–Independent Work Report Fall,2013–

BeagleCache: A Low-Cost, eMMC-based Caching Proxy for the Developing World

Abstract

The recent release of the BeagleBone Black–a $45, 1GHz computer the size of a credit card and packed with 512MB of DRAM and 2GB of onboard eMMC flash–has put forth a unique opportunity for the development of memory-intensive applications designed to run on a low-cost device. The BeagleBone Black's low cost and size make it the ideal platform for deployment in developing world countries. Furthermore, eMMC is one of the least-researched forms of flash memory. Thus we present BeagleCache, an inexpensive and easily-deployable caching proxy running off an eMMC persistant store.

1. Introduction

In developing countries, internet bandwidth can be both prohibitively slow and expensive compared to that of more affluent countries. For example, according to Du et al.[6], who studied internet traffic at Internet cafes and kiosks in Cambodia and Ghana, the monthly cost of a 512 Kbps DSL connection costs $400. In Ghana, a 512kbps connection costs $650 a month. For comparison, in the United States, a DSL connection of 2 Mbps costs on average only $25 a month. This discrepancy is largely due to the lack of physical links in developing countries, where access is provided through costly satellite or cellphone connections. Acceleration techniques such as Web caching, WAN acceleration, bandwidth shifting, and prefetching have been proposed to aid access in the developing world[8], but often the biggest challenge researchers face in implementing such solutions is in their deployment. Even if the physical transportation of low-cost computers were not a problem, one

would still likely require a trained systems administrator to set up the devices. The BeagleBone Black (BBB) is thus an extremely attractive piece of hardware for such applications because it's smaller than 3.4 inches square and ships with several onboard peripherals, including an ethernet and USB port as well as 2GB of eMMC persistant storage.

In this paper, we investigate the performance of Polipo[1], a lightweight caching proxy, on the BBB. Because the performance characteristics of eMMC are relatively understudied (and at the time of this writing, eMMC has never been studied in a caching proxy application), the purpose of this investigation is two-fold: first, to determine whether or not the BBB's performance running Polipo is suitable for deployment in developing world countries, and second, to study the performance of eMMC as a Web cache.

2. Background

2.1. Flash Memory

3. Related Works

3.1. Caching Proxies for Flash Memory

3.2. Web Acceleration Techniques for Developing World Countries

4. Experimental Setup

4.1. BeagleCache Software

The BeagleBone Black ships with Angstrom Linux, a bare-bones flavor of Linux designed for embedded devices. However, because of both its large support community and convenient package manager, we decided to install the Arch Linux distribution designed for ARM processors[2]. As a filesystem, we originally considered using the filesystem Yaffs (Yet Another Flash Filesystem)[5], which is specifically designed for use in embedded NAND and NOR devices, and was in fact the filesystem originally used with Android devices using eMMC (Android 2.3 switched to ext4

with journaling)[7]. However, because Arch Linux on ARM does not have built in support for mounting a Yaffs filesystem, we decided to use a standard ext4 filesystem without journaling. Lee and Won found in their characterization of IO on Android Smartphones that in typical applications (web browser, SMS, camera, camcorder), ext4 journaling–the process whereby all data intended to be written on disk is first written to an on-disk journal that can be replayed in the case of a crash[3]–represents 40% to 50% of write operations[7]. Thus, between disabled journaling and the eMMC's built-in FTL, we believe ext4 is a suitable filesystem for BeagleCache.

As for our caching proxy software, we decided to use Polipo, an open-source http caching web proxy. Polipo is an extremely small–(GIVE SIZE HERE)–simple, and configurable proxy. The size of its RAM cache can be configured in terms of bytes or objects (or both). It (optionally) writes objects to disk either when it runs out of memory in RAM or when it is idle for a user-specified amount of time. When storing objects on disk, Polipo stores objects in a standard filetree structure (MORE ON THIS). It is single-threaded and non-blocking.

*** MORE *** In order to compile Polipo for the BeagleBone Black quickly, we cross-compiled it using a gcc ARM extension.

4.2. Testing Software

In testing the performance of BeagleCache, we connected the BeagleBone Black to a MacBook Air running OSX 10.9.1 via ethernet. The MacBook served as both client and server, with the client's requests forwarded through the BBB http proxy. Because the MacBook's processor is roughly twice as fast as the BBB's (1.8 GHz Intel Core i5), we believed it would be able to flood the BBB and any observed latency would be on the BBB and not the MacBook. As a server, we wrote a small node.js application that served a 10KB file for any request it received. This file consisted of exactly 10,000 copies of the character 'j'. When stored in memory, Polipo allocated 12K bytes for each of these objects.

As for the client, we decided to use an open-source http benchmarking and load testing utility called Siege[4].

4.3. Paper Formatting

There are no minimum or maximum length limits on IW reports. We are including this template because we think it will be helpful for citing things properly and for including figures into formatted text. If you are using LaTeX [**?**] to typeset your paper, then we strongly suggest that you start from the template available at http://iw.cs.princeton.edu – this document was prepared with that template. If you are using a different software package to typeset your paper, then you can still use this document as a reasonable sample of how your report might look. Table 1 is a suggestion of some formatting guidelines, as well as being an example of how to include a table in a Latex document.

Field	**Value**
Paper size	US Letter 8.5in $\times$ 11in
Top margin	1in
Bottom margin	1in
Left margin	1in
Right margin	1in
Body font	12pt
Abstract font	12pt, italicized
Section heading font	12pt, bold
Subsection heading font	12pt, bold

Table 1: Formatting guidelines.

Please ensure that you include page numbers with your submission. This makes it easier for readers to refer to different parts of your paper when they provide comments.

We highly recommend you use bibtex for managing your references and citations. You can add bib entries to a references.bib file throughout the semester (e.g. as you read papers) and then they will be ready for you to cite when you start writing the report. If you use bibtex, please note that the references.bib file provided in the template example includes some format-specific incantations at the top of the file. If you substitute your own bib file, you will probably want to include these incantations at the top of it.

4.4. Citations

There are various reasons to cite prior work and include it as references in your bibliography. For example, If you are improving upon prior work, you should include a full citation for the work in the bibliography [**?**, **?**]. You can also cite information that is used as background or explanation. In addition to citing scholarly papers or books, you can also create bibtex entries for webpages or other sources. Many online databases allow you to download a premade bibtex entry for each paper you access. You can simply copy-paste these into your references.bib file.

Figures and Tables. Ensure that the figures and tables are legible. Please also ensure that you refer to your figures in the main text. Make sure that your figures will be legible in the expected forms that the report will be read. If you expect someone to print it out in gray-scale, then make sure the figures are legible when printed that way.

Main Body. Avoid bad page or column breaks in your main text, i.e., last line of a paragraph at the top of a column or first line of a paragraph at the end of a column. If you begin a new section or sub-section near the end of a column, ensure that you have at least 2 lines of body text on the same column.

4.5. Ethics

Your independent work report should abide by the basic standards of scholarly ethics and by the Princeton Honor Code. If you have any doubts about how to cite other work, how to quote or include text or images from other works, or other issues, please discuss them with your project adviser or with the IW coordinators.

References

[1] [Online]. Available: http://www.pps.univ-paris-diderot.fr/~jch/software/polipo/

[2] "Arch linux on arm." [Online]. Available: http://archlinuxarm.org/platforms/armv7/ti/beaglebone-black

[3] "Ext4 filesystem documentation." [Online]. Available: https://www.kernel.org/doc/Documentation/filesystems/ext4.txt

[4] Siege manual. [Online]. Available: http://www.joedog.org/siege-manual/

[5] "Yaffs filesystem." [Online]. Available: http://www.yaffs.net/

[6] B. Du, M. Demmer, and E. Brewer, "Analysis of www traffic in cambodia and ghana," in *Proceedings of the 15th International Conference on World Wide Web*, ser. WWW '06. New York, NY, USA: ACM, 2006, pp. 771–780. [Online]. Available: http://doi.acm.org/10.1145/1135777.1135894

[7] H.-J. Kim and J.-S. Kim, "Tuning the ext4 filesystem performance for android-based smartphones," in *Frontiers in Computer Education*, ser. Advances in Intelligent and Soft Computing, S. Sambath and E. Zhu, Eds. Springer Berlin Heidelberg, 2012, vol. 133, pp. 745–752. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-27552-4_98

[8] V. S. Pai, A. Badam, S. Ihm, and K. Park, "First-class access for developing-world environments," in *Proceedings of the 5th International Conference on Future Internet Technologies*, ser. CFI '10. New York, NY, USA: ACM, 2010, pp. 30–34. [Online]. Available: http://doi.acm.org/10.1145/1853079.1853089