Performance and mobility in the mobile cloud

Jakub Krzywda
Dept. of Computing Science
Umeå University
SE-901 87 Umeå, Sweden
Email: jakub@cs.umu.se

William Tärneberg
Dept. of Electrical and Information Technology
Lund University
Ole Römers väg 3, 223 63 Lund, Sweden
Email: william.tarneberg@eit.lth.se

***Abstract*—In an mobile cloudtopology the cloud resources are geographically dispersed throughout the mobile network. Services are actively located with close proximity to the user equipment. Georgraphically migrating a service from data centreto data centrewith its user equipmentimposes a load on the affected data centres. Consequently, user equipmentmobility provides a fundamental problem to the mobile cloudparadigm.**

This paper determines the fundamental service performance issues in system of mobile users with disperead data centres, in relation to the placement of the mobile cloudhost nodes and explores the user equipmentand provider utility of subscribing to an mobile cloudnode at a certain network depth.

***Keywords*—Cloud, Mobility, Mobile infrastructure, User experience consistency, Omnipresent Cloud, Infinite cloud, Edge cloud, Latency, Throughput, Virtualization, Geo-distributed resources, VM migration**

I. Introduction

Mobile service and function are at an increasing rate beeing virtualized and augmented to the cloud. Applications are soon more ofthen than not seamliessly executed, partially or fully in the cloud. Alongside applications, fundemental user equipmentresources, such as storage and CPU, are being virtualized. In this paradigm, the border between what is beeing executed locally and remotely is blurred as developers are given more powerful tools to tap into remote ubiquitous generic virtual resources. This resource paradigm, has overwhelmingly increased the capabilities of mobile applications, simplifying hardware and enabled collaborateive computing.

Nevertheless, as we begin to rely more on remote resources we also grow more dependant on the communication delay introduced the intermediate network and by the geographical separation of the user equipmentand the data centre. Latency sensitive applications such as process controlls, storage, and compute offloading will quickly faulter if subject to a significant and varying delay.

The virtual resources are accessed through increasingly congested mobile access networks. With more devices arecrowding the mobile networks, and aplications are generating and receiving more data, this congestion translates to delay. Additionally, the geographic distance to the data centre introduces a propagation delay.

The mobile cloudparadigm, put forward by [1], [2], attemps to remedy the aformentioned congestion and delay by locating cloud resources at various stategic nodes in and adjacent to the mobiel access network. At one extreme data centreresources can propsedly be located in at the edge of the network, adjacent or integrated into an radio base station, catering for the user equipmentsreciding in its cells. Alternatively, or complimenarty, data centrescan be integrated with resources in the proposed forthcoming virtualized radio access networks.

The concept of geo-disributed resources has been worked on

The geograpghic proximity between the user equipmentand the data centreis an esscential parameter when eliminating application service delay, to that effect, services are migrated with the user equipment, through the network to minimize this incurred delay. Services, or rather the VMs that host the service is migrated to the node that is available, provides the lowest delay, and least global network congestion. However, by doing so might minimize the experiance delay for the user equipment, but will incurr a migration overhead in data centreand in the network a VM is migrated. Conceivably, various schemes and cost functions can be deployed to minimize both the delay experianced by the user and the added resource strain to th data centreand the network.

User mobility is a fundamental dynamic property of mobile cloud, and it is essential to understand how user equipmentmobility affects the perceived service performance and what load it imposes on the network.

This paper provides an investigation into the fundamental effects of user user equipmentin the mobile cloudin relation to the number of subscribers, the abstract placement of the servers, and the numer of services. An optimal or reasonable technical bounds for the mobile cloudtopology is not yet to be determined. This paper disregards the deeper techincal and topological constrains of existing mobile systems in order to provide fundamental resulst that can be employed to shape the forthcoming mobile nework generations.

II. Desiard model

As the the topology of the mobile cloudand future mobile networks is yet to be determined, and due to the fact that we want to research the effects of mobility without a socioeconomic model, the network

III. Simulation

Multiple runs for each DC/mobile cloudplacement mode.

Service The traffic generated by and the usage pattern of a simple web application is characteristic of any smaller mobile application. The HTTP traffic model in [3] provides a small scale closed loop

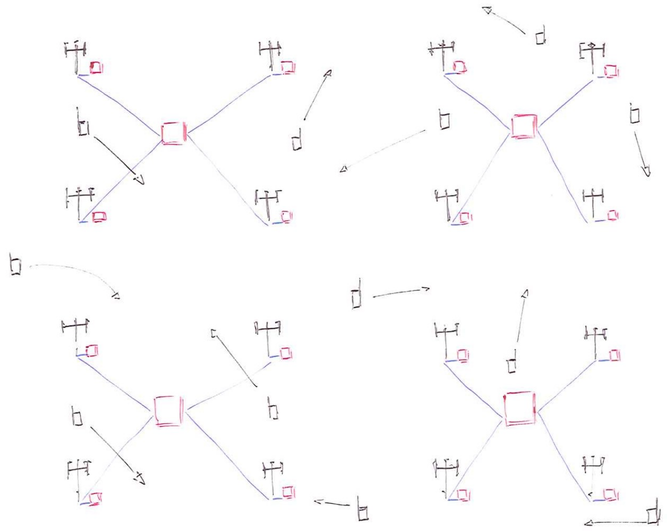

Fig. 1: Performance model

traffic model that is representative of light mobile traffic.

Mobility The 2 dimensional, multi model, mobility model [?] will provide the uniform mobile network with an relevant distribution of users.

Mobile Access Network Provided by Williams SIMJava framework. Handover are instantanious and move

Core network No delay, no routeing

Server The server provides VM and DC models that encompass, inucrred VM migration performance degredation in DC and in VM, resulting in a different service time.
Possible service hosting schemes:
- One service model, one VM is empolyed to host that service for each user.
- One service model, each VM hosts multiple but each number of users, behaving as multiple services while still being compatible.

At all placement modes:

- Measure RTT for all packetc at UE
- Measure DC load
- Measure ratio of requests generated vs. processed in mobile cloudnode
- Identify the incurred VM migration load

IV. Simulation

Multiple runs for each DC/mobile cloudplacement mode.

Service The traffic generated by and the usage pattern of a simple web application is characteristic of any smaller mobile application. The HTTP traffic model in [3] provides a small scale closed loop traffic model that is representative of light mobile traffic.

Mobility The 2 dimensional, multi model, mobility model [?] will provide the uniform mobile network with an relevant distribution of users.

Mobile Access Network Provided by Williams SIMJava framework. Handover are instantanious and move

Core network No delay, no routeing

Server The server provides VM and DC models that encompass, inucrred VM migration performance degredation in DC and in VM, resulting in a different service time.
Possible service hosting schemes:
- One service model, one VM is empolyed to host that service for each user.
- One service model, each VM hosts multiple but each number of users, behaving as multiple services while still being compatible.

At all placement modes:

- Measure RTT for all packetc at UE
- Measure DC load
- Measure ratio of requests generated vs. processed in mobile cloudnode
- Identify the incurred VM migration load

V. Results

VI. Conclusions

VII. Future research

- Optimal service/VM migration/placement in relevant topology
- Performance in LTE newtork topology using LTE-SIM [4]

References

[1] Abhishek Chandra, Jon Weissman, and Benjamin Heintz. Decentralized edge clouds. *Internet Computing, IEEE*, 17(5):70–73, 2013.

[2] Ericsson. Ericsson and akamai establish exclusive strategic alliance to create mobile cloud acceleration solutions. Press Release, February 2011. http://www.ericsson.com/news/1488456.

[3] Zhen Liu, Nicolas Niclausse, and César Jalpa-Villanueva. Traffic model and performance evaluation of web servers. *Performance Evaluation*, 46(2):77–100, 2001.

[4] G. Piro, L.A. Grieco, G. Boggia, F. Capozzi, and P. Camarda. Simulating lte cellular systems: An open-source framework. *Vehicular Technology, IEEE Transactions on*, 60(2):498–513, Feb 2011.