PHP Optimization Using Hip Hop Virtual Machine

Chaitali Tambe, Pramod Pawar, Dashrath Mane
Vivekanand education society institute of technology, Chembur

Abstract— **HipHop compiler is PHP based compiler created by facebook in order to optimize the php code and save resources on the facebook server. HPHPc translates the php code into executable c++ code. Later facebook realized that HPHPc had some snags which included resource consumption and parallel maintenance of HPHP and HPHPi. In order to overcome the drawbacks of HPHPc HipHop virtual machine was developed in 2011 which is a process virtual machine based on just-in-time (JIT) compilation, serving as an execution engine for PHP and Hack programming languages. HHVM provides complete support for the entire PHP language. This paper shows how HHVM is superior to other compilers like PHP, PHP 7 and how various framework uses HHVM to optimize php. In addition the lock-down approach shows the behavior of each framework with hhvm and show how performance optimization can be obtained for heavy websites.**

Index Terms - **Hack, HHVM, HipHop, HPHPc, JIT, PHP.**

I. INTRODUCTION

During recent years php has emerged as prominent language for website development. over millions of websites created used php as their core language to develop web applications. As of January 2013, PHP was installed on more than 240 million websites and 2.1 million web servers originally created by Rasmus Lerdorf in 1994, the reference implementation of PHP is now produced by The PHP Group. PHP is used as development language for major organizations including the Facebook.

As Facebook's popularity increased, it was necessary for the website to be scalable HipHop for PHP (HPHPc) is a PHP transpiler created by Facebook. HipHop Compiler Converts PHP sourcecode into highly optimized C++ code, which includes built in multi-threaded webserver. HPHPc has its own configuration file syntax also support virtual hosts. The main objective behind HipHop was to optimize resources on Facebook servers. HipHop Compiler allowed resource optimization of facebook server during its increasing popularity.

HHVM converts executed PHP or Hack code into HHBC as a high-level bytecode. HHVM support for the entire PHP language and wide range of databases. HHVM reduces the conversion time.

II. HIPHOP COMPILER

HipHop for PHP (HPHPc) is a PHP transpiler created by Facebook. It is a source to source compiler which translates php code to c++ and stores it as a executable code as opposed to the PHP's usual execution path of PHP code being transformed into opcodes and interpreted. HPHPc consists mainly of C++, C and PHP source code, and HPHPc is open source and has free licensing.

The main reason behind HipHop was to save resources on Facebook servers, given the large PHP codebase of facebook.com. PHP applications that were developed using HipHop terminology resulted in increased speed of application. Several drawbacks of standalone php applications were eliminated by using HPHPc, one of which was performance.

A. Limitation of HPHPc:

- HPHPc did not fully support the PHP language, including the create_function() and eval() constructs.
- It involved a specific time and resource consuming deployment process that required a bigger than 1 GB binary to be compiled and distributed to many servers in short orders.
- In addition, maintaining HPHPc and HPHPi in parallel was becoming cumbersome.

III. HHVM

HipHop Virtual Machine (HHVM) is a process virtual machine based on just-in-time (JIT) compilation, serving as an execution engine for PHP and Hack programming languages. By using the principle of JIT compilation, executed PHP or Hack code is first transformed into intermediate HipHop bytecode (HHBC), which is then dynamically translated into the x86-64 machine code, optimized and natively executed [2].

A.What can HHVM do?

HHVM is mainly focused on server-side scripting, so you can collect form data, generate dynamic page content, or send and receive cookies. But HHVM can do much more.
There are two main areas where HHVM scripts are used.

- Server-side scripting:

  This is the most traditional and main target field for HHVM. You need three things to make this work. The HHVM parser, a web server and a web browser. You need to run the web server, with a connected HHVM installation. You can access the Hack program output with a web browser, viewing the Hack page through the server. All these can run on your home machine if you are just experimenting with Hack programming

- Command line scripting:

  You can make an HHVM script to run it without any server or browser. You only need the HHVM parser to use it this way. This type of usage is ideal for regularly executed scripts or testing new libraries and functionality. These scripts can also be used for simple text processing tasks.

HHVM is fully supported in various versions of the Linux operating system, including Ubuntu, CentOS, Debian, etc. HHVM can be installed on Mac OS X to run in interpreted mode (no JIT, but that is being actively developed). Windows support is actively being developed. HHVM has also support for most of the web servers today. This includes any web server that can utilize the FastCGI protocol, like Apache, lighttpd and nginx.

So with HHVM, you have the freedom of choosing an operating system and a web server. Furthermore, you also have the choice of using procedural programming or object oriented programming (OOP), or a mixture of them both. And, of course, you can choose to use Hack or PHP as your programming language.

With HHVM you are not limited to output HTML. HHVM's abilities include outputting images, PDF files and even Flash movies (using libswf and Ming) generated on the fly. You can also output easily any text, such as XHTML and any other XML file. HHVM can auto generate these files, and save them in the file system, instead of printing it out, forming a server-side cache for your dynamic content.

One of the strongest and most significant features in HHVM is its support for a wide range of databases. Writing a database-enabled web page is incredibly simple using one of the database specific extensions (e.g., for mysql), or using an abstraction layer like PDO, or connect to any database supporting the Open Database Connection standard via the ODBC extension. Other databases may utilize cURL or sockets.

HHVM also has support for talking to other services using protocols such as LDAP, IMAP, HTTP, and others. You can also open raw network sockets and interact using any other protocol. HHVM has support for the WDDX complex data exchange between virtually all Web programming languages.

HHVM has useful text processing features, which includes the Perl compatible regular expressions (PCRE), and many extensions and tools to parse and access XML documents. HHVM standardizes all of the XML extensions on the solid base of libxml2, and extends the feature set adding SimpleXML, XMLReader and XMLWriter support [5].

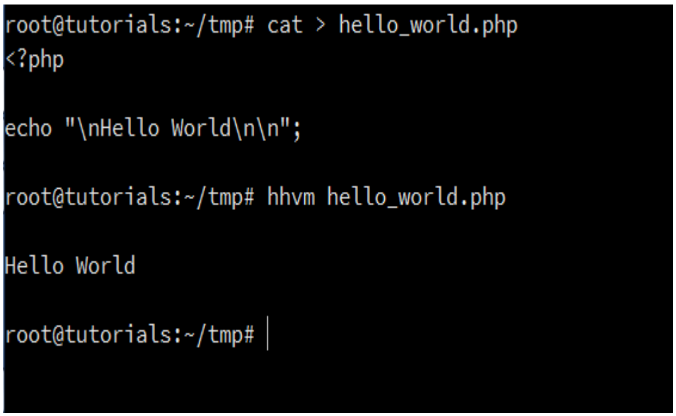
B.Installing HHVM:

Installing HHVM is quite straightforward and shouldn't take more than a few minutes. Executing the following 4 commands from the command line will have HHVM installed and ready [6]:

a. wget -O - http://dl.hhvm.com/conf/hhvm.gpg.key | apt-key add -
b. echo deb http://dl.hhvm.com/ubuntu saucy main | tee /etc/apt/sources.list.d/hhvm.list
c. apt-get update
d. apt-get install hhvm

To confirm that HHVM has been installed, type the following command:

hhvm –help



```
root@tutorials:~/tmp# cat > hello_world.php
<?php

echo "\nHello World\n\n";

root@tutorials:~/tmp# hhvm hello_world.php

Hello World

root@tutorials:~/tmp# |
```

Figure1. Running Hello World program [6].

IV. ADVANTAGES OF HHVM OVER HPHPC

A. Performance:

Performance of facebook.com code running on HHVM was even surpassed by around 15%.

B. The development environment:

HHVM reduced page load times by over 3x compared to HPHPi, all while keeping the rapid workflow that HPHPi provided. HHVM enable PHP developers to debug, tune, and iterate with the same system run in production.

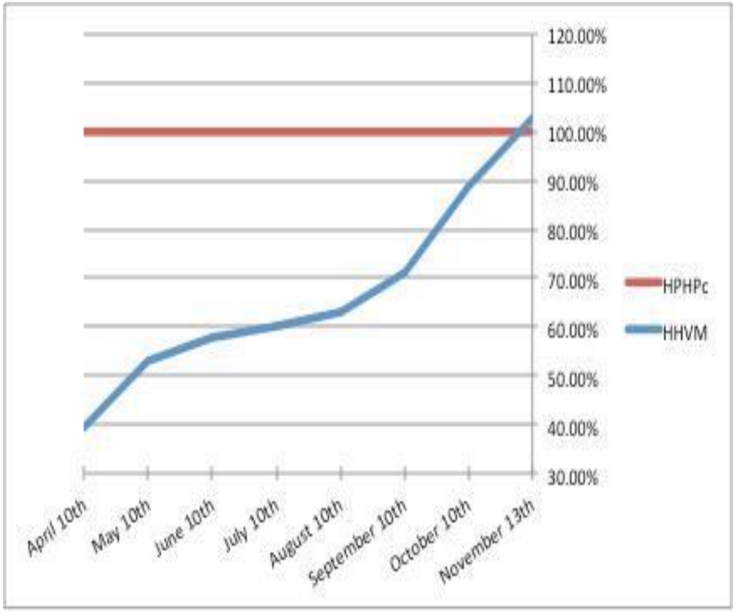


Figure2. HHVM Performance graph [3].

C. Open source:

HHVM is an Open Source and it is an important step forward in reducing friction for adoption.

D. Improving the JIT:

HHVM runs PHP programs by converting the PHP source into bytecode (HHBC) and executing the bytecode using both a bytecode interpreter and an x64 JIT compiler that seamlessly interoperate with each other. The JIT is used where possible, and the interpreter is used as an execution engine of last resort.

V. HHBC

HipHop ByteCode (HHBC) captured PHP's semantics correctly and happened to fit in fairly well with PHP's evaluation order.

A. Advantages:

- Run-time efficiency:

  The design of HHBC should be congruous to implementing an efficient execution engine, whether it be an interpreter or a just-in-time compiler.

- PHP 5.5 compatibility:

  It should be possible to compile valid PHP 5.5 source code into HipHop bytecode in a way that preserves the semantic meaning of the source.

- Simplicity:

  The design of HHBC should avoid features that could be removed or simplified without compromising PHP 5.5 compatibility, run-time efficiency,or design cleanliness.

VI. HACK

Hack is a programming language for HHVM. Hack reconciles the fast development cycle of a dynamically typed language with the discipline provided by static typing, while adding many features commonly found in other modern programming languages. Hack provides instantaneous type checking by incrementally checking your files as you edit them. It typically runs in less than 200 milliseconds, making it easy to integrate into your development workflow without introducing a noticeable delay [7].

A. Advantages:

- Strict Typing:

  PHP is a loosely typed language and type-related errors may not be recognized until runtime. Strict typing nullifies the need for a lot of type-related unit tests and encourages developers to catch type-related errors sooner in the development process.

- Instantaneous Type Checking:

  To make the development process and error-catching process even easier, type-checking server with its HHVM engine is included.This server runs locally and monitors Hack code as it is written. Developer’s code editors and IDEs can use this type-checking server to immediately report syntax or type-related errors during code development.

VII. FRAMEWORKS USED FOR PHP DEVELOPMENT

A. Drupal

Drupal is a free and open-source content-management framework written in PHP and distributed under the GNU General Public License. It is used as a back-end framework for at least 2.1% of all Web sites worldwide ranging from personal blogs to corporate, political, and government sites including WhiteHouse.gov and data.gov.uk. It is also used for knowledge management and business collaboration.

The standard release of Drupal, known as Drupal core, contains basic features common to content management systems. These include user account registration and maintenance, menu management, RSS feeds, taxonomy, page layout customization, and system administration. The Drupal core installation can serve as a simple Web site, a single- or multi-user blog, an Internet forum, or a community Web site providing for user-generated content.

Drupal core contains the essential building blocks providing you a flexible foundation to get started with Drupal [8].

B.MediaWiki

MediaWiki is a free and open-source wiki application. It was developed by the Wikimedia Foundation and runs on many websites, including Wikipedia, Wiktionary and Wikimedia Commons. It is written in the PHP programming language and uses a backend database.

The first version of the software was deployed to serve the needs of the Wikipedia encyclopedia in 2002.Wikipedia and other Wikimedia projects continue to define a large part of the requirement set for MediaWiki. The software is optimized to efficiently handle large projects, which can have terabytes of content and hundreds of thousands of hits per second.Because Wikipedia is one of the world's largest websites, achieving scalability through multiple layers of caching and database replication has been a major concern for developers [9].

C.Wordpress

WordPress is a free and open-source tool and a content management system (CMS) based on PHP and MySQL. Features include a plugin architecture and a template system. WordPress was used by more than 23.3% of the top 10 million websites as of January 2015.WordPress is the most popular blogging system in use on the Web, at more than 60 million websites.

It was first released on May 27, 2003, by its founders, Matt Mullenwegand Mike Little, as a fork of b2/cafelog. The license under which WordPress software is released is the GPLv2 (or later) from the Free Software Foundation [10].

VIII. ISSUE OF HPHPC

As the development of HipHop progressed, it was realised that HPHPc could not substantially increase the speed of PHP applications in general. Increases the code conversion time and resource consumption. A stated goal of HPHPc was to provide a high level of compatibility for Zend PHP, but most Zend-based PHP programs run unmodified on HPHPc.

As an addition to HPHPc, Facebook engineers also created a "developer mode" of HipHop (interpreted version of a PHP execution engine, known as HPHPi) and the HipHop debugger (known as HPHPd). These additions allow developers to run PHP code through the same logic provided by HPHPc while making it possible to interactively debug PHP code by defining watches, breakpoints, etc. Running the code through HPHPi yields lower performance when compared to HPHPc.

IX. PROPOSED SOLUTION(HHVM)

HHVM is developed by Facebook, with the project's source code hosted on GitHub and licensed under the terms of the PHP License and Zend License.

HHVM brings many benefits over HPHPc, one of them is almost complete support for the entire PHP language as defined by its official 5.4 version, including the support for create_function() and eval() constructs.

HHVM there is only one execution engine for both production and development environments, meaning that maintaining HPHPi separately is no longer required, and there exists both production and development integration with HPHPd.

At the same time, development and deployment processes are much simpler when compared to HPHPc, as the lengthy binary build time is no longer present.

X. COMPARISON OF HHVM COMPILER OVER OTHER COMPILERS

The results published by HHVM team which demonstrated pre and post-lockdown HHVM which used PHP 5.6.9, and commits from PHP, and HHVM master. For the pre and post-lockdown comparison stable releases of MediaWiki, Drupal 7, and WordPress were used, post-lockdown numbers include a patch to MediaWiki that was written during the lockdown. The second set of comparisons uses the same stable releases, with the MediaWiki patch applied for all engines. Data for each engine, framework pair was collected in ten independent runs, and RPS was quantified. These results obtained showed low variability between runs and have a great deal of confidence in the reproducibility.

In the lockdown, the performance of MediaWiki improved by 19.4%, and WordPress by 1.8%. Unfortunately, Drupal wins were no longer measurable once the Drupal database was patched to fix the aforementioned plug-in bug. In add ition, we increased RPS for simple pages by 5.2% (this was mostly a measure of the overhead incurred by a request). The lockdown offered an opportunity to evaluate methodology, and work with the community on realistic configurations for the sample data used in the benchmarking created during the demonstration. Number important changes to the benchmarking tools and framework configurations were made throughout the course of the lockdown. The results of the lockdown are summarized below, normalized to pre lockdown RPS numbers.

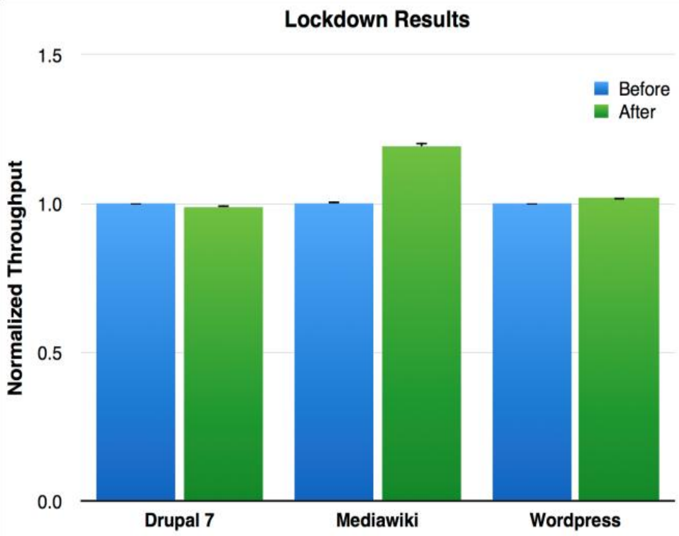


Figure3. Lockdown Results [11].

Across the board it was found that HHVM performs best on applications where CPU time is maximized. In particular HHVM was 1.55 times faster than PHP 7 on a MediaWiki workload, 1.1 times faster on a Drupal 7 workload, and 1.19 times faster on a WordPress workload. As none of these frameworks take advantage of the asynchronous I/O architecture available in HHVM (i.e., async), it’s not surprising that the greatest performance benefits come from the efficient execution of PHP code possible with a JIT compiler. The following figure summarizes the performance difference between PHP 5, PHP 7, and HHVM. Results were normalized to PHP 5 RPS, and Drupal 8 has been included. We benchmarked Drupal 8 with caching both enabled and disabled. In general the results for Drupal 8 were more stable with the cache disabled.

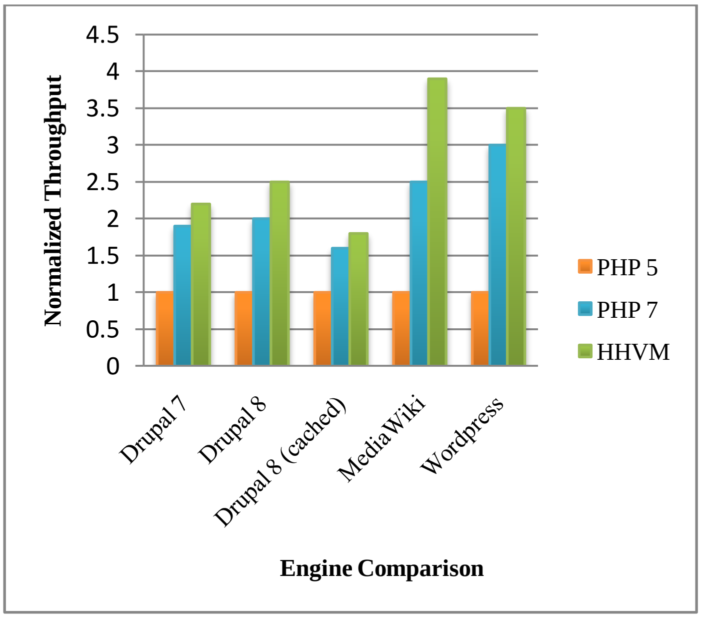


Figure4. Engine Comparison [11].

For all benchmarks ten independent runs were performed and used the mean RPS result. Standard deviation was also measured and was highly consistent between runs, and it was seen that with the proper configuration and hardware optimized result are easily produced.

During lockdown a 19.4% RPS improvement for MediaWiki workloads was achieved, and a 1.8% RPS improvement for WordPress. It was demonstrated that HHVM is 55.5% faster than PHP 7 on a MediaWiki workload, 18.7% faster on a WordPress workload, and 10.2% faster on a Drupal 7 workload. Improvements made to HHVM to better serve open source frameworks will ship with the next release [11].

XI. CONCLUSION

This research analyzes the HPHPc issue and suggests a solution for using HHVM. The issues are related to the HPHPc not fully support the PHP language and resource consuming deployment process. In addition, maintaining HPHPc and HPHPi in parallel was becoming cumbersome.

HHVM overcome from this issue, HHVM reduced page load times and drastically reduce the number of servers. HHVM's development and deployment processes are much simpler than HPHPc. HHVM is open source and improved JIT performance.

REFERENCES

[1] http://en.wikipedia.org/wiki/HipHop_for_PHP
[2] http://en.wikipedia.org/wiki/HipHop_Virtual_Machine
[3] https://www.facebook.com/notes/facebook-engineering/speeding-up-php-based-development-with-hiphop-vm/10151170460698920
[4] https://github.com/facebook/hhvm/wiki/Running-PHP-programs-with-HHVM
[5] http://docs.hhvm.com/manual/en/install-intro.intro.hhvm.php
[6] https://www.digitalocean.com/community/tutorials/how-to-install-hhvm-hiphop-virtual-machine-on-an-ubuntu-13-10-vps
[7] http://hacklang.org/
[8] https://en.wikipedia.org/wiki/Drupal
[9] https://en.wikipedia.org/wiki/MediaWiki
[10] https://en.wikipedia.org/wiki/WordPress
[11] http://hhvm.com/blog/9293/lockdown-results-and-hhvm-performance

ISSN: 2278 – 1323