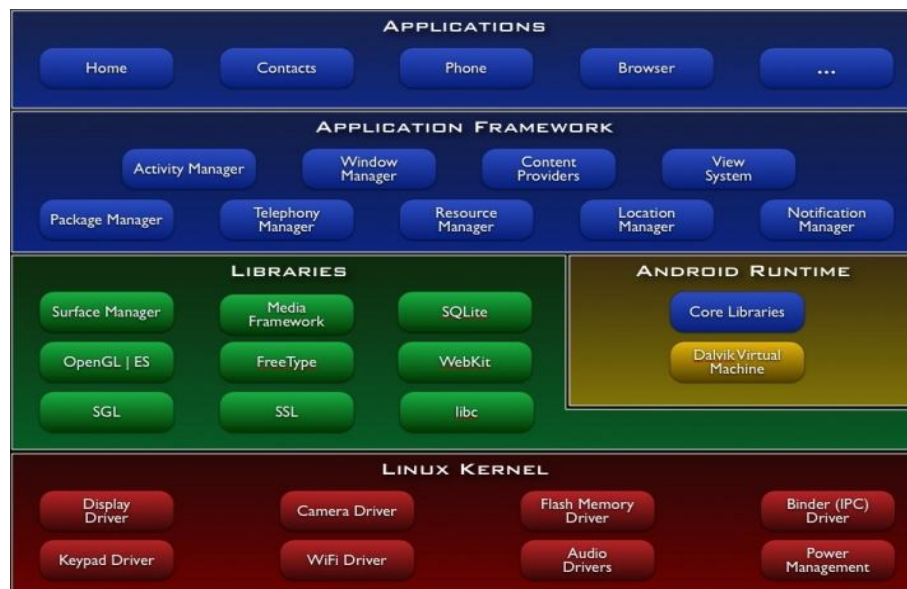

Figure 1. Android Architecture (eLinux, 2011)

1.2 Amazon EC2

Amazon EC2 is an IaaS cloud service that provides resizable computing capacity in Amazon Web Services Cloud (AWS). EC2 allows launching as many virtual machines as needed, configure security and networking, and manage storage. EC2 enables to scale up or down to handle changes in requirements or spikes in popularity, reducing needs to forecast traffic, see Figure 2. Amazon EC2 provides a web-based interface that allows users to configure their environment, virtual machines, operative systems, memory, storage and even the data center location so virtual machines can be deployed closer to the service target. Amazon EC2 also supports several instance types including micro, high CPU, high memory, cluster GPU, cluster compute, high memory cluster and high I/O instances (Bhagavathi, 2013).
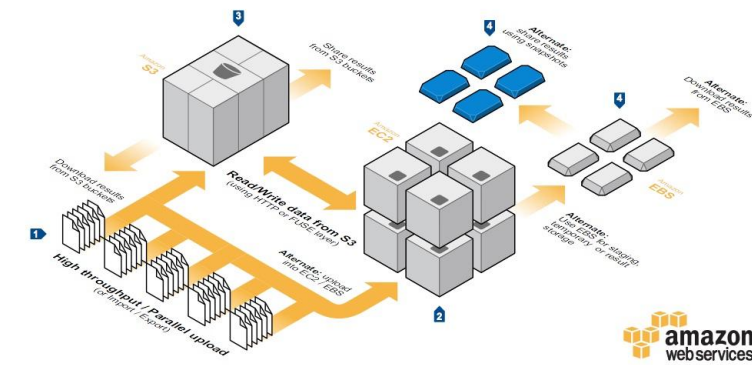
Figure 2. Amazon Web Services Architecture (Amazon AWS, 2013)

1.3 Microsoft Windows Azure

Azure is an Internet-scale cloud computing and services platform hosted in datacenters created by Microsoft Corp. An open and flexible cloud platform that provides a web-based console that allows user to create and manage cloud services, virtual machines as well as to configure networking, create security policies, administer storage, replication and so on, see Figure 3. It delivers a 99.95% monthly SLA and enables to build and run highly available applications without focusing on the infrastructure (Microsoft, 2013). A significant difference between Amazon EC2 and Windows Azure is that Azure has built in a network load balancer and resiliency to hardware failure.
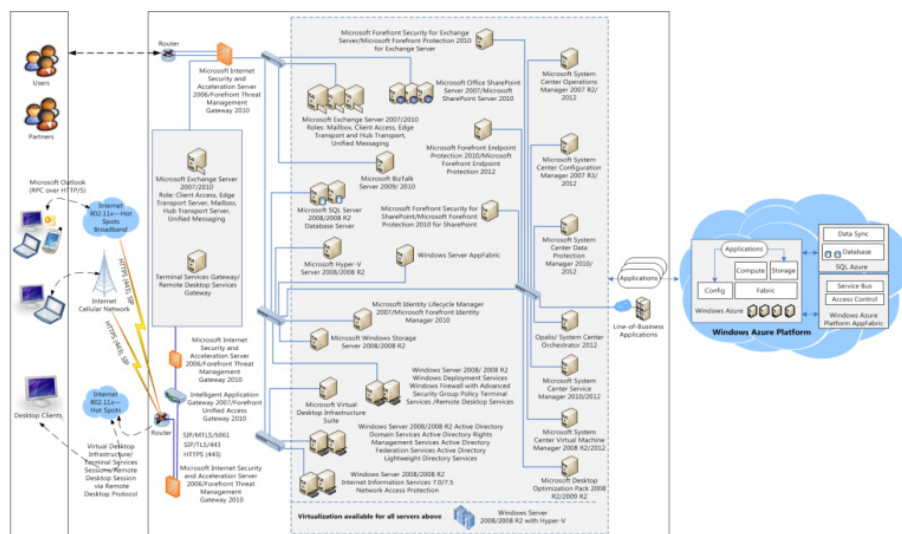

Figure 3. Microsoft Windows Azure Architecture (Azure, 2013)

In this paper, we evaluate offloading two different workloads to two public cloud-computing platforms; Amazon EC2 and Microsoft Windows Azure as well as it is evaluated the performance when using Wi-Fi and 4G as communication links. On the cloud platform sides, small, medium and large VM instance types were created on both Amazon EC2 and Microsoft Windows Azure for performing these workloads and collecting information on response times and energy consumption on a mobile phone.

The remaining sections in the paper discuss the related works in section II, our experimentation in section III followed by results discussion in section IV and conclusions in section V.

2. Motivation and Related Work

2.1 Efficiency of Mobile Clients in Cloud Computing

Miettinen *et al.* in 'Energy efficiency of mobile clients in cloud computing' discussed the potential saving of mobile clients energy when offloading certain mobile workloads to the cloud. They also discussed the trade-off

between energy consumed by computation and the energy consumed by the additional communication. They showed how the trade-offs are highly sensitive to the exact characteristics of the workload and data communication patterns.

2.2 Benefits of Offloading Mobile Computation

Kumar *et al.* in 'Cloud Computing for Mobile Users: Can Offloading Computation Save Energy?' stated that mobile cloud computing can enhance the computing capability of mobile systems and it is a possible solution for extending battery. They also discuss some factors in which mobile cloud computing may not be beneficial: privacy and security, and reliability and dependency on the wireless network and the cloud service.

2.3 Advances in Mobile Cloud Computing

Mirzoev *et al.* in 'Advances in Mobile Cloud Computing' stated that current mobile devices run high computation applications that tend to head up and drain battery power sooner. Additionally, they discussed that by allowing mobile devices to offload heavy processing to cloud systems, they can mimic traditional computer performance and enable devices to use less battery power, increase uptime and improve response times. Therefore, this computer framework dramatically improves the overall mobile computing experience.

2.4 Virtual Cloud Computing Provider for Mobile Devices

Huerta-Canepa *et al.* presented the guidelines for a framework to create virtual mobile cloud computing providers. They showed the feasibility of a new scheme of sharing mobile resources to perform common workloads among different mobile users. The power of this approach increases when there is a high availability of nearby devices with common goals/activities**.**

3. Methodology

3.1 Workloads

We used two different workloads on three different experiments in order to determine their suitability to be offloaded to the cloud. Since the decision whether or not offload processing to the cloud depends mainly on two different workload characteristics; see Figure 4. We designed one light-communication / intensive-computation and one intensive-communication / intensive-computation workload.
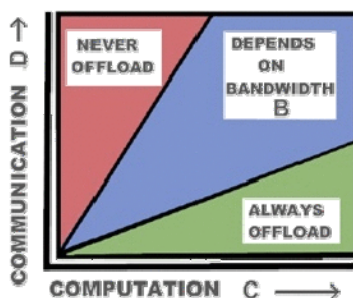

Figure 4. Decision graph with variables Computation (C) and Bandwidth (B). (Kumar, 2010)

We executed three experiments, collected data and analyzed it in order to determine the suitability of offloading intensive-computation workloads to cloud computing providers.

3.2 Light-Communication/Intensive-Computation Workload

In this study, the Android app calculates the next prime number of a given Java long number (64KB) on both, phone and cloud platforms. The algorithm used locally is the same used on the Java server application running on the both cloud providers.

This workload was designed due the data to be transmitted to the cloud has not significant size, but computing a next prime number of a large number it is a computing-intensive process. Response times and energy consumption measures were collected in order to be compared and analyzed.

3.3 Intensive-Communication/Intensive-Computation Workload

In this study, a text file was generated by a developed Java application. The file structure consists in 16-characters alphanumerical indexes randomly generated next to their pairs a regular Java integer number

ordered ascending. The Android app performs a sequential search algorithm which finds the matching pair for a given file index in this big text file on both phone and clouds platforms. The 16-characters alphanumeric file indexes were intentionally created to make the processing of comparing the given index to each file index a more computing-intensive process.

This workload was designed due the data to be transmitted to the cloud or performed locally is more communication-intensive as well as computing-intensive. Response times and energy consumption measures were collected in order to be compared and analyzed.

3.4 Methodology

This research evaluated the performance on a mobile device; a Samsung S3 phone, and two different cloud computing providers; Amazon EC2 and Microsoft Windows Azure as platforms for data-intensive computation. The study performs and analyzes three experiments in order to determine the suitability of offloading certain workloads to the cloud. This research is done by obtaining response times and energy consumption on the mobile device when performing two different kinds of workloads; light-communication / intensive-computation and intensive-communication / intensive-computation. These studies reflect what workload characteristics are more suitable to offload to the cloud.

3.4.1 Local vs. Remote Processing Comparison

In this study, the mobile app performed both workloads locally on the Android mobile phone as well as it offloaded them to three different VMs sizes on both Amazon EC2 and Windows Azure. Response times and energy consumption measurements were both collected while performing local and remote processing.

Based on these measurements, T-Test statistical analysis was performed in order to determine what kinds of workloads are more suitable to be offloaded to the cloud.

The mobile phone specifications and Cloud Computing providers VMs specifications are shown in Table 1, 2 and 3.

Table 1. Android Mobile Phone

Mobile OS	Memory	Battery	Processor
Android 4.1.2	1 GB	10.74	Quad Core 1.4Ghz

Table 2. Amazon EC2

Instance	Memory	Processor*
Small	1.6 GB	1 EC2 Compute Unit
Medium	3.7 GB	5 EC2 Compute Unit
Large	15 GB	2 EC2 Compute Unit

Table 3. Windows Azure

Instance	Memory	Processor*
Small	1.75 GB	1 EC2 Compute Unit
Medium	3.5 GB	5 EC2 Compute Unit
Large	7 GB	2 EC2 Compute Unit

Note. *1 EC2 Compute Unit provides the equivalent CPU capacity of a 1.0-1.2 GHz 2007 Opteron or 2007 Xeon processor

3.4.2 Mobile Memory Saturation Breakpoint

The goal of this study was to find a breakpoint where the mobile app could not handle workloads due their size exceeds the app memory heap size assigned by the Dalvik VM. This was performed by varying intensive-communication / intensive-computation workloads to a large VM on Amazon EC2 using both Wi-Fi and 4G. The workloads consisted in different data file size (1MB, 10MB, 25MB, 42.5MB and 45MB)

3.4.3 Amazon EC2 vs. Microsoft Windows Azure

The goal of this study was to compare response times and energy consumption measurements collected by offloading mobile workloads to Amazon EC2 and Microsoft Windows Azure. Based on these measurements,

T-Test statistical analysis was performed in order to determine what cloud platform performed better.

3.5 Metrics

Response times were measured in milliseconds for each experiment. Locally speaking, it is time it took the mobile app while performing the assigned task. Remotely, it is the time it took from offloading the assigned workload, processing it and obtaining the result back from the cloud provider. The final response time measurements are an average of performing the experiment five times under similar conditions in order to obtain a more accurate value.

Energy consumption values were measured in mAh for each experiment. Locally speaking it is the energy consumed by the app while performing the assigned task. Remotely, it is the energy consumed by the app when offloading the assigned workload, processing it and obtaining the result back from the cloud provider. The final energy consumption measurements are an average of performing the experiment five times under similar conditions in order to obtain a more accurate value.

4. Results Analysis and Discussion

The previous studies captured fair amount of data for analysis, comparison and to determine what kinds of workloads are more suitable to be offloaded to more powerful computing platforms. A statistical analysis was performed on the obtained results using the T-TEST.

In the first study, T-TEST consumed two datasets as inputs; one for Local execution and one for Amazon EC2 or Windows Azure for both response times and energy consumptions on each VM size.

In the second study, T-TEST consumed two datasets as inputs; one for Amazon EC2 and one for Windows Azure for both response times and energy consumptions on each VM size. The result of this test is a p-value where $p < 0.05$ indicates statistically significant difference between data sets.

4.1 Local vs. Remote Processing Comparison

4.1.1 Light Communication / Intensive Communication

While most of the p values are not significant ($p > 0.05$) for overall data set when comparing local computation with Amazon EC2 and Azure, there is a significant difference in computing time and energy consumption when the workload is more computing-intensive. This is supported by collected data as well as for the plotted graphics, as seen in Figures 5 and 6. The reason is that there is a communication penalty in response times and energy consumption when offloading this kind of workload to both cloud computing platforms using Wi-Fi or 4G. Results indicated that it is up to 28 times faster and it saves up to 37 times more energy on the mobile phone when the workload is more computing-intensive and it is offloaded to either cloud platform, see Table IV.

The results indicate the suitability of the cloud for computing-intensive workloads.

Table 4. Local vs. Remote Processing Comparison-Light-Communication/Intensive-Computation Response times and Energy consumption results

Workload	Response Times	Energy Consumption
Light-Communication	Up to 28 times faster	Up to 37 times more energy saving

4.1.2 Intensive Communication / Intensive Communication

Most of the p values are significant ($p < 0.05$) for overall data set when comparing local computation with Amazon EC2 and Azure. Only when using 4G and small virtual machines the P-values are not significant due this poor configuration. Therefore, there is a significant difference in computing time and energy consumption when offloading this kind of workloads, see Figures 7 and 8. After using Wi-Fi and a large VM on EC2, the response time is up to 1.6 times faster and there is a 53% of energy saving on the mobile device, See Table V.

The results indicate the suitability of the cloud for computing-intensive workloads.

Table 5. Local vs. Remote Processing Comparison - Intensive-Communication/Intensive-Computation Response times and Energy consumption results

Workload	Response Times	Energy Consumption
Intensive-Communication	Up to 1.6 times faster	Up to 53% more energy saving

4.2 Mobile Memory Saturation Breakpoint

After offloading a varying intensive-communication / intensive-computation workload to Amazon EC2 using Wi-Fi and 4G, a breakpoint found shows that there are certain workload sizes not suitable to be offloaded due loading them into memory could exceed the assigned memory heap size causing that the app crashes. After transmitting different file sizes, the results show that the app performs well when transmitting file sizes below 42.5 MB over any communication link, see Figure 9. When trying to transmit files over 42.5 MB the app misbehave; it crashes due it reaches its memory heap size. Moreover, when the app tries to allocate more memory the Android memory manager kills the app throwing an *out of memory error.* Therefore, the memory heap size assigned by the Dalvik VM for an app is a very important factor when deciding to offload certain workload.

Additionally, it is vital to clarify that heap size is device-dependent and the Dalvik VM increases it on demand until it reaches its maximum assigned memory size. Although it is possible to manually set the maximum heap size in runtime it is strongly not recommended due it may degrade the overall system performance as well as turn other apps to become unstable.

4.3 Amazon EC2 and Microsoft Windows Azure

While most of the p values are not significant ($p > 0.05$) for overall data set when comparing Amazon EC2 to Microsoft Windows Azure, there is a slightly significant difference in computing time and energy consumption when offloading to Amazon EC2, see Figures 10, 11, 12 and 13.

Based on the previous measurements, Microsoft Windows Azure response times are slightly higher than the ones on Amazon EC2. Therefore, it also means that energy consumption tends to be higher when performing processing on Microsoft Windows Azure.

4.4 Wi-Fi vs. 4G

After collecting and analyzing all the previous response times and energy consumption measures offloading workloads to both cloud providers using Wi-Fi and 4G, the outcome was that Wi-Fi produces significant better response times and saves more energy compared to using 4G for the two different   intensive-communication workloads on both cloud computing platforms, see Table VI . The reason is that transmission time over Wi-Fi takes less than transmitting over 4G.

Table 6. Wi-Fi vs. 4G Response times and Energy consumption averages results

	Response Times	Energy Consumption
Light-Communication	Wi-Fi 1.6 times faster	Wi-Fi 19% Energy Saving
Intensive-Communication	Wi-Fi 1.8 times faster	Wi-Fi 48% Energy Savings

5. Conclusions

This paper benchmarked the benefits in response times and energy consumption when offloading mobile computing-intensive workloads to Amazon EC2 and Windows Azure using Wi-Fi and 4G network technologies. Therefore, two different workloads impacting the possible decision were used to determine the suitability of offloading processing.

5.1 Local vs. Remote Processing Comparison

There appeared to be a significant benefit in response time and energy consumption when processing computing-intensive workloads on the cloud due these remote platforms can be performed them faster no matter the trade-off between energy consumed by computation and the energy consumed by communication. Additionally, it is vital to remark that the workload size is a determining factor when deciding whether or not offloads, no due the transmission time, but for the constraint given for the assigned memory heap size that doesn't allow allocating more memory than the assigned by the Dalvik VM per app. Although it is possible to manually set the maximum heap size in runtime it is strongly not recommended due it may degrade the overall system performance as well as turn other apps to become unstable.

5.2 Amazon EC2 vs. Microsoft Windows Azure

While the T-Test statistical analysis didn't determine any significant difference between Amazon EC2 and Windows Azure for overall data set there appeared to be slightly better response time and energy consumption in Amazon EC2 compared to Windows Azure.

5.3 Wi-Fi vs. 4G

There appeared to be that Wi-Fi produces significant better response times and saves more energy compared to using 4G for the two different intensive-communication workloads on both cloud computing platforms. The reason is that transmission time over Wi-Fi takes less than transmitting over 4G.

For future work this research can be taken as base to execute similar tests on different cloud providers, more powerful VMs, mobile devices, mobile OS, and more computing-intensive workloads. This research can also be used as reference for developing a middleware interface that automatically and transparently makes the decision whether or not offload mobile processing using the conclusions of this research as start point.

Acknowledgement

References

Amazon, A. W. S. (2013). *Large Scale Computing & Huge Data Sets*. Retrieved October 3, 2013, from http://media.amazonwebservices.com/architecturecenter/AWS_ac_ra_largescale_05.pdf

Bhagavathi, K. (2013). *Performance Evaluation of Data-Intensive Computing in the Cloud*. Cloud Commons Online.

Brahler, S. (2013). *Analysis of the Android Architecture*. Retrieved September 4, 2013, from http://os.ibds.kit.edu/downloads/sa_2010_braehler-stefan_android-architecture.pdf

Ehringer, D. (2010). *The Dalvik Virtual Machine Architecture*. Retrieved September 8, 2013, from http://show.docjava.com/posterous/file/2012/12/10222640-The_Dalvik_Virtual_Machine.pdf

Elinux. (2011). *Android Architecture*. Retrieved September 21, 2013, from http://elinux.org/Android_Architecture

Huerta, C., & Dongman, L. (2010). *A Virtual Cloud Computing Provider for Mobile Devices*. ACM Workshop on Mobile Cloud Computing & Services: Social Networks and Beyond. MCS'10, 2010, San Francisco, California, 2010.

Kumar, K. (2010). *Cloud Computing for Mobile Users: Can Offloading Computation Save Energy?* Retrieved June 16, 2011, from http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=05432121

Mandeep, K. (2015). *A Review on Mobile Cloud Computing: Issues, Challenges and Solutions*. Retrieved January 17, 2016, from http://www.ijarcce.com/upload/2015/june-15/IJARCCE%208.pdf

Microsoft. (2013). *Microsoft Azure*. Retrieved September 12, 2013, from http://www.windowsazure.com/en-us/overview/what-is-windows-azure

Miettinen, A. P. (2010). *Energy efficiency of mobile clients in Cloud Computing*. Nokia Research Center. Retrieved June 10, 2011, from http://research.nokia.com/files/hotcloud10_miettinen_nurminen.pdf

Mirzoev, T. (2015). *Advances in Mobile Cloud Computing*. Retrieved January 17, 2016, from http://v1.wcsit.org/media/pub/2015/vol.5.no.7/Advances%20in%20Mobile%20Cloud%20Computing.pdf

Ranganathan, P. (2010). *Recipe for Efficiency: Principles of Power-Aware Computing*. Retrieved June 24, 2011, from http://delivery.acm.org/10.1145/1730000/1721673/p60-ranganathan.pdf?ip=66.177.3.238&CFID=330
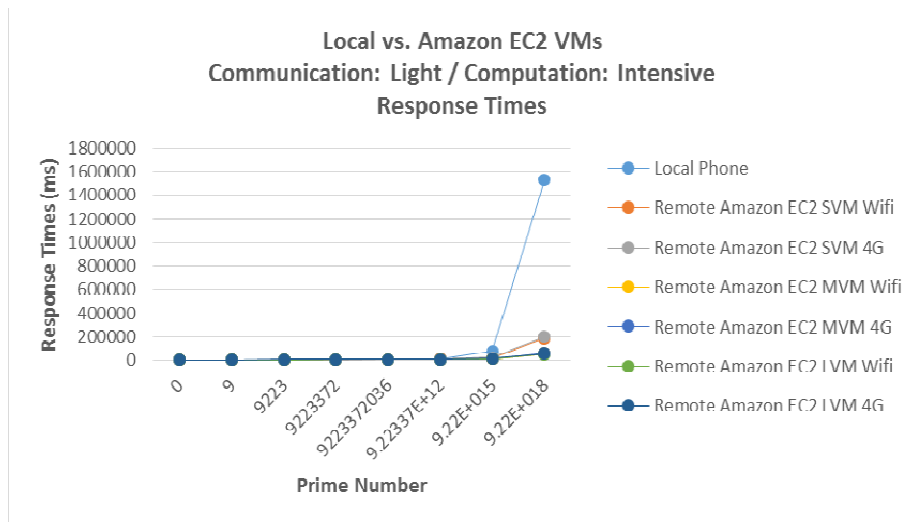
Appendix A

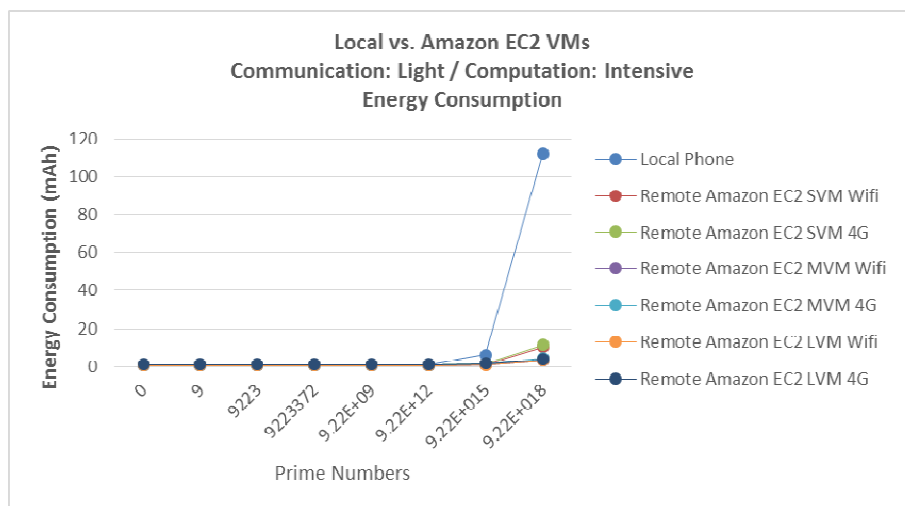Figure A1. Local vs. Amazon EC2 – Light-Communication Response times

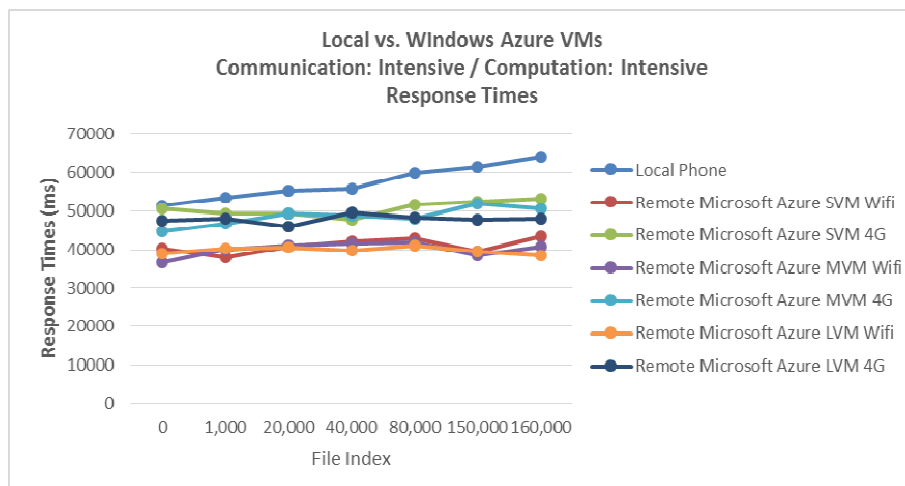Figure 2A. Local vs. Amazon EC2 – Light-Communication Energy Consumption

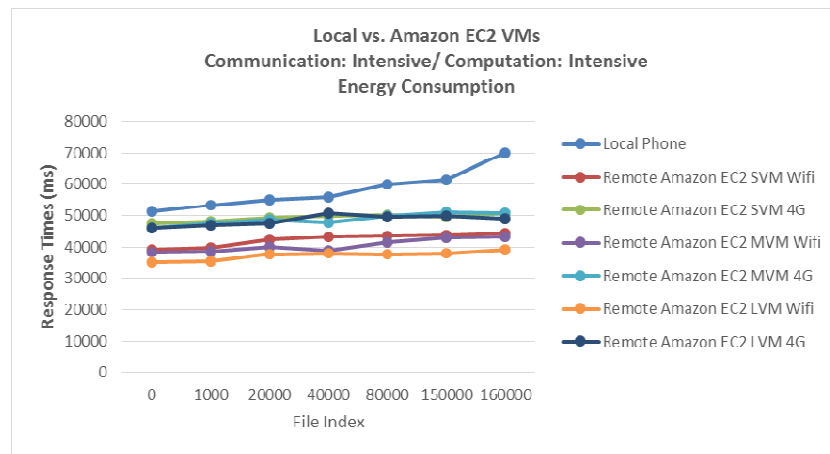Figure 3A. Local vs. Windows Azure – Intensive-Communication Response times

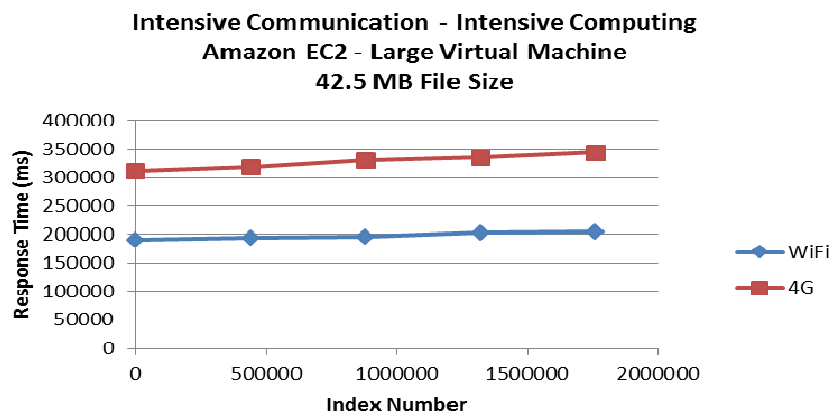Figure 4A. Local vs. Windows Azure– Intensive-Communication Energy Consumption

Figure 5A. Mobile Memory Saturation Breakpoint – Intensive-Communication on Amazon EC2 Large VM. 42.5 MB file execution
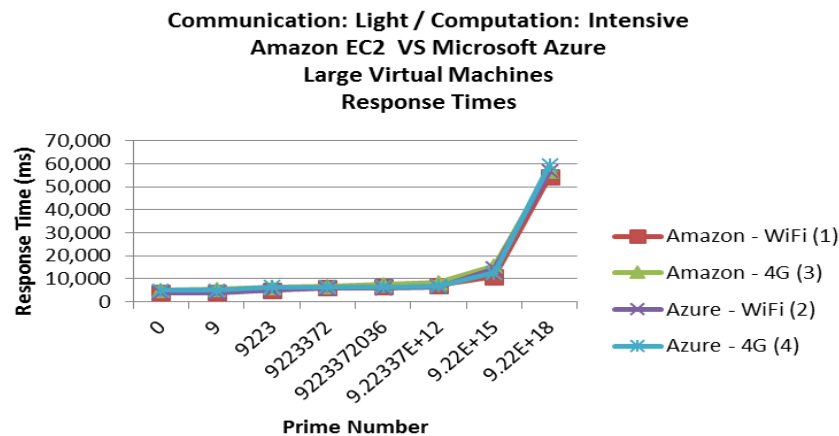

Figure 6A. Amazon EC2 vs. Windows Azure – Light-Communication Response time
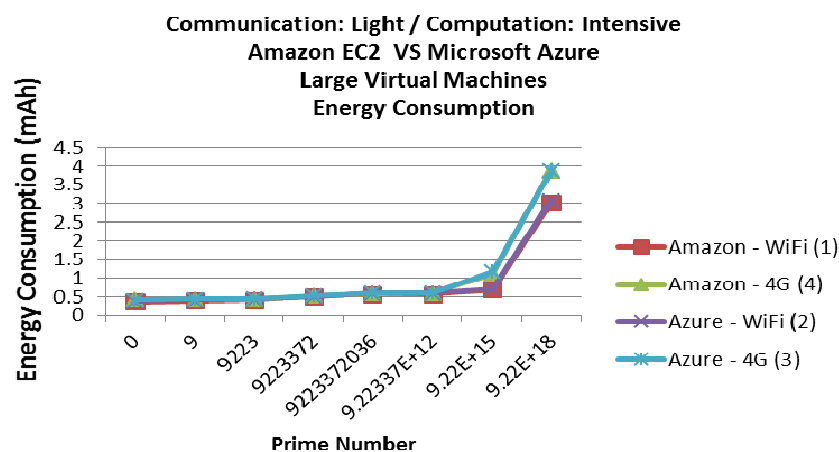
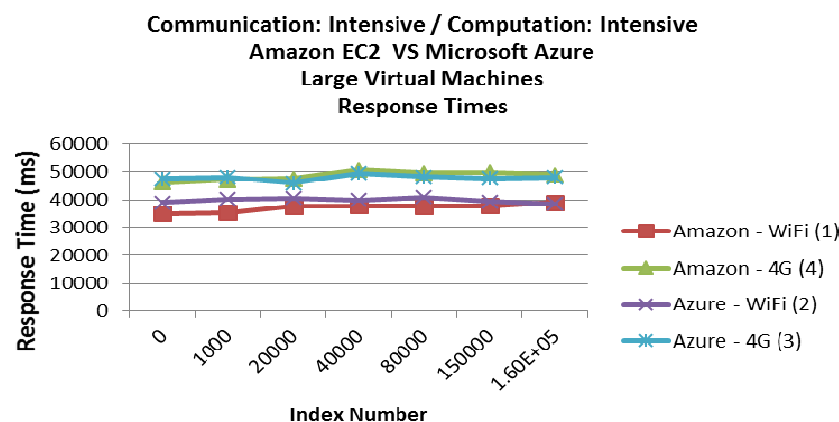Figure 7A. Amazon EC2 vs. Windows Azure – Light-Communication Energy Consumption

Figure 8A. Amazon EC2 vs. Windows Azure – Intensive-Computation Response time
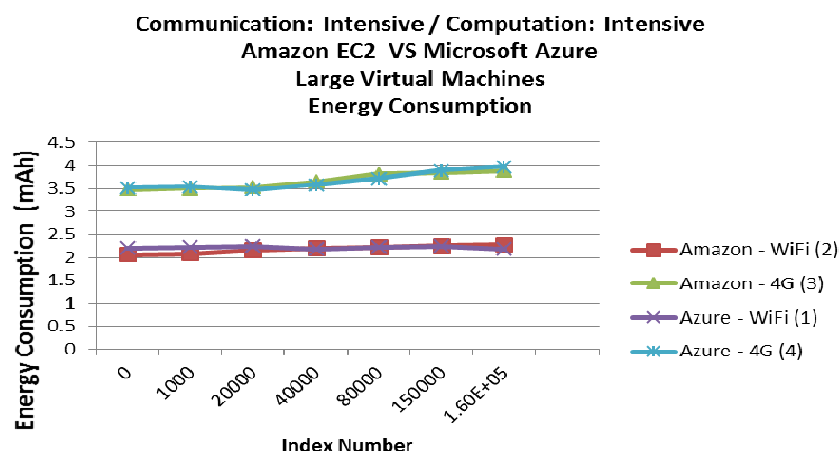

Figure 9A. Windows Azure – Intensive-Computation Energy consumption

Copyrights