Future Cellular Networks Empowering Mobile Cloud Computing

Niko Kortström
Department of Computer Science
University of Helsinki
Finland
Email: niko.kortstrom@cs.helsinki.fi

***Abstract*—The abstract goes here.**

I. Introduction

Introduction goes here.

II. Mobile Cloud Computing

Mobile devices lack the computational power residing in personal computers. According to Satyanarayanan et al. this will always be the case because improving size, weight and battery life are higher priorities in designing them[5]. Of course while a design focus, battery life is still an issue when using mobile devices. There are also many other issues like problems with connectivity. Most applications can still be ported to run on mobile devices. Many of them are simple enough that the lack of resources is not an issue at all. Sometimes it can also be a good solution to limit the usability of an application to better fit it into its new environment. This can mean offering only some qualities of the desktop app. An approach like this can be called creating companion apps. For example a huge multiplayer games companion app can simply offer a connection to the in-game chat or your characters inventory.

A hot topic called internet of things can benefit a lot from mobile cloud computing. When computational capacity is embedded into things not originally build for it, their resources are most likely close to non-existent. Using them as mere thin clients, providing inputs to applications running on the cloud is a good option until computational resources of these things evolves.

There are numerous problems that can't be solved solely on the limited resources of modern mobile devices. Some of these problems are image processing, natural language processing and multimedia search[1]. This is where mobile cloud computing comes into play. Using a much more powerful device or collection of devices, in other words a cloud, to carry out the brunt of computing solves this issue. In this situation, a mobile device could possibly be used just as a mean to display results of computation. There are several types of clouds that mobile devices can make use of.

A. Remote Cloud

This type of mobile cloud computing is mostly similar with traditional cloud computing. A user offloads some computations to a powerful remote cloud. The difference is that in this case the user interacts with the cloud via a mobile device instead of, for example a laptop or a desktop computer. Everyday services that make use of the remote cloud include some of the most widely used applications and web sites in the world, such as Facebook, Google search and Outlook.

B. Cloudlets

Cloudlets form a far less known type of mobile cloud computing. Compared to remote clouds, cloudlets can be seen as less powerful but more closely located clouds. Whereas remote clouds can be seen as huge clusters or warehouses full of computational power, cloudlet can be just a single laptop moving with the user or a desktop computer situated at a public location.

C. Cloud of Mobile Devices

Mobile devices can also be used to form clouds amongst themselves. For this to be possible, users must be willing to submit their mobile device's computational resources to be used by someone else. Another possibility is that multiple users are interested in the result of same computational task. In this case mobile devices will divide the task into smaller pieces and distribute it between the interested mobile devices. The computational capacity of mobile devices is increasing rapidly. Much of this capacity is often not being used. Having a way to harness this unused power could be a great benefit to both, the users offering and making use of it.

Main benefit of the more closely located clouds is not having to use long range connections. There can be many different reasons that it just isn't a valid option at the time to make use of remote data centers. It might be that it is too battery draining, pricey or time consuming to transmit over long ranges. There are also cases that long range connectivity isn't available as depicted in the disaster scenario of Wang et al.[2].

Bringing cloud computing closer to the user has been envisioned to happen via public cloudlets located as densely as Wi-Fi access points in todays world[5]. This would require numerous different parties to want to establish cloudlets, meaning that they need to gain some sort of benefit out of this. Wi-Fi access points are often deployed in public places to attract customers and offer them something to do while waiting for example. Similar to this case, if applications making use of cloudlets are interesting enough, people could seek out vicinity of cloudslets. Obviously to gain financial benefits, cloudlet owners could also sell their customers' personal information to third parties or owners of different facilities could allow third parties to set up their own cloudlets a fee.

Probably the most well known peer-to-peer systems are BitTorrent applications used to distribute entertainment. An application making use of a cloud of mobile devices can also be seen as a kind of peer-to-peer system. BitTorrent systems often encourage users to upload their already downloaded files to other users. Offering them higher download data rates in return is one widely used incentive. An example of a possible incentive[3] in case of mobile computation distribution is for service providers to offer discounts for users who distribute computation in cases of congested network bandwidth.

There are also reasons why users should offload some of their computation when possible, even if they aren't paid for it. Usually offloading computation directly helps with an inherent limiting factor of mobility: battery life. This is quite clear as heavy computation consumes battery fast, minimal computation consumes it a lot slower. However, the issue isn't quite as simple since we also have to account for the amount of battery used by network connection. Connections can limit the possibilities of mobile devices cloud usage if not available but they can also consume a lot of power when used. We must make sure that energy cost of offloading computation isn't too high compared to the energy save.

Trade-offs of mobile cloud computing must be considered. For this purpose there are a lot of different models[1] to closely look into the cost-benefit relationship of mobile cloud computing. These models can possibly be used by humans to deside if it would be useful to offload some of the computation of a specific application. Since benefit of offloading computation is often environment dependant, it is also useful for the application itself to be able to determine whether and how to offload.

When discussing moving computation from one entity to another we obviously need to have some kind of method to transfer the data efficiently. Three different methods are considered suitable for this purpose[1]: client-server communication, virtualization and mobile agent. Client-server communication requires the cloud to have some kind of pre-installed interface available for the client to call. When using virtualization, an image of the current memory state of the client is transferred as an image to the server where its execution is resumed. Mobile agent approach includes splitting up the program code from suitable points and sending it ahead to the cloud to be executed. These points can be chosen by the programmer beforehand or the application can for example use an analyzer to find them[2].

Distributing computing to multiple mobile devices also offers other possibilities. One such is also distributing the applications input or sensing capabilities. Distributing a problem to be solved with the help of multiple users is called crowdsourcing. Users can collectively form larger data sets to be processed. Larger data sets often mean more accurate and reliable results, no matter what data we are talking about. There are many[4] possible applications that could benefit greatly from such an approach.

Modern mobile devices are able to collect enormous amounts of information about our habits, schedules and interests etc. We also use mobile devices to access vital services such as banking. Making applications residing in mobile devices interact with the cloud, directly causes additional security threats to both data about our everyday life and different kinds of personal information. Making sure that no data is disclosed to unwanted parties is a requirement that should absolutely be met before making use of cloud in an application. Additional aspect to consider is that in traditional cloud computing, cloud is usually owned and maintained by some well knows corporation that is considered quite trustworthy. In cases of cloudlets and mobile clouds, targets of offloading cannot be always known beforehand. This makes ensuring security and privacy even more difficult.

All these different kinds of clouds, or more precisely combining them, also means rethinking the architecture of applications that we develop. Not only applications might have to be able to choose whether to offload computing, they may also need to pick the best destination for offloading. Having different layers of cloud is a new factor to keep in mind. Programmers need to be able to make applications smart enough for them to get the maximal benefit. This isn't simple as there are many things to compare, like available resources, data rates and network congestion.

Cloudlet and mobile device solutions also contain additional challenges due to mobile devices moving constantly along with their users. Some of the great benefits of these approaches are saving energy and radio access bandwidth by using bluetooth for example. Transmitting the results of offloaded computation while it is still possible, before the device that issued the task moves out of range is essential. How to know make sure this happens is a great challenge for application designers[3]. Losing all the results of offloaded computation right before it's finished can result in much worse processing time, compared to not distributing computing at all. One way to solve this problem is location based offloading[4]. Of course querying users' location adds to the computational needs so it would have to be used sparingly. This is again one problem that could use a powerful artificial intelligence to be solved.

III. Improving Mobile Networks

At the moment mobile networks are using the fourth generation techology, LTE. In essence, user equipment (mobile phone etc.) connect to the radio access network. Radio access network is communicating with the mobile core that is connected to the Internet. This is a simplified version of the way our mobile phones get access to the Internet in our everyday life.

A. Requirements of Fifth Generation Mobile Networking

Fifth generation mobile communication technology is to be released for consumer use in 2020. There are numerous important and really challenging qualities that are required to be met before the next generation cellular communications techology can be released. 5g standards are yet to be created although the work on them should start fairly soon (http://www.ericsson.com/research-blog/lte/release-14-the-start-of-5g-standardization/). Still, many expectations have already been visioned.

Gohil et al. [6] offer their views on what will and should be included in next generation radio access networks. This paper seems like an early review of their opinions rather than evidence based research as to what 5g will really be. It still offers a good glance at the subject and introduces some other related work. Some of the technical qualities highlighted in the paper include seamless cooperation of different radio access technologies, combining cellular and ip connectivity and new protocol stack architecture which all sound interesting and powerful concepts. They also list some practical features associated with 5g. Advanced billing interfaces, large data rates, better consistency and reduced latency are some of these.

Andrews et al. [7] offer similar but more recent and in my opinion much more accurate description of the fifth generation of cellular networking. Requirements set for 5g are described pretty accurately in the paper. They include areas of data rate, latency and energy and cost efficiency. It is also stated that while great challenged to be achieved, all of the goals don't necessarily have to work in practise at the same time. Example is given that streaming high-definition video needs the high data rates but can do with a higher latency. Then again the situation is reversed for driverless cars, low latency is crucial but data rates can be lower. Coming of internet of things also suggests a need for supporting huge amounts of connections from really heterogenous device base. Low rate device-to-device connections functioning side by side with high rate mobile user connections require a new approach to mobile core management and control functions.

Three clear requirements are set for the data rates needed by ever-advancing technology in the near future.

- Data rates should be increased by approximately thousand times for each area unit.
- The worst reasonably expectable user data rate should range from 100 Mbps to 1 Gbps. At the moment this value is around 1Mbps so the advance must be at least hundred fold.
- Best possible rate, while mostly a marketing number, should also be increased greatly, to tens of gigabits per second.

Fourth generation radio access networks offer roundtrip latencies of around 15 ms. It is stated that current applications can tolerate this well but some of the real time applications visioned for the era of 5g networks need faster response times. Examples would be virtual reality interaction that should always fulfill certain rates in order for human mind not to register responses being a little delayed. Roundtrip latency of around 1ms should be delivered to be enough for all the applications envisioned for near future.

Another much discussed subject is energy efficiency. It would obviously be ideal to reduce energy and financial costs but it is stated that this isn't on the top of requirements of 5g. It is, however, required that the costs wouldn't rise. It is stated [7] that environmental, logistical, cost ad battery technology cannot withstand ever-increasing power consumption. In practise this means that as data rates are increased by for example a factor of hundred, energy and monetary cost per bit should be reduced to one percent of the current cost to keep the overall cost constant.

B. Reaching for the Goals

Many techniques designed to overcome these challenges. There are many promising approaches have already been developed quite far. Still, as even the standards are missing, the used technologies are under development and quite surely not set in stone yet.

For energy efficiency, different approaches can be combined to achieve the maximal gains in. Some of the solutions suggested are: [7].

- allocating resources in a energy efficient way
- not having unnecessary basestations and running algorithms that turn basestations off when not needed and on again on request
- using renewable energy to power basestations, for example by attaching solar panels to them
- energy efficient hardware solutions

Some ways for operators to deal with financial costs related to moving into 5g are also presented. Needed additional infrastructure could be constructed by some outside parties. Right to use them for offering connections to end users would then be rented to operators. Increased backhaul requirements gain a lot of help because of overall spreading of fiber connections. Some other ways to help deal with them are better wireless backhaul solutions and optimizing the backhaul operations.

The paper [7] lists ultra-densification, mmWave and massive multiple-input multiple-output (massive MIMO) along with virtualization as the defining factors of future radio access networks. The combination of these approaches are seen as the keys to provide ubiquous, high-rate, low-latency connections to all mobile users.

Densification means simply making cells served by a radio access point smaller. This approach has been used in previous radio access improvements and ultra-densification just means taking it to yet another level. The most important benefit of this approach is that spectrum reserved by a user connection via a basestation accommodates a smaller area. In other words, smaller coverage area per basestation means less users connecting via that particular basestation. Obviously this means that there are more basestation resources and more wave frequencies available for each user.

The small cell approach doesn't come without downsides. Some of the associated challenges listed are supporting mobility between cells and increased costs of installing and maintaining all the basestations. Keeping the energy and financial costs low enough for the business to be profitable, small cell devices have to be cheaper and more energy-efficient than today's basestations.

Smaller areas served by a single unit mean more need to transport computing from one basestation to another while the mobile user is on the move. Efficient alghorithms are needed to handle the additional computing caused. Another challenge that cannot be avoided is adding yet another technology into the mix. The fact that the system must be able to support more technologies at the same time increases required complexity.

At the moment wireless communication is mainly propagated via microwave frequencies starting from hundreds of MHz and ranging to few GHz. This isn't too big of a spectrum and it is becoming increasingly congested. Improvements like radio cloud and ultra-densification do help with this issue but don't solve it completely.

The new wave frequencies planned for next generation wireless networking reside at 30 to 300 GHz range. This range is called mmWave because wave length at these frequencies varies from one millimeter to ten. The mmWave frequencies have not been used before because it has been considered unfit for longer range wireless communication. This is due to it being very susceptible to interferance from weather effects and various objects, such as walls, between the transmitter and the receiver.

Rapport et al. [8] did extensive research on mmWave, its viability and effectibility. They conducted important experiments with mmWave both indoors and outdoors. Different kinds of blockages and interferences were also included. Their tests verified that buildings indeed affect the waves a lot as results from Texas were better than those from New York. However with proper network planning and density the mmWave should work just fine for propagating data. Their tests concluded that 200 meter cell radius is enough to achieve good coverage using mmWave.

Some other challenges related to benefiting from this technology are narrowness of the beams. In addition to coverage challenges, initializing the connection between mobile users and the basestation becomes harder. Building the systems to provide coverage while keeping the cells from interfering with each other is no simple task with beams that can be described to resemble flashlights [7]. The narrowness of the beam also forces the user equipment to possibly scan numerous different directions before locating the the beam.

In practice MIMO means making use of multiple antennas on both, transmitting and receiving ends. More antennas transmitting can be used to reach higher rates. Modern technology also allows multiple basestations to cooperate, even allowing them to turn some of the interference into signals containing data. MIMOs are said to use two to four antennas for transmitting to the mobile device and maximum of eight to transmit within one basestations area.

Envisioned massive MIMOs can consist of hundreds of antennas within one basestation sector. Realizing basestations with these new MIMOs requires a great deal of restructuring the physical architechture of a basestation. Another thing to consider carefully is how to have MIMOs work together with new small cell and mmWave technologies to achieve maximal benefits.

Virtualization has already been somewhat taken into use with mobile core but there is still a lot of work to be done to extend it farther. There are two important terms related to virtualizing the mobile networks, software defined networking and network function virtualization. Implementations before network function virtualization have always closely paired network functions with the hardware. Now the aim is to move much of these functions to be executed in the cloud, offering improved flexibility and management possibilities. This is the essential meaning of network function virtualization. Software defined networking provides the possibility for external applications to modify the network's capabilities, creating intelligent networks.

Lindholm et al. [9] have studied how to implement the virtualization in practise. In other words, they proposed which parts of the LTE mobile network to virtualize should be virtualized and how to possibly change it's overall architecture. Their approach is based on publish-subscribe model where mobile core subscribes to changes in some functions of radio access network and vice versa. This seems like a well suited and efficient way to handle the communication in a modern mobile networking system.

On the hardware level virtualization allows a new design approach, where a traditional radio access point is somewhat divided in two. Remote radio heads handle physically exchanging data with the user equipment. Data centers are used to connect to multiple remote radio heads and handle the processing for them. This is described in more detail for instance by Rost et al. [10].

They view the mobile network stack divided to seven layers:

- Network management
- Admission/congestion control
- Radio resource management
- Medium access control
- Physical layer
- Radio frequency

As much as the first six of the are seen to be possible to be virtualized in future implementations. They also refer to flexible functional splitting, meaning that every layer shouldn't just be automatically virtualized. As discussed, virtualization allows creation of more intelligent networks. This is where it can be realized, as we can choose which parts would be useful to be virtualized at runtime. This is yet another gain from the new architecture, in a way to make the architecture itself better.

As we have discussed, virtualization is a big part of the ongoing mobile network development. Most have heard of system architectures consisting of different services, such as infrastructure as a service, platform as a service and software as a service. Nikaein et al. [11] go as far as to introduce a similar consept related to mobile networks, radio access network as a service. This is a term that to me really shows demonstrates the possible gains from this new approach. The amount of flexibility and managebility gained seems to get new proportions when approached like this.

In addition to the new qualities introduced to improve mobile networks, there are also possibilities to improve the existing ones once the architectural changes have been implemented. One such approach is improving the caching performed by radio access networks [12]. Ao et al. introduce their views on how to increase benefits from caching with the help of small cells.

Multiple radio access points covering an area mean multiple caches available to deliver cached data quickly to the receiving users. They give an example of caching video files based on their popularity (the most popular video being cached to most radio access points). This way some of the ever-increasing data download can be handled even more efficiently in a way that could possibly reduce the stress from core network by a lot.

IV. Mobile Cloud Computing in the Context of Improved Cellular Networks

So how does this all pan out? How do the improved mobile networks affect existing and new mobile cloud computing applications? Is there some considerations to be acquired from the field of mobile cloud computing for designing the next generation networks? This is of course just one technological relationship but in my opinion it's important to discuss the possible interactions of the two topics.

First I'd like to discuss the cloudlet approach to mobile cloud computing. I somewhat question the need for public cloudlets for computation offloading. It is definately true that computational needs are exceeding the capabilities of mobile devices. However I would argue that with the coming of 5g networking many of the benefits of cloudlets can be answered with traditional cloud computing.

Reduced latencies and increased data rates make transmitting virtual machines or remote calling functions from longs distances even more viable. Perhaph viable enough to answer the needs of present and even near future applications. The challenges of the cloudlet approach, building the infrastructure and modifying application architectures to be able to use cloudlets effectively are no small tasks. In addition, the computational power of huge datacenter far exceeds that of cloudlets. It might be a lot more practical and even computationally effective to stick with offloading to remote cloud instead of opting for this new approach.

Standardization for 5g is most likely based on exactly the requirements visioned for future needs of mobile computing that the cloudlet and mobile cloud approaches are also trying to solve. If research indeed is successful in meeting the set goals, it would be intuitive to think that there is no real need for other solutions.

Having no need for widespread implementation of the newer mobile cloud solutions doesn't still mean that there wouldn't be specific scenarios where they would be of great benefit. Local proximity is already an important point of many applications (for example some dating applications). Taking this a step further and opting only for close range connections could be valid selling point. Some cloudlet applications could also be visioned, perhaps even own applications set up in public cloudlets. This could mean having to join others in a specific public location to collaborate in whatever is the goal of said application. This would introduce yet another social factor to world of mobile computing.

One reason to offload where ever possible is saving money. As discussed before avoiding roaming costs or even saving on data transfers when you are paying by usage. Still, widespread free wi-fi connections offer possibilities for traditional cloud usage, albeit data rates are often less than optimal.

Another aspect speaking for close range computation offloading is saving battery. Closer range connection techniques do require less battery and in this way it would be really useful to be able to offload close whenever you can. Still, it is hard to determine if benefits overcome the costs in case of cloudlets.

In my opinion mobile clouds offer a more viable option compared to cloudlets. The thing is, when the computation is offloaded to other mobile devices, the infrastructure already exists. It should also be quite clear that there are a lot of spare resources available most of the time. There is no question that a way to harness these resources would profit all parties.

Of course a lot of work needs to be done to design applications that can make use of other mobile devices' resources. Also is is obvious that mobile devices don't possess even the computational powers that cloudlets do. Their sheer numbers might however prove to offer similar or even better possibilities, should people opt to participate in enabling this approach.

Peer-to-peer applications are really widely used and there is also extensive research on the subject. This should prove to be of great help in implementing clouds of mobile devices. Already existing research can offer good guidelines to start designing your application from or you could possibly even reuse some protocols.

The previously discussed disaster scenario is an obvious example of a case where the mobile network infrastructure is of no use. This is where cloud of mobile device gets an opportunity to shine. As cloudlets could be of great effect as emergency personnel could move them around with them, mobile networks could also be set up quite quickly.

Mobile networks are gaining a lot of flexibility via virtualization thus, it could even be viable to set up remote radio heads after disasters. Having most of the network functions already operational in a data center and just installing the radio frequency transmitters/receivers could possibly be done fairly quickly. Obviously it would not be as quick a way to respond as having mobile devices form their own cloud, it's something to consider.

Security issues, disaster scenario or just every day usage, are probably higher when offloading to peers and cloudlets set up by non-experts. They might not be high profile targets and so don't face as many threats but it's hard to see their security maching that of well established and well reknown cloud providers.

Challenges related to near-constant mobility of the mobile users are also a challenge that is already being solved in cellular networks. In case of mobile clouds it would have to be figured out all over again. It would probably also be a lot harder, as there might not be similar high level management layers available.

V. Conclusion

The conclusion goes here.

References

[1] N. Fernando, S. W. Loke and W. Rahayu, *Mobile cloud computing: A survey*. Future Generation Computer Systems Volume 29 Issue 1, January 2013.

[2] Y. Wang, I. Chen and D. Wang, *A Survey of Mobile Cloud Computing Applications: Perspectives and Challenges*. Wireless Personal Communications: An International Journal Volume 80 Issue 4, February 2015.

[3] M. Miluzzo, R. Cceres and Y. Chen, *Vision: mClouds Computing on Clouds of Mobile Devices* . MCS '12 Proceedings of the third ACM workshop on Mobile cloud computing and services, 2012.

[4] G. Chatzimilioudis, A. Konstantinidis, C. Laoudias and D. Zeinalipour-Yazti, *Crowdsourcing with Smartphones*. IEEE Internet Computing Volume 16 Issue 5, September 2012.

[5] M. Satyanarayanan, P. Bahl, R. Caceres and N. Davies, *The Case for VM-Based Cloudlets in Mobile Computing*. IEEE Pervasive Computing Volume 8 Issue 4, October 2009

[6] A. Gohil, H. Modi and S. K. Patel, *5G Technology of Mobile Communication: A Survey*. Intelligent Systems and Signal Processing (ISSP) International Conference, 2013

[7] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong and J. C. Zhang, *What Will 5G Be?*. IEEE Journal on Selected Areas in Communications Volume 32 Issue 6, June 2014

[8] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi and F. Gutierrez, *Millimeter Wave Mobile Communications for 5G Cellular: It Will Work!*. IEEE Access Volume 1, May 2013

[9] H. Lindholm, L. Osmani, H. Flinck, S. Tarkoma and A. Rao, *State Space Analysis to Refactor the Mobile Core*. AllThingsCellular '15 Proceedings of the 5th Workshop on All Things Cellular: Operations, Applications and Challenges, August 2015

[10] P. Rost, C. J. Bernardos, A. De Domenico, M. Di Girolamo, M. Lalam, A. Maeder, D. Sabella and D. Wbben, *Cloud Technologies for Flexible 5G Radio Access Networks*. IEEE Communications Magazine Volume 52 Issue 5, May 2014

[11] N. Nikaein, R. Knopp, L. Gauthier, E. Schiller, T. Braun, D. Pichon, C. Bonnet, F. Kaltenberger and D. Nussbaum, *Demo Closer to Cloud-RAN: RAN as a Service*. MobiCom '15 Proceedings of the 21st Annual International Conference on Mobile Computing and Networking, September 2015

[12] W. Ao and K. Psounis, *Distributed Caching and Small Cell Cooperation for Fast Content Delivery*. MobiHoc '15 Proceedings of the 16th ACM International Symposium on Mobile Ad Hoc Networking and Computing, June 2015