Micro Data Server using Raspberry Pi for Video Streaming

Research Proposal

P.J.E. Velthuis
University of Twente
P.O. Box 217, 7500AE Enschede
The Netherlands
p.j.e.velthuis@student.utwente.nl

ABSTRACT

This document describes a research proposal in the area of cloud computing services. Cloud computing is a trend IT that customers move computing and data away from desktop and portable PCs into large data centers. These data centers require a lot of power and cooling. Nowadays 30 % of the data coming from these data centers is video streaming. The Raspberry Pi is a device that might improve these problems. The main research question is therefore: Is it possible to improve availability algorithms for video streaming by using small scale Raspberry Pis data center? To answer this question a micro Raspberry Pi video streaming service will be build.

Keywords

Cloud computing, Raspberry Pi, micro data server, video streaming, load balancing

1. MOTIVATION

The amount of cloud computing services is increasing fast [4]. But despite the attention from the research community, research and development of Cloud Computing services is still in its early days [38].
Cloud computing is a trend IT that customers move computing and data away from desktop and portable PCs into large data centers [11]. In the future most internet users access internet services over lightweight portable devices. This requires a lot of data bandwidth. Video streaming is something that requires a lot of bandwidth. This bandwidth usage causes bottlenecks and small data servers can be build after these bottlenecks. This makes research on this topic very interesting.
In this research a small data server consisting of Raspberry Pit's will be investigated. A reason for this is that data centers require a lot of space and cooling. There are now new technologies such as for example the powerful ARM processor. Many companies want to explore the possibilities of for example the Raspberry Pi and his ARM Processor [29]. Some data centers already offer some cloud computing using the Raspberry Pi. A reason for this is the small location that the Raspberry Pi needs, and the low power usage of a Raspberry Pi [34, 29]. A Raspberry Pi has a power usage between the 3-5 Watt. A normal server has a power usage between the 75 and 250 Watt [30, 6]. However the Raspberry Pi has less computing power then a normal server. The low power consumption and its computing power could mean that it is better to use a Raspberry Pi for specific small tasks that do not demand a whole server. During this research there will be a investigation on the performance of the Raspberry Pi as a small data server.
A Raspberry Pi can be the small data center for the future [38]. The Raspberry is a relatively cheap device for 40 euro. This makes it cheaper to do research in compared to a normal server. Building a cloud like this can be a cost effective scale model[38]. It's a ideal testbed for testing distributed software.

2. SPECIFIC PROBLEM

The Netherlands have one million subscribers for Netflix [10]. Netflix is a popular video streaming service that makes HD movies watching possible. For this Netflix makes use of a content distribution network (CDN). On the internet this is known as a on demand service [2]. Netflix makes use of MPEG-DASH a protocol that makes streaming over HTTP possible [24]. The problem is that Netflix is responsible for 29.7% of the peak downstream traffic in US [2].
In the Netherlands more and more users are making use of video streaming as Youtube and Netflix. This makes it interesting to do more research in video streaming by cloud computers. Most of the research nowadays happen on expensive large servers [38]. For this it can be very useful to see if it is possible to do research on a small computer like a Raspberry Pi. The problem is that we do not know if video streaming on a small raspberry pi cluster is possible.

3. RESEARCH QUESTIONS

The main research question

Is it possible to improve availability algorithms for video streaming by using small scale Raspberry Pis data center?

This research will be part literature study and a part of it will be building a small Raspberry Pi cloud to measure the availability of the video streaming. Here below there are several questions to come to a good answer for the main research question.

3.1 Why are small scale data centers being used?

In this research there will be a small scale cloud data center with the Raspberry Pi. The main motivation behind this is that it is ways to expensive to simulate large scale cloud computing. Resources for this motivation: [8, 30, 9]
Another reason is the small location that is needed using a small scall data center. On the size of the data center is currently done a lot of research [29].
Small scale cloud computing is cloud computing with smaller computation amounts than normal. This can mean that it is better for different tasks. More motivation is that it does not use a lot cooling and power [38]. Resources available for small scale cloud computing: [29, 4, 35, 1, 8, 38, 7, 33, 16, 12, 13, 28, 25, 6, 9]

3.2 Is it useful to use Raspberry Pis for video streaming?

The Netherlands have one million subscribers for Netflix [10]. Netflix is a video streaming service that makes HD movies watching possible. This requires a lot of internet bandwidth and it could be useful to do the streaming in small scale data centers nearby. To investigate what Netflix really is there are the following resources: [10, 2]
Netflix makes use of MPEG-DASH a protocol that makes streaming over HTTP possible. In this research there will be an investigation in the video streaming as a cloud service in a small scale data center. To investigate what that is there are the following resources: [20, 37, 32, 19]

3.3 What is the perfect setup for a Raspberry Pi video streaming cluster?

This research question will look into previous cloud cluster projects: [1, 8, 38, 7, 9]
From this a perfect setup for the hardware of the Raspberry Pi's will be build. There will be taken a look into the different software that has to be used. For example: [20, 37, 5, 23, 26, 3]

3.4 How does the availability change using different load balancing techniques with the Raspberry Pi?

With this subquestion we will look into the different load balancing techniques. In a video streaming service it's very important to have high performance and availability. That's why it's interesting to do more research on this. Resources for load balancing on Raspberry Pi: [36, 17, 27, 31, 22]
Things that are interesting to look into are for example bandwidth usage, power usage, i/o throughput. For this the following resources will be used: [23, 26, 3]

4. RESEARCH METHODS

This research will investigate a cloud computer consisting of Raspberry Pi's. The research method for this would be the Design Science research method. this is a method to solve field problems. This research will make use of the design Science method proposed by Hevner [15].

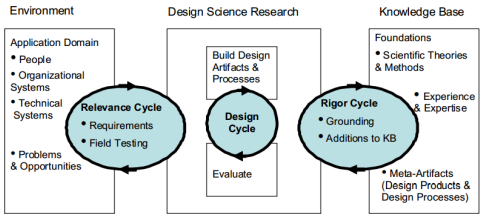

Figure 1: Design Science

For this design research the research will specify the requirements, existing solutions, new solutions and the result of this new solution. After that a conclusion will be given.

5. RESEARCH APPROACH

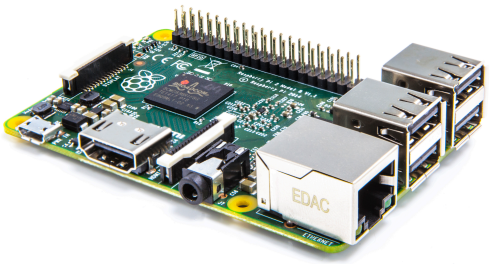

Figure 2: Raspberry Pi 2 model B

Ethernet	100 MBps
USB	4 x USB 2.0
Video out	HDMI 1.4
Audio	2 x analog
CPU	900MHz quad-core ARM Cortex-A7
card slot	Micro SD

Table 1: Specificaties

For the experiment we need a Raspberry Pi B or Raspberry Pi 2 that can make use of the Ethernet. This is because a Raspberry Pi can be used as a webserver and this is needed to have this in order to make video streaming possible. This webserver will make video streaming possible. Beside the Raspberry Pi a several other things will be needed. For this we have a cost structure:

Device	cost piece	use	amount	total cost
Raspberry Pi 2	39,50	cloud cluster	8	316
UTP kabel	1,10	data kabel	8	8,80
Micro usb kabel	1	charging	8	8
switch	40	internet	1	40
				372,80

Table 2: Cost summary [18]

So first he video streaming webserver will be made. After this we will make a cluster with one Raspberry Pi as a load balancer and the other two as a video stream pi. The university of Southampton has made a cluster possible [8]. So I think it is also possible for video streaming. With this there several performance tests can be done. For example

some test in streaming in different quality with load balancing. The setup of the raspberry Pi can be seen in the figure below 3.

pi setup.jpg pi setup.bb

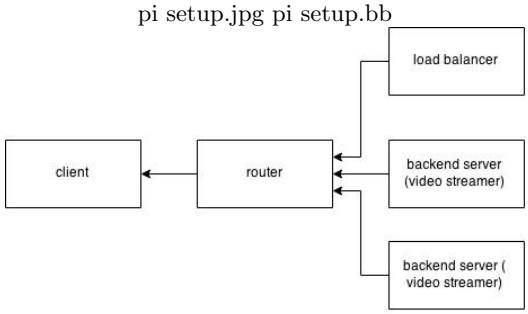

Figure 3: Setup

Input
Data from SD card
Data from Internet
Electrical Power
Cooling
Measurement tools
System
Operating system (Raspbian)
Measurement software
Processing power
Output
data over network (video stream)
data over HDMI

Table 3: Variables

The system will make use of several variables 3. We can distinguish three types which we need to keep in our mind during the research. We have input variables that define the input. The system has not a lot of variables, because you can choose for example a operating system, but not change the software in the operating system. By using the software of the system we can get the different result in the output variables.

6. PLANNING

Week	**Start Date**	**Activity**
6-13	6 Feb	Collect literature and software
12	16 March	Deadline peer review and proposal
14	23-30 March	**Deadline proposal**
15	23-30 March	**Accept Reject proposal**
15-14	30 March	Start building test environment
16	6 April	Answer subquestion 1
17	13 April	Answer subquestion 2
18	21 April	Answer subquestion 3
16-18	6 April	working on test environment
17	1 May	Answer subquestion 4
18	4 May	conclusion and check coherency
19	11 May	Finalize paper
22	1 June	**Deadline paper**
23-25	8 june	presentation preperation
26	22 June	**Conference**

Table 4: Planning

Here above is the table 4. The main research question is split up into several sub questions. These subquestions have each one week for answering. After that there is time to write the conclusion and answer the main question. Then there is still some time left to finalize the paper.

7. STATE OF THE ART

The amount of cloud computing services is increasing fast [4]. Despite the attention from the research community, research and development of Cloud Computing services is still in it's early days [38]. Today there is a lot of customers that use video streaming and this will become even more. Customers want to have everything has to be accessible through the network around the clock [39]. One thing that people want on demand are there film series. Currently Netflix has 30% of the downstream in the United States [19]. This downstream is huge and that means that there is optimization in this branch possible. To make the video on demand service of Netflix possible they use servers from Amazon.

The Raspberry Pi is made for research and education purposes [14]. The Raspberry Pi is a cheap device costing around 40 euro. It has a power consumption of only three watt and has quite good processor. For this reason it might be very useful for small scale cloud computing. The raspberry Pi is a small device and is excellent for a lot of small devices in a relatively small place. The Raspberry Pi doesn't use cooling and it can be used for very rapid elasticity and on-demand self-service. Because of it's cooling features and low power consumption it might be a good alternative for the nowadays high power consuming data centers. The raspberry pi is really small so it is a lot easier to place extra pi in a data center compared to a normal server. The Raspberry Pi has one drawback and that is that the processing power is relatively low. By using load balancing there will be taken a look at what load balancing techniques are successful for the Raspberry pi.

In this research we will make a small raspberry cloud to do research on video streaming using HTTP. This will be done using the design Science method proposed by Hevner [15]. This research will try to find out if the raspberry has enough processing power and bandwidth to be a suitable streamer to multiple web clients. The cloud performance can now be better researched, because everything is at one location and the information about what processes are running on the pi are well defined. In this way we can take a better look at algorithms used in video streaming. This is mostly because there is not a lot of overhead.

8. REFERENCES

[1] P. Abrahamsson, S. Helmer, N. Phaphoom, L. Nicolodi, N. Preda, L. Miori, M. Angriman, J. Rikkila, X. Wang, K. Hamily, et al. Affordable and energy-efficient cloud computing clusters: The bolzano raspberry pi cloud cluster experiment. In *Cloud Computing Technology and Science (CloudCom), 2013 IEEE 5th International Conference on*, volume 2, pages 170–175. IEEE, 2013.

[2] V. Adhikari, Y. Guo, F. Hao, M. Varvello, V. Hilt, M. Steiner, and Z.-L. Zhang. Unreeling netflix: Understanding and improving multi-cdn movie delivery. In *INFOCOM, 2012 Proceedings IEEE*, pages 1620–1628, March 2012.

[3] Apache. ab - apache http server benchmarking tool. `http://httpd.apache.org/docs/2.2/programs/ab.html`, 2015. Last accessed March 2, 2015.

[4] M. Armbrust, O. Fox, R. Griffith, A. D. Joseph, Y. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, et al. M.: Above the clouds: a berkeley view of cloud computing. *Communications of the ACM*, 2009.

[5] askubuntu. Changing permissions for folders and files in /var/www/. `http://askubuntu.com/questions/416528/changing-permissions-for-folders-and-files-in-var-www`, 2014. Last accessed March 8, 2015.

[6] A. Beloglazov, J. Abawajy, and R. Buyya. Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing. *Future generation computer systems*, 28(5):755–768, 2012.

[7] A. Beloglazov and R. Buyya. Energy efficient allocation of virtual machines in cloud data centers. In *Cluster, Cloud and Grid Computing (CCGrid), 2010 10th IEEE/ACM International Conference on*, pages 577–578. IEEE, 2010.

[8] P. S. Cox. Steps to make raspberry pi supercomputer. `http://www.southampton.ac.uk/~sjc/raspberrypi/pi_supercomputer_southampton_web.pdf`, 2012. Last accessed February 20, 2015.

[9] S. J. Cox, J. T. Cox, R. P. Boardman, S. J. Johnston, M. Scott, and N. S. Obrien. Iridis-pi: a low-cost, compact demonstration cluster. *Cluster Computing*, 17(2):349–358, 2014.

[10] de Volkskrant. Nederland op vijfde plek met bijna miljoen netflix-abonnees. `http://www.volkskrant.nl/media/nederland-op-vijfde-plek-met-bijna-miljoen-netflix-abonnees~a3851031/`, 2015. Last accessed February 27, 2015.

[11] M. D. Dikaiakos, D. Katsaros, P. Mehra, G. Pallis, and A. Vakali. Cloud computing: Distributed internet computing for it and scientific research. *Internet Computing, IEEE*, 13(5):10–13, 2009.

[12] I. Drago, M. Mellia, M. M Munafo, A. Sperotto, R. Sadre, and A. Pras. Inside dropbox: understanding personal cloud storage services. In *Proceedings of the 2012 ACM conference on Internet measurement conference*, pages 481–494. ACM, 2012.

[13] Dropbox. Share the power of great sync with external collaborators. `https://www.dropbox.com/business/uses/file-sharing`. Last accessed March 2, 2015.

[14] R. P. FOUNDATION. What is a raspberry pi? `http://www.raspberrypi.org/help/what-is-a-raspberry-pi/`, 2015. Last accessed February 27, 2015.

[15] A. R. Hevner. A three cycle view of design science research. *Scandinavian journal of information systems*, 19(2):4, 2007.

[16] C. Höfer and G. Karagiannis. Cloud computing services: taxonomy and comparison. *Journal of Internet Services and Applications*, 2(2):81–94, 2011.

[17] N. HOWTOS. Load balancing with nginx. `http://www.networkinghowtos.com/howto/load-balancing-with-nginx/`, 2013. Last accessed February 27, 2015.

[18] IAPC. Assortiment. `https://www.iapc.utwente.nl/assortiment`, 2015. Last accessed March 8, 2015.

[19] K. W. R. James F. Kurose. *Computer Networking A Top-Down Approach*. Pearson, sixth edition edition, 2012.

[20] jelleke60. How-to: Raspberry pi media server (easy). `http://www.raspberrypi.org/forums/viewtopic.php?t=56149`, 2013. Last accessed February 27, 2015.

[21] T. D. Journal. The datacenter journal. raising data center power density. `http://www.datacenterjournal.com/it/raising-data-center-power-density/`, October 2013. Last accessed February 27, 2015.

[22] koff1979. Raspberry pi owncloud (dropbox clone). `http://www.instructables.com/id/Raspberry-Pi-Owncloud-dropbox-clone/`, 2015. Last accessed February 28, 2015.

[23] LINUX. nmon for linux. `http://nmon.sourceforge.net/pmwiki.php`, 2012. Last accessed March 2, 2015.

[24] J. Martin, Y. Fu, N. Wourms, and T. Shaw. Characterizing netflix bandwidth consumption. In *Consumer Communications and Networking Conference (CCNC), 2013 IEEE*, pages 230–235. IEEE, 2013.

[25] A. P. Miettinen and J. K. Nurminen. Energy efficiency of mobile clients in cloud computing. In *Proceedings of the 2Nd USENIX Conference on Hot Topics in Cloud Computing*, HotCloud'10, pages 4–4, Berkeley, CA, USA, 2010. USENIX Association.

[26] S. Moon. 18 commands to monitor network bandwidth on linux server. `http://www.binarytides.com/linux-commands-monitor-network/`, 2014. Last accessed March 2, 2015.

[27] Nginx. Using nginx as http load balancer. `http://nginx.org/en/docs/http/load_balancing.html`, 2015. Last accessed February 28, 2015.

[28] ownCloud. owncloud. `https://owncloud.org/`. Last accessed March 2, 2015.

[29] PCextreme. Raspberry pi colocatie. `http://raspberrycolocatie.nl/`, October 2013. Last accessed February 27, 2015.

[30] PCextreme. Waarom doen we dit. `http://raspberrycolocatie.nl/why/`, 2013. Last accessed February 27, 2015.

[31] N. Peers. How to build the best raspberry pi media server. `http://www.techradar.com/news/digital-home/media-servers/how-to-build-the-best-raspberry-pi-media-server-1163133/2`, 2015. Last accessed February 27, 2015.

[32] L. Plissonneau and E. Biersack. A longitudinal view of http video streaming performance. In *Proceedings of the 3rd Multimedia Systems Conference*, pages 203–214. ACM, 2012.

[33] L. Qian, Z. Luo, Y. Du, and L. Guo. Cloud computing: An overview. In *Cloud Computing*, pages 626–631. Springer, 2009.

[34] raspberry hosting. Raspberry pi hosting. raspberry pi in the heart of europe. `https://raspberry-hosting.com/en`, 2015. Last accessed February 27, 2015.

[35] M. Richardson and S. Wallace. *Getting Started with Raspberry Pi.* " O'Reilly Media, Inc.", 2012.

[36] E. Sverdlov. How to set up nginx load balancing. `https://www.digitalocean.com/community/tutorials/how-to-set-up-nginx-load-balancing`, 2012. Last accessed February 27, 2015.

[37] tpy's. Live 1080p video streaming from the raspberry pi to browsers using nginx and rtmp. `http://tpy.tw/?p=109`, 2014. Last accessed February 27, 2015.

[38] F. P. Tso, D. R. White, S. Jouet, J. Singer, and D. P. Pezaros. The glasgow raspberry pi cloud: A scale model for cloud computing infrastructures. In *Distributed Computing Systems Workshops (ICDCSW), 2013 IEEE 33rd International Conference on*, pages 108–112. IEEE, 2013.

[39] L. Youseff, M. Butrico, and D. Da Silva. Toward a unified ontology of cloud computing. In *Grid Computing Environments Workshop, 2008. GCE'08*, pages 1–10. IEEE, 2008.