My Desktop
Prepare & Submit Proposals
Proposal Status
Proposal Functions
Awards & Reporting
Notifications & Requests
Project Reports
Submit Images/Videos
Award Functions
Manage Financials
Program Income Reporting
Grantee Cash Management Section Contacts
Administration
Lookup NSF ID

Preview of Award 1527612 - Annual Project Report

Cover |
Accomplishments |
Products |
Participants/Organizations |
Impacts |
Changes/Problems

Cover

Federal Agency and Organization Element to Which Report is Submitted:	4900
Federal Grant or Other Identifying Number Assigned by Agency:	1527612
Project Title:	CSR: Small: Collaborative Research: Designing Hierarchical Edge Cloud for Mobile Computing
PD/PI Name:	Wei Gao, Principal Investigator
Recipient Organization:	University of Tennessee Knoxville
Project/Grant Period:	10/01/2015 - 09/30/2018
Reporting Period:	10/01/2016 - 09/30/2017
Submitting Official (if other than PD\PI):	Wei Gao Principal Investigator
Submission Date:	10/02/2017
Signature of Submitting Official (signature shall be submitted in accordance with agency specific instructions)	Wei Gao

Accomplishments

*** What are the major goals of the project?**

Cloud computing can be leveraged to bridge the gap between the increasing complexity of mobile applications and the limited capabilities of mobile devices, by remotely executing mobile applications at the cloud. However, the efficiency of such remote execution is hindered by excessive network latency accessing data centers and significant overhead of provisioning and managing large amounts of Virtual Machines (VMs). Traditional solutions reduce the cloud access latency by deploying servers at the network edge, but ignore the impact of mobile users' workload patterns on the efficiency of cloud operation. Instead, this project aims to design the edge cloud as a tree hierarchy of geo-distributed servers, so as to efficiently exploit the cloud resources for handling the peak load from mobile users. This research will benefit end users with various mobile devices by facilitating practical integration of these devices into the cloud. The results from this research are likely to foster new research directions on edge cloud design and mobile cloud computing. The project will engage under-represented students in the research activities, and the scholarly discovery of this project will be disseminated broadly to the community.

This project aims to satisfy the performance requirements of remote program execution by designing the edge cloud in a hierarchical manner and hence ensuring efficient utilization of cloud resources. More specifically, this project consists of three closely intertwined research thrusts: (i) developing algorithms and systems to optimize the placement of mobile workloads among edge cloud servers and efficiently serve the mobile peak load; ii) mitigating the impact of user mobility on the performance of remote program execution, by developing efficient mobility-aware VM migration techniques; iii) developing an experimental testbed, as a unique research facility, to emulate and investigate the impact of mobile workload peak on edge cloud operations.

*** What was accomplished under these goals (you must provide information for at least one of the 4 categories below)?**

Major Activities:

Modern cloud computing services, such as Amazon EC2 and Microsoft Azure, are solely hosted by data centers and incapable of efficiently executing mobile applications due to the following reasons. First, mobile applications require immediate response, and hence suffer from the excessive network latency accessing the remote data centers. Second, data centers provide virtualized cloud resources as Virtual Machines (VMs), each of which serves an enterprise user with high volumes of workloads or responds to a type of web requests. As a result, data centers also handle each mobile application using a separate VM no matter how small its amount of workload is, but incur significant overhead for global VM provisioning and management due to the huge number of mobile applications using the cloud. Such overhead may even exceed the expense of mobile program execution itself and overload the data centers during the peak hours. To address these challenges, an edge cloud, which is an intermediate cloud layer being deployed at the network edge, is indispensable to efficient remote execution of mobile programs.

To realize such vision of edge cloud, during the past year we have been focusing on constructing a personal mobile cloud, which is able to efficiently execute mobile programs at remote cloud locations and eliminate the restriction of individual mobile devices' capability in computation and communication. Our basic idea of designing such a personal mobile cloud is to incorporate and interconnect all the mobile devices owned by a user via wireless links in a generic manner, so that these devices can flexibly share system resources with each other, augmenting the mobile computing capability provided to the user. For example, wearables can save their local battery by exploiting the computational power of nearby stronger devices, while providing their sensory data to these devices and facilitate their context-aware applications. We have developed a generic resource sharing framework across heterogeneous types of mobile hardware and software, while hiding all the low-layer details of device driver operations from user applications. Interconnection between mobile devices, then, is realized via remote access and invocation of these OS services.

Education Activities:

One PhD student has worked on the project. Some of the research results have been integrated with the education curricula at University of Tennessee, Knoxville.

Specific Objectives:

Recent diversification of mobile computing devices allows a mobile user to own multiple types of devices for different application scenarios, but also results in various restrictions on the performance and usability of these devices. A viable solution to such restriction is to incorporate and interconnect mobile devices towards a personal mobile cloud where these devices can complement each other via cooperative resource sharing. Based on this insight, we aim to develop a resource sharing framework to address these challenges and generically interconnect heterogeneous mobile devices. Our basic idea is to mask the hardware and software heterogeneity in mobile systems by exploiting the existing mobile OS services as the interface of resource sharing, and further develop the resource sharing framework as a middleware in the mobile OS. We planned to implement our design over various mobile platforms with diverse characteristics and resource limits, so as to ensure that our design can efficiently

support generic resource sharing among heterogeneous mobile devices without incurring significant system overhead or requiring individual system modification.

The major challenge of realizing such a personal mobile cloud is the heterogeneity of mobile computing devices, which resides in both hardware and software aspects and prevents these devices from being interconnected in a generic manner. First, the increasing variety of hardware components being mounted on today's mobile devices results in fundamental difference in the drivers, I/O stacks and data access interfaces being used by these hardware. Even for the same type of hardware, access to the hardware data from a remote system could fail if the hardware drivers are provided by different manufacturers and incompatible with each other. Such incompatibility is usually a result of customized SoC designs adopted by different hardware manufacturers. For example, the accelerometer drivers for the Qualcomm Snapdragon chipsets are definitely incompatible with the Samsung Exynos chipsets. Second, the complexity of today's mobile applications has been dramatically increased, leading to heterogeneity in both their requested types of mobile system resources and their specific ways of accessing these resources. Existing solutions, unfortunately, are limited to interconnecting mobile devices with respect to an individual mobile application or a specific type of shared hardware. Therefore, they will need a large amount of reprogramming efforts to interconnect heterogeneous mobile devices, by rewinding the wheel for each individual hardware or software component of these devices. Such reprogramming efforts do not only impair the usability of mobile computing system in versatile environments, but also incur additional overhead to the operation of mobile OS and hence reduce the mobile system performance.

Significant Results:

We envision that the key to addressing the aforementioned challenges and realizing generic interconnection across heterogeneous mobile devices is to develop an efficient framework for resource sharing between these devices, which appropriately masks the hardware and software heterogeneity in mobile systems from each other. Development of such a resource sharing framework, however, is challenging due to the close interaction between mobile hardware and software. A framework at the lower layer of mobile OS hierarchy unifies the heterogeneous resource requests of mobile applications, but has to tackle with individual hardware drivers which are operated in intrinsically different ways and hence incurs a tremendous amount of re-engineering efforts. Sharing system resources at the application layer, on the other hand, is able to access mobile hardware through a generic OS interface, but has to be associated with specific data transfer protocols and hence has limited generality.

Instead, our work develops such a generic resource sharing framework towards a practical personal mobile cloud as a middleware in the mobile OS, which exploits the existing mobile OS services to share resources between mobile devices. These services hide the low-layer details of device driver operations while providing unified data access APIs to user applications. Interconnection between mobile devices, then, could be realized via remote access and invocation of these OS services. Since these services are executed as a standalone system process by the OS kernel and are separated from application processes, remote service invocation can be done via inter-process communication (IPC) between mobile systems without involving complicated issues such as memory referencing and synchronization. As a result, any new device can be incorporated into the mobile cloud by inserting our framework into its OS, without modifying the OS kernel, our framework itself, or the source code of any mobile application.

We have implemented our design on Android OS with less than 5,000 Lines of Codes (LoC) over various mobile platforms including smartphones, tablets and smartwatches, and demonstrated the efficiency of sharing various types of hardware (GPS, accelerometer, audio speaker, camera) between remote mobile devices. The evaluation results show that our design can efficiently support ubiquitous resource access between remote systems with arbitrary mobile applications accessing these resources,

without incurring any significant system overhead. Our proposed framework is also fully compatible with existing application-level remote messaging protocols (e.g., MQTT and XMPP), and hence can also be efficiently exploited for mobile application development.

Key outcomes or Other achievements: The results of our work “Interconnecting Heterogeneous Devices in the Personal Mobile Cloud,” has been accepted by the highly competitive *IEEE Conference on Computer Communications (INFOCOM 2017)*, which has an acceptance ratio of 20%.

*** What opportunities for training and professional development has the project provided?**

One PhD student has worked on the project, and the research results have been published at top-tier conference proceedings.

*** How have the results been disseminated to communities of interest?**

Our research work in this project has resulted in one conference paper. The publications will help people better understand our novel designs on the generic resource sharing framework towards the personal mobile cloud, and further implement these designs to improve the performance and efficiency utilizing the public cloud resources in practice. We have also given seminar and summer camp talks to high school students to stimulate their interest in engineering majors.

*** What do you plan to do during the next reporting period to accomplish the goals?**

We plan to further extend our developed technique of generic resource sharing among personal mobile devices, and synergistically integrate such resource sharing into the design and implementation of hierarchical edge cloud. More specifically, we consider the scenario where a mobile user moves and hence connects to different edge cloud servers. In this case, synthesizing a new VM for serving the multiple mobile devices owned by this user at each edge cloud server would be expensive. Another alternative is to migrate an existing VM among these servers, but may incur a large amount of local computational overhead at multiple devices. Instead, we will investigate the possibility of utilizing the collective computational power among co-located mobile devices owned by the same user for such mobility support, and then improve the performance of remote program execution after VM migration. Furthermore, we also plan to implement our proposed hierarchical edge cloud designs, optimization algorithms, and systematic techniques as an experimental testbed over practical mobile platforms and cloud computing facilities, and to evaluate their effectiveness in real-world mobile computing scenarios. The main contribution of this testbed will be that it provides a unique research facility for emulating and investigating the impact of mobile workload peak on edge cloud operations.

Products

Books

Book Chapters

Inventions

Journals or Juried Conference Papers

Yong Li and Wei Gao (2017). Interconnecting Heterogeneous Devices in the Personal Mobile Cloud. *Proceedings of the 36th IEEE Conference on Computer Communications (INFOCOM)*. . Status = PUBLISHED; Acknowledgment of Federal Support = Yes ; Peer Reviewed = Yes

Licenses

Other Conference Presentations / Papers

Other Products

Other Publications

Patents

Technologies or Techniques

Thesis/Dissertations

Websites
Project website
http://www.pitt.edu/~weigao/reporting/edge_cloud.html

Participants/Organizations

What individuals have worked on the project?

Name	Most Senior Project Role	Nearest Person Month Worked
Gao, Wei	PD/PI	2
Li, Yong	Graduate Student (research assistant)	4

Full details of individuals who have worked on the project:

Wei Gao
Email: weigao@utk.edu
Most Senior Project Role: PD/PI
Nearest Person Month Worked: 2

Contribution to the Project: Manage the project team. Design and implement the generic resource sharing framework towards the personal mobile cloud.

Funding Support: This grant

International Collaboration: No
International Travel: No

Yong Li
Email: yli118@vols.utk.edu
Most Senior Project Role: Graduate Student (research assistant)
Nearest Person Month Worked: 4

Contribution to the Project: Design and implement the generic resource sharing framework towards the personal mobile cloud.

Funding Support: This grant

International Collaboration: No
International Travel: No

What other organizations have been involved as partners?
Nothing to report.

What other collaborators or contacts have been involved?
Nothing to report

Impacts

What is the impact on the development of the principal discipline(s) of the project?

Integration of mobile devices into the cloud dramatically extends the capacities of these devices and fundamentally transforms the way mobile computing applications and services are developed and operated. This integration, however, also imposes serious challenges on the cloud capacity and the efficiency of cloud resource utilization. The transformative nature of the proposed research is to rethink how the edge cloud should be designed to efficiently support remote execution of mobile programs, by turning various analytical modeling and optimization techniques into actionable system design strategies. The research can also spawn a new area of research on hierarchical designs of edge cloud. Finally, the analysis techniques, the evaluation methodology and systems developed in this research will be valuable for future undertakings.

What is the impact on other disciplines?

The edge cloud is a typical example of computer systems with heterogeneous types of workloads in the system's execution. Being able to efficiently improve the performance and reduce the expense of executing these workloads has a direct and immediate impact on a large variety of distributed computing and cyber-physical systems.

What is the impact on the development of human resources?

Many of the research results have been integrated into the undergraduate curricula at the University of Tennessee, by adopting many perspectives of the research results for undergraduate students' course projects and senior design topics. The project has supported one PhD student working on his dissertation. The involvement of the graduate and undergraduate students into this research will prepare them for leadership roles in computer science research, academia, and industry.

What is the impact on physical resources that form infrastructure?
Nothing to report.

What is the impact on institutional resources that form infrastructure?
Nothing to report.

What is the impact on information resources that form infrastructure?
Nothing to report.

What is the impact on technology transfer?
Nothing to report.

What is the impact on society beyond science and technology?
Nothing to report.

Changes/Problems

Changes in approach and reason for change
Nothing to report.

Actual or Anticipated problems or delays and actions or plans to resolve them
Nothing to report.

Changes that have a significant impact on expenditures
Nothing to report.

Significant changes in use or care of human subjects
Nothing to report.

Significant changes in use or care of vertebrate animals
Nothing to report.

Significant changes in use or care of biohazards
Nothing to report.