

BUILDING LEAN, STANDALONE WEB SERVERS, AND ROUTING ENGINES

PROJECT REPORT

Presented to

Dr. Chris Pollett

Department of Computer Science

San Jose State University

In Partial Fulfillment

Of the Requirements for the Class

CS 297

By

Ajita Shrivastava

Fall 2023

ABSTRACT

Scalability, optimization, and rapid development are among the primary concerns in web application development. This project introduces a collection of single-file, low-dependency, pure PHP servers and routing engines that have adopted an event-driven, asynchronous I/O approach. Currently comprising two server classes – WebSite and GopherSite – the project can function seamlessly within traditional servers like Apache, nginx, or lighttpd, or operate independently as standalone servers for applications utilizing its routing capabilities. The server is implemented to be request event-driven, supporting asynchronous I/O for efficient web traffic handling. In response to the increasing complexity of email setups and the reliance on external software, this research endeavors to develop single-file, low-dependency, pure PHP servers and routing engines implementing SMTP and IMAP protocols. The aim is to simplify email server configurations, minimize external software dependencies, and democratize the process of setting up email servers. Creating lightweight servers using PHP ensures ease of integration into diverse projects while maintaining robustness. Thus, redefining the landscape of server technology, providing a user-friendly and efficient solution to email server management.

TABLE OF CONTENTS

I. Introduction	4
II. Deliverable I: Implementation of routes with atto servers.....	6
III. Deliverable II: Conversion of HTTP 1.1 to HTTP/2	9
IV. Deliverable III: Additional feature implementation: HELP Command	9
V. Deliverable IV: HTTP/2 Detection.....	13
VI. Conclusion.....	15
References.....	17

I. INTRODUCTION

In the current global landscape, where web application development stands as a cornerstone of digital interaction, the imperatives of scalability, optimization, and swift development have become paramount. Imagine the internet as a vast network of highways, and web applications as the cars navigating through. This project introduces another kind of road system - think of it as efficient pluggable highways. These highways, i.e. single file PHP servers, are designed to manage web traffic smoothly and quickly. They're like magical paths that help websites by providing a small yet efficient PHP server. But that's not all – this project also tackles another challenge: emails. Setting up email servers can be like solving a puzzle with too many pieces. The project simplifies this puzzle, making it easier for everyone to set up and manage their email servers. It's like giving you a user-friendly tool to handle your emails without needing a tech expert. Ultimately, this project serves as a versatile micro framework, ensuring a streamlined experience for developers to use under traditional servers like Apache, nginx, or lighttpd.

This report consists of four sections, each detailing one of the completed deliverables from this semester. Following these sections are the conclusion and future work. Each section elaborates on the deliverable's aim, the solution approach, and the results achieved. Like any research project, the first deliverable serves as an exercise to comprehend the existing technology and gain familiarity with the various aspects of this project. It entails creating routes using the web server to understand the structure and functioning of the server. The second deliverable provides a comprehensive understanding of the project's next phase: the transition from HTTP

1.1 to HTTP/2. Here, a simple program is developed to illustrate the conversion of a given HTTP 1.1 request into an HTTP/2 binary request format. The third deliverable introduces a new feature to the existing platform—middleware priority and ordering. Implemented as modular functions within the web server class, this addition enhances the project's capabilities to prioritize and order the execution of middleware functions. The fourth deliverable delves into extraction of the negotiated protocol. The implementation involves attempting to connect to a server to determine the protocol negotiated between the client and server for further communication.

In essence, this project and its deliverables establish a foundation of knowledge on the subject, provide an understanding of the features, and foster comfort of working on the project. This inherently charts the course for future project implementation. Finally, the section on future work outlines a direction for the future implementation of the project.

II. DELIVERABLE 1: IMPLEMENTATION OF ROUTING WITH ATTO SERVER

The primary goal of this deliverable is to create a seamless system for directing web traffic within the Atto Server. By configuring routes in the `index.php` file, we aim to enable the server to display different web pages, providing a practical understanding of its operational dynamics.

In setting up this routing system, we began by crafting the `index.php` file. This file acts as the central control for defining how incoming requests should be handled and directed to specific pages. The code written in this file establishes the groundwork for effective routing.

Subsequently, running the server was a crucial step. By navigating to the folder containing the `index.php` file and executing the command `php index.php`, we ensured that the server was up and running, actively listening on port 8000.

To observe the routing in action, users simply needed to open their web browser and visit `http://localhost:8000`. This direct interaction with the main page provided a tangible demonstration of how the server manages different routes.

The successful implementation of these measures yielded a fully functional routing system within the Atto Server. The main page can be accessed through `http://localhost:8000`, and from there, navigate to various web pages, each associated with a distinct route as shown in the below screenshots.

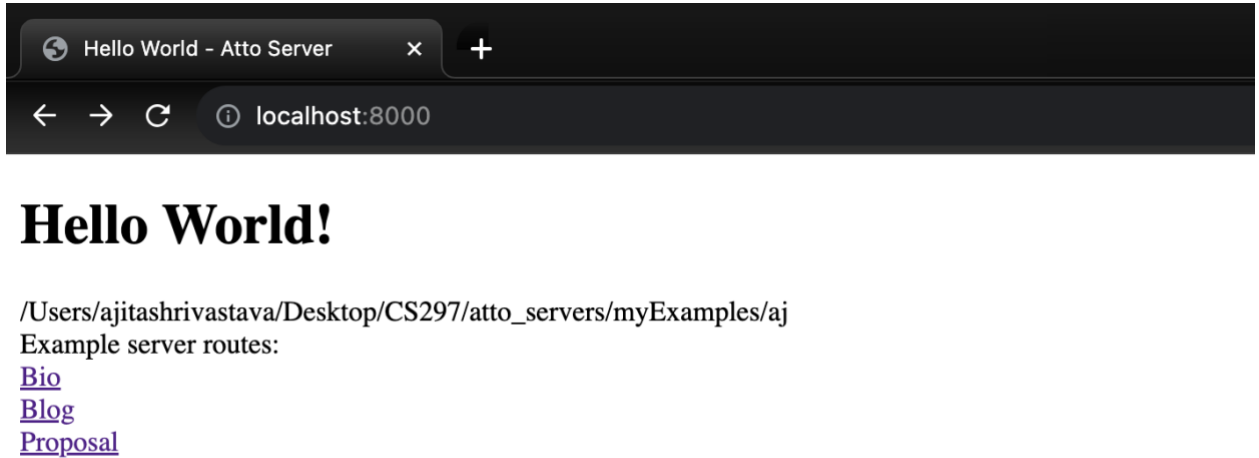


Figure 1: Results of the main page displaying the routes

Bio Page Route



Figure 2: Results of the bio link route

Blog Page Route

```
$test->get('/blog', function() { ... });
```

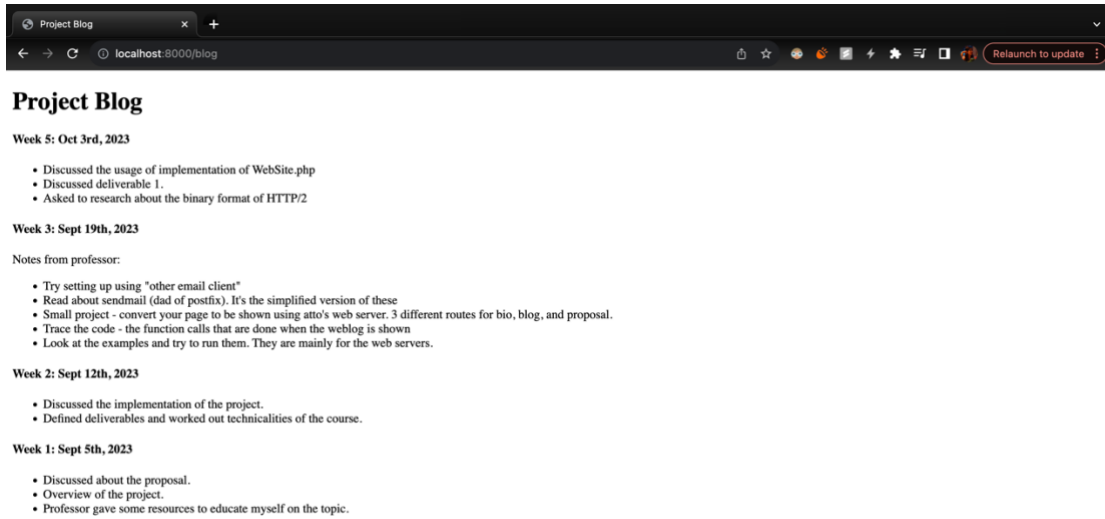


Figure 3: Results of the project link route

Proposal Page Route

```
$test->get('/proposal', function() { ... });
```

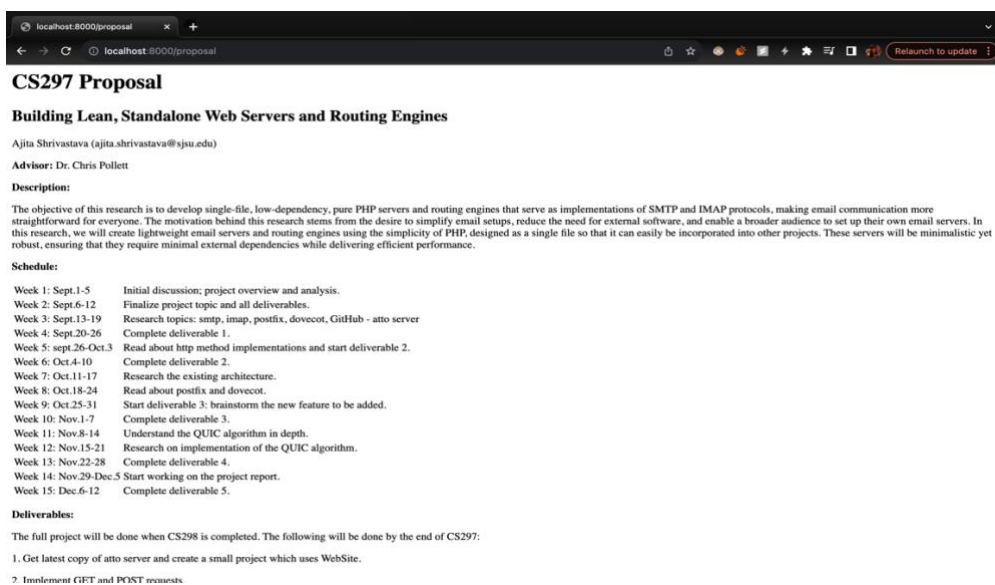


Figure 4: Results of the proposal link route

III. DELIVERABLE 2: CONVERSION OF HTTP1.1 TO HTTP/2

This deliverable involves the development of a program to convert an HTTP 1.1 request into the binary format of HTTP/2, offering insights into the conversion process. The program comprises three primary steps.

In the first step, the format is transformed by adding necessary headers and assigning their values. This involves extracting key components of the request, such as 'method', 'path', 'protocol', 'headers', and 'body'. Two functions contribute to this process:

`http1_request_parse($request_string)` parses the request to create an array, and

`http1_to_http2_hex($request_array)` replaces extracted parts with the new format.

Following the new format, a request in HTTP/2 includes headers with binary values assigned to them. The `createHeadersFrame($request_array, $lastFrame)` function generates a frame for these headers. For example, the conversion of frame types is tabulated, with each type assigned a specific hexadecimal value.

The subsequent function, `createDataFrame($request_array)`, creates a frame for the data stored in the original request by converting it into key-value pairs. This ensures the appropriate fitting into the HTTP/2 format.

Finally, each frame is converted into binary, simplifying the process and facilitating further analysis. This comprehensive conversion process provides a clear understanding of how an HTTP 1.1 request evolves into the binary format of HTTP/2.

```

ajitashrivastava@Ajitas-MacBook-Pro deliverable2 % php parser.php
Xdebug: [Step Debug] Could not connect to debugging client. Tried: localhost:9003 (through xdebug.client_host/xdebug.client_port).
Example usage 1:
*****
Original request:
GET /doc/test.html HTTP/1.1
Host: www.test101.com
Accept: image/gif, image/jpeg, */*
Content-Length: 35

bookId=1234&author=paulo+Coelho

Parsed request array:
Array
(
    [method] => GET
    [path] => /doc/test.html
    [protocol] => HTTP/1.1
    [headers] => Array
        (
            [Host] => www.test101.com
            [Accept] => image/gif, image/jpeg, */*
            [Content-Length] => 35
        )
    [body] => bookId=1234&author=paulo+Coelho
)

Frames created:
Array
(
    [headerFrame] => Array
        (
            [length] => 136
            [type] => 0xf
            [flags] => Array
                (
                    [end_stream_flag] => 0
                    [end_header_flag] => 1
                    [padding_flag] => 0
                    [prior_flag] => 0
                )

            [P] => 0
            [streamIO] => 18943
            [paddingLength] => not set
            [E] => not set
            [StreamDependency] => not set
            [weight] => not set
            [requestHeaders] => :method = GET
:scheme = HTTP/1.1
:path = /doc/test.html
:Host = www.test101.com
:Accept = image/gif, image/jpeg, */*
:Content-Length = 35

            [padding] => not set
        )
)

```


IV. DELIVERABLE 3: ADDITIONAL FEATURE IMPLEMENTATION

In this deliverable, the implementation of the HELP command feature in MailSite.php is achieved. This feature enables the server to provide assistance to clients by offering detailed information about available commands. The HELP command does not affect the core functionalities of MailSite, such as the reverse-path buffer, forward-path buffer, or the mail data buffer. It can be invoked by clients at any point during communication. The HELP command's response follows a structured format, starting with a header indicating the beginning of available commands. It then lists each command along with its corresponding description, concluding with an end marker. The implementation involves several steps:

The first step involves creating a JSON file named 'cmds.json' that contains a list of all commands and their descriptions. This file format is chosen for its suitability in storing key-value pair information. Next, the 'parseHelp' function is implemented. This function reads the 'cmds.json' file, decodes its contents, and formats them into a printable string. It handles any file-related errors that may occur during the process. The 'parseHelp' function is then integrated into the 'parseRequest' function, which is responsible for parsing client requests. The response generated by 'parseHelp' is added to the 'out_stream' for transmission to the client. Lastly, the 'processRequestStreams' function is updated to accommodate the HELP command. Proper checks are implemented to handle the new command, and the server's response, including the HELP information, is included in the 'out_stream' and returned to the client.

Overall, the addition of the HELP command feature enhances the usability and user experience of MailSite, providing clients with valuable assistance and improving communication between the client and server. Below are some illustrations of the commands json file and the results.

```
ajitashrivastava@Ajitash-MacBook-Pro ~ % telnet localhost 25
Trying 127.0.0.1...
Connected to localhost.
Escape character is '^]'.
220 127.0.0.1 ATTO MAIL SERVER
help
214- Available commands:
[AUTH]
    Used for authentication during the SMTP session. It is typically used when the server requires the client to authenticate before sending emails.

[EHLO]
    Initiates the SMTP session with the server and requests extended SMTP capabilities. It is used to identify the client to the server and to negotiate the supported features.

[HELO]
    Initiates the SMTP session with the server. It is a simpler version of EHLO and is used when extended capabilities are not required.

[NOOP]
    Does nothing but can be used as a keep-alive mechanism during the SMTP session to prevent the connection from timing out.

[QUIT]
    Ends the SMTP session with the server and closes the connection. It is used to gracefully terminate the session.

[RSET]
    Resets the current SMTP session state. It is used to abort the current email transaction and return to the initial state.

[STARTTLS]
    Initiates a TLS-encrypted SMTP session. It is used to secure the connection between the client and server by encrypting the data transmission.

[MAIL]
    Specifies the sender's email address in the email transaction. It is used to indicate who is sending the email.

[RCPT]
    Specifies the recipient's email address in the email transaction. It is used to indicate who is receiving the email.

[DATA]
    Indicates the start of the email message body. It is used to transmit the content of the email message.

[VRFY]
    Verifies if a given email address is valid and exists on the server. It can be used to check the validity of email addresses before sending emails.
```

Figure 6: Results of the HELP command invoked

V. DELIVERABLE 4: HTTP/2 DETECTION

This deliverable entails the development of a PHP program aimed at establishing a connection with a server and retrieving the negotiated protocol between the client and server. The ‘negotiated protocol extraction’ plays a vital role in the project as it aims to eventually support all versions of HTTP. This part will help determine the appropriate responses accordingly.

The initial step involves the creation of a TCP/IP socket using PHP's socket functions, enabling a channel for communication. The program tries to establish a connection to the designated server through the utilization of the `socket_connect()` function. This connection attempt either successfully establishes communication channels or prompts an error message in the event of connection failure.

Upon successfully establishing a connection, the program proceeds to extract the negotiated protocol from the handshake packets exchanged during the initial communication between the client and server. This extraction process entails dispatching a request to the server and meticulously capturing the protocol information encapsulated within the handshake packets. Through methodical dissection and analysis of the handshake protocol, the program discerns the mutually agreed-upon protocol designated for subsequent interactions between the client and server.

By systematically executing these procedures, the PHP program achieves its fundamental objective of connecting to the server and retrieving the negotiated protocol. This accomplishment serves as a foundation for facilitating seamless and efficient communication channels between

the client and server entities, thereby fostering enhanced interoperability and data exchange capabilities across network environments.

```
ajitashrivastava@Ajitash-MacBook-Pro src % php DetectingH2.php
Xdebug: [Step Debug] Could not connect to debugging client. Tried: localhost:9003 (through xdebug.client_host/xdebug.client_port).
Creating a socket...
Socket created...
Connecting to the socket...
Connected...
Fetching the negotiated protocol...

Protocol in use : HTTP/1.1
ajitashrivastava@Ajitash-MacBook-Pro src %
```

Figure 7: Results of the detection program

VI. CONCLUSION

In conclusion, the progression of this project has been instrumental in advancing the realm of web application development and email server management. The development of single-file, low-dependency, pure PHP servers and routing engines, exemplified by WebSite and GopherSite, has ushered in a new era of scalability, optimization, and rapid development in web environments. By embracing an event-driven, asynchronous I/O approach, these servers offer developers unparalleled flexibility in managing web traffic within both traditional server setups and standalone applications.

Additionally, the project's exploration into email server management represents a significant leap forward in simplifying the complexities associated with setting up and managing email servers. Through the implementation of SMTP and IMAP protocols within single-file PHP servers, the project aims to streamline email server configurations, reduce external software dependencies, and democratize the process of email server management.

Each deliverable achieved throughout the project's lifecycle has played a crucial role in propelling its objectives forward. From laying the groundwork with atto servers to navigating the intricacies of detecting HTTP/2 and negotiating protocols, each task has contributed to a deeper understanding of server technology and its practical implications for web development.

As the project looks towards the future, the insights gleaned from these milestones provide invaluable guidance for continued development and innovation. With a steadfast focus on meeting the evolving needs of developers and users alike, this project stands poised to shape the future of web application development and email server management.

REFERENCES

- [1] Riabov, Vladimir V. "SMTP (Simple Mail Transfer Protocol)." River College, 2005.
- [2] Mullet, Dianna, and Kevin Mullet. *Managing Imap*. O'Reilly Media, Inc., 2000.
- [3] Gourley, David, and Brian Totty. *HTTP: The Definitive Guide*. O'Reilly Media, Inc., 2002.
- [4] Pollard, Barry. *HTTP/2 in Action*. Simon and Schuster, 2019.
- [5] Belshe, Mike, Roberto Peon, and Martin Thomson. "Hypertext Transfer Protocol Version 2 (HTTP/2)." No. rfc7540. 2015. RFC Manual.
- [6] Langley, Adam, et al. "The QUIC Transport Protocol: Design and Internet-Scale Deployment." *Proceedings of the Conference of the ACM Special Interest Group on Data Communication*, 2017.
- [7] Carlucci, Gaetano, Luca De Cicco, and Saverio Mascolo. "HTTP over UDP: An Experimental Investigation of QUIC." *Proceedings of the 30th Annual ACM Symposium on Applied Computing*, 2015.
- [8] Dent, Kyle D. *Postfix: The Definitive Guide: A Secure and Easy-to-Use MTA for UNIX*. O'Reilly Media, Inc., 2003. *Postfix Overview*.
- [9] Hildebrandt, Ralf, and Patrick Koetter. *The Book of Postfix: State-of-the-Art Message Transport*. No Starch Press, 2005.