Cyber-Foraging for Improving Survivability of Mobile Systems

Sebastián Echeverría†, Grace A. Lewis†, James Root†, Ben Bradshaw†
†Carnegie Mellon Software Engineering Institute
Pittsburgh, PA USA
{seecheverria, glewis, jdroot, bwbradshaw}@sei.cmu.edu
∗†Universidad de los Andes
Santiago, Chile

Abstract—Cyber-foraging is a technique for dynamically augmenting the computing power of resource-limited mobile devices by opportunistically exploiting nearby fixed computing infrastructure. Cloudlet-based cyber-foraging relies on discoverable, generic, forward-deployed servers located in single-hop proximity of mobile devices. In particular, we define tactical cloudlets as the infrastructure to support computation offload and data staging at the tactical edge. However, the characteristics of tactical environments, such as dynamic context, limited computing resources, disconnected-intermittent-limited (DIL) network connectivity, and high levels of stress pose a challenge for the continued operations of mobile systems that leverage cloudlets in tactical environments. This paper presents an architecture and experimental results that demonstrate that cyber-foraging using tactical cloudlets increases the survivability of mobile systems, defined as the capability of a system to be able to continue functioning in spite of adversity.

I. INTRODUCTION

Cyber-foraging is an area of work within mobile cloud computing that leverages external resources (i.e., cloud servers, or local servers called surrogates) to augment the computation and storage capabilities of resource-limited mobile devices while extending their battery life [1]. There are two main forms of cyber-foraging. One is computation offload, which is the offload of expensive computation in order to extend battery life and increase computational capability. The second is data staging to improve data transfers between mobile devices and the cloud by temporarily staging data in transit.

Cloudlet-based cyber-foraging relies on discoverable, generic, forward-deployed servers located in single-hop proximity of mobile devices [2]. We further define tactical cloudlets as cloudlets that support computation offload and data staging at the tactical edge [3]. However, the characteristics of tactical environments, such as dynamic context, limited computing resources, disconnected-intermittent-limited (DIL) network connectivity, and high levels of stress pose a challenge for the continued operations of mobile systems that leverage cloudlets in tactical environments.

Battery savings and better response times (lower latency) are demonstrated benefits of cyber-foraging and contributors to mission success [3]. However, mission success also requires cloudlet-deployed capabilities to be available when they are required, despite the challenges of tactical environments.

This paper presents our implementation of tactical cloudlets with a specific focus on survivability, defined as the capability of a system to continue functioning in spite of adversity. Section II presents a summary of related work in cyber-foraging and software systems survivability. Section III presents the architecture of our tactical cloudlet implementation. Section IV focuses on the tactical cloudlet features that promote survivability of mobile systems in the field. Finally, Section V concludes the paper and outlines next steps.

II. RELATED WORK

A. Cyber-Foraging

We recently conducted a systematic literature review (SLR) on architectures for cyber-foraging systems that identified 57 primary studies and a total of 60 cyber-foraging systems in these studies (52 computation offload and 8 data staging systems) [4]. The SLR showed that there is indeed active work in this area but also that there is (1) a lack of understanding of system qualities attributes beyond energy, performance, network usage, and fidelity of results, and (2) a lack of focus on system-level concerns that are required when moving from experimental prototypes to operational systems such as ease of distribution and installation, survivability and security. The goal of the work presented in this paper is to fill in some of these gaps.

As far as the use of cyber-foraging and cloudlets in tactical environments, there is indeed a motivation for their use as outlined in [5]. Satyanarayanan et al propose proximate, VM-based cloudlets as a way to overcome the lack of a reliable, high-bandwidth, end-to-end network, which is difficult to guarantee in hostile environments. More specifically, there is recent work in deploying cloudlet-like servers at the tactical edge to support mobile devices. Wang et al [6] propose the use of mobile micro-clouds and in particular mobility-induced service migration so that computation can follow its mobile users. Sookoor et al [7] propose the use of smartphones as data collection platforms that leverage mobile high-performance computers deployed on humvees for data processing. This work focuses on optimal selection of the processing node based on minimizing job completion time. Even though computation migration and optimal cloudlet selection are two features that are implemented in our tactical cloudlets to promote
survivability, the uniqueness of our work is the combination of lightweight versions of these features with additional cloudlet management capabilities that in addition support very simple cloudlet deployment in the field.

B. Software Systems Survivability

Survivability is traditionally defined in military systems as the capability to avoid or withstand the interaction between a system and a given hostile environment [Richards2007]. More specifically, the Air Force Cyber Vision 2025 document [8] lists survivability as an enduring principle and defines it in terms of fitness/readiness, awareness, anticipation, speed (responsiveness), agility, and evolvability. While this is only one definition of survivability, it is an example of the many system attributes that fall under the category of survivability.

Survivability of software systems is viewed differently among research and practice communities, but in general it refers to the capability of a system to be able to continue functioning in spite of adversity. For example:

- The Self-Healing Systems (Autonomic Computing) community defines survivability as the capability of a system to continue functioning in spite of system disruptions. More specifically, self-healing is the capability to discover, diagnose, and react to system disruptions with the goal of maximizing availability, survivability, maintainability, and reliability of a system [9].
- The Security community defines survivability as the capability of a system to continue functioning in spite of faults caused by cyber-attacks and to continue to provide essential services, even if in degraded mode [10].
- The Software Architecture community views survivability as a system quality that is related to other system qualities such as dependability, availability, reliability, fault-tolerance, and trustworthiness [11]. The goal of work in this area is to define architectural elements that promote survivability of software systems.

Our work takes the view of the software architecture community, which is to define architectural elements that promote survivability. In particular, it is aligned with Thuraisingham et al who simply define survivability as being able to adapt to changing environments [12].

III. TACTICAL CLOUDLETS

Forward-deployed, discoverable, virtual-machine-based tactical cloudlets can be hosted on vehicles or other platforms to provide infrastructure to offload computation, provide forward data-staging for a mission, perform data filtering to remove unnecessary data from streams intended for dismounted users, and serve as collection points for data heading for enterprise repositories. The forward-deployed, single-hop proximity to mobile devices promotes energy efficiency as well as lower latency (faster response times) [3].

The architecture for our tactical cloudlet implementation is presented in Figure 1. The main elements of the architecture are:

- Client: Mobile device running Android 4.x that hosts three main components:
  - Cloudlet-Ready App(s): Mobile apps that are set up to offload computation or data to a cloudlet.
  - Cloudlet Client GUI: Mobile app that is used to access the app store capability described in Section IV-A.
  - Cloudlet Client Lib: Library that is used by the two components above to discover cloudlets, retrieve data from cloudlets, and offload computation or data. It interacts with the Cloudlet Host using HTTP.
- Cloudlet Host: Linux server that runs a tactical cloudlet. The main components are:
  - PyCloud Lib: Python component that implements the core cloudlet functionality.
  - Cloudlet API: Python component that is used by the Cloudlet Client Lib to start Services as Service VMs.
  - Cloudlet Manager: Python web application that is used by an administrator to manage Services (along with their VM Images) and Cloudlet-Ready Apps.
  - Service Repository: Each capability that is made available to mobile apps is considered a Service (Section IV-A). Each service has associated metadata (Service Metadata), the actual capabilities packaged as VM disk and memory images (VM Images), and one or more Cloudlet-Ready Apps that can use the capability.
- Admin (PC): Browser that is used to access the Cloudlet Manager web application.

The implementation and additional documentation for tactical cloudlets is available on GitHub at https://github.com/SEI-AMS/pycloud.

IV. TACTICAL CLOUDLET FEATURES THAT PROMOTE SURVIVABILITY

Sheard states that implementing survivability in a software system requires knowledge of known or predictable threats [13]. Banjeree et al state that survivability deals with three basic kinds of threats: attacks, failures, and accidents [14]. However, systems in tactical environments have to deal with an additional threat which is the environment itself. To account for the characteristics of tactical environments we mapped these characteristics to system requirements and then to cloudlet features described in the following subsections.

A. Pre-Provisioned Cloudlets with App Store

In previous work [3] we presented several options for cloudlet provisioning and selected to combine pre-provisioning (Cached VM) with app store capabilities (Cloudlet Push). This combination leads to lower energy consumption on the mobile device during provisioning, places less requirements on mobile devices, and simplifies provisioning in tactical environments. The following characteristics provide the rationale for meeting system requirements:

- Capabilities as Services: Based on the definition of a service in the context of service orientation, each Service
VM provides a self-contained capability and exposes a simple interface. Service Metadata is used by the cloudlet discovery protocol to inform mobile devices of available capabilities (Section IV-C).

- **Virtual Machines as Service Containers:** VMs can be started and stopped as needed based on number of active users (which is typically bounded in tactical environments because group size is known) therefore providing scalability and elasticity to efficiently support as many active users as possible.

- **Request-Response Nature of Interactions Between Clients and Cloudlets:** In the case of computation offload, tactical cloudlets are best fit for applications based on stateless, request-response, client/server interactions. This type of interaction enables easy detection of failed communication between mobile devices and cloudlets, as well as minimal effect on mobile devices if computation needs to be restarted or migrated.

### B. Standard Packaging of Service VMs

To support ease of deployment and re-deployment of cloudlets, there is a standard format for Service VMs so these can be easily loaded from the cloudlet disk drive, an enterprise Service VM repository, a thumb drive (if allowed), or a mobile device connected via USB to the cloudlet. The file format is .csvm, which is a .tar.gz file that contains a JSON file with Service metadata, and two image files (disk and state) that compose the VM Image associated to the Service.

- **Service Metadata**
  - Service ID (string): unique identifier
  - Service Port (integer): port on which the server inside the VM is listening
  - Version (string): version number
  - Description (string): text description
Cloudlet services through Multicast DNS to the Cloudlet Multicast IP address.
3) The query reaches the Discovery Service of all cloudlets in the network through Multicast DNS.
4) Each Discovery Service replies with a DNS-SD response indicating the IP and port of its cloudlet.
5) The client requests cloudlet metadata through an HTTP GET request to each cloudlet that it has discovered.
6) Each cloudlet collects the required metadata and replies to the HTTP request with the information.

The cloudlet metadata model consists of a number of objects that are transmitted in JSON format and can be used by a mobile device to select an optimal cloudlet for offloading:

- **CPU metadata (assumes all cores are equal)**
  - `cpu.usage` (float, in %): % of CPU in use on the cloudlet, as a whole
  - `cpu.maxCores` (int): number of cores in the system
  - `cpu.speed` (double, MHz): frequency that the cores are working at
  - `cpu.cache` (int, KBs): the amount of cache per core

- **Memory metadata**
  - `memory.maxMemory` (int, bytes): total amount of physical memory on the cloudlet
  - `memory.freeMemory` (int, bytes): amount of free memory on the cloudlet

The first ranking algorithm we implemented is called CPUBasedRanker. It evaluates the cloudlet metadata to calculate the percentage of CPU power available to execute the offloaded code. This is done using the formula below which represents the percentage of CPU available on the cloudlet at that particular moment.

\[
(100.0 \times \text{cpu.maxCores}) - \text{cpu.usage}
\]

We performed experiments to evaluate the behavior of the CPUBasedRanker. These were performed on 3 machines acting as cloudlets, with the following characteristics:

- **Cloudlet 1**: i7-4700MQ @ 2.40GHz, 32 GB, 0.1% base CPU load, 1.5 GB base mem usage. Ubuntu 12.04 LTS.
- **Cloudlet 2**: Core 2 Q9550 @ 2.83GHz, 4 GB, 0.2% base CPU load, 1.1 GB base mem usage. Ubuntu 14.04 LTS.
- **Cloudlet 3**: Core 2 Q9550 @ 2.83GHz, 32 GB, 0.1% base CPU load, 0.7 GB base mem usage. Ubuntu 14.04 LTS.

The purpose of the experiments was to evaluate the response time of a request to a cloudlet and test if the algorithm correctly selected the less loaded cloudlet for each scenario. We used three test apps: FACE-OPENCV (face recognition), SPEECH (speech recognition) and OBJECT (object recognition). Results are shown in Table I. App-Ready Time measures the roundtrip request-response time in seconds of the first request, including the setup of the Service VM on the cloudlet. Average Response Time measures the average roundtrip request-response time in milliseconds of each subsequent request (all the Service VMs are capable of handling multiple users). Each row corresponds to a different scenario in the experiments.
which the three cloudlets in the network are working under different loads. Table I shows that the CPUBasedRanker indeed always selects the cloudlet with the lowest load. However, it not always selects the cloudlet that would deliver the smallest response time because not all the cloudlets in our experiments are equal. Cloudlet 1 is the most powerful cloudlet and has the smallest app-ready and response times. Data from additional experiments showed that even at higher loads Cloudlet 1 would still have smaller app-ready and response times than Cloudlets 2 and 3. Therefore, the CPUBasedRanker is ideal only if all cloudlets have the same characteristics.

This led us to another approach for a ranking algorithm which is to calculate a value that measures the overall performance of the cloudlet, which is not to calculate a single overall performance value, but to calculate a value for CPU power or memory only, in order to measure the computational performance or the memory performance of the system, respectively. Because a Cloudlet-Ready App can choose the selection algorithm to use, it could select based on the feature that is more useful for its particular needs. As such, we could have the following ranking algorithms:

- **CPUPerformanceRanker**: This could be done with an over-simplification: measuring performance using the CPU frequency as a proxy. Though this is not very accurate, it could be used as a rough comparison. A simple formula that could be used to calculate the available CPU performance of a cloudlet is:

\[
\text{CPU performance} = (100.0 \times \text{cpu.maxCores}) - \text{cpu.usage} \times \text{cpu.speed}
\]

This is the same formula used to calculate the percentage of available CPU performance in the CPUBasedRanker, but using that percentage to calculate the "available speed" for the offloaded code.

- **MemoryPerformanceRanker**: Information on free memory and CPU cache could be combined to measure the available memory performance of a cloudlet. The following formula could be used, in which A and B are constants that could be calculated to give different weights to cache and RAM memory:

\[
(\text{cpu.cache} \times \text{cpu.maxCores}) \times A + \text{memory.freeMemory} \times B
\]

### D. Cloudlet Management Component

The Cloudlet Manager shown in Figure 1 is a lightweight, web-based interface that enables easy deployment and redeployment of capabilities. It provides the following functionality:

- Service VM creation, edit and deletion
- Service VM import and export
- Service VM Instance start, stop and migration
- Cloudlet-Ready App repository (i.e., app store)

### E. Cloudlet Handoff/Migration

A characteristic of cloudlets in resource-constrained environments is that (1) they can be mobile because they could reside on vehicles and (2) clients can easily become disconnected due to mobility and to intermittent network connectivity. A strategy to deal with these problems is to enable manual handoff of data and computation between cloudlets that are within range of each other. Manual handoff would enable scenarios in which a user is migrating capabilities from a fixed cloudlet to a mobile cloudlet to support field operations, as well as reintegration back to the fixed cloudlet. Our tactical cloudlet implementation uses QEMU+KVM through libvirtd, as shown in Figure 1. The migration process leverages libvirtd to control the migration of a Service VM. The steps of the migration process are:

2. The Remote Cloudlet Server (pycloud) adds the Service VM Instance Metadata to its list of Service VM Instances in the Service Repository.
3. The Cloudlet Server pauses (suspects) the Service VM Instance so that the disk image is not changed while it is being sent.
5. The Remote Cloudlet Server points (rebases) the Service VM Instance Disk Image File to the path of the Service base VM image file on that Remote Cloudlet.
6. The Cloudlet Server calls the libvirtd function to migrate the VM (XML description and state of the VM, and memory).
7. The Remote Libvirt Daemon receives and stores the VM.
8. The Cloudlet Server notifies the Remote Cloudlet Server that the migration has finished.

To test migration times we used two of the test apps from Section IV-C — SPEECH and OBJECT — plus a third app called FLUID which is a fluid simulation app. The reason for including FLUID is because it is a continuous interaction app rather than a request-response app so we could see the effects of migration on this type of application. Migration was performed between Cloudlet 1 and Cloudlet 3 as defined in Section IV-C. Both cloudlets were connected to the same switch with a 1Gbps link. The roundtrip time for a ping between the two machines is on average 0.2 ms. Service VM Instance image file sizes and average migration times for the test applications are shown in Table II. Note that the VM disk image is not a full image, but a qcow2 Service VM instance image that only contains the differences between the disk image of the Service and the current running instance. Migration time is measured from the moment that the migration button is pushed in the Cloudlet Manager until the Service VM is running again on the Remote Cloudlet Server.

The VM image file sizes are approximately the same, as well as the VM migration times. For the SPEECH and OBJECT apps that have a request-response interaction with the cloudlet the user simply experiences a longer response time if executing the application at the moment of migration. For the
FLUID application which constantly receives data from the cloudlet (continuous interaction) the app freezes momentarily during the migration. The migration times for FLUID are slightly better than for the other cases, even though FLUID has the largest VM image sizes. This may be caused by the way in which the memory state is transferred, which is done directly by the libvirt daemons without creating a saved state file. The information in the FLUID memory state may be easier to compress or to efficiently transfer than for the other apps thus decreasing the migration times slightly.

V. CONCLUSIONS AND NEXT STEPS

The paper presents an architecture and experimental results that demonstrate that cyber-foraging using tactical cloudlets increases the survivability of mobile systems, defined as the capability of a system to be able to continue functioning in spite of adversity. The characteristics of tactical environments were mapped to system requirements for survivability and then to tactical cloudlet features. The next steps in this area of our research are to (1) develop and evaluate a set of rankers for different service characteristics, and (2) add capabilities for automated migration that would enable load balancing, similar to what is done in cloud data centers for resource optimization, or to enable migration to a more powerful or nearby cloudlet to improve response time and provide continued operations. Other areas of our research are exploring trusted identities for secure communications and the use of tactical cloudlets for efficient data staging.

ACKNOWLEDGMENT

This material is based upon work funded and supported by the Department of Defense under Contract No. FA8721-05-C-0003 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center. This material has been approved for public release and unlimited distribution (DM-0002371).

REFERENCES