

Yaws - Yet Another Web Server

Claes Wikstrom
klacke@hyber.org

June 25, 2012

Contents

1	Introduction	3
1.1	Prerequisites	4
1.2	A tiny example	4
2	Compile, Install, Config and Run	6
2.0.1	Compile and Install	6
2.0.2	Configure	7
3	Static content	10
4	Dynamic content	11
4.1	Introduction	11
4.2	EHTML	11
4.3	POSTs	16
4.3.1	Queries	16
4.3.2	Forms	16
4.4	POSTing files	17
5	Mode of operation	20
5.1	On-the-fly compilation	20
5.2	Evaluating the Yaws Code	21
6	SSL	22
7	Applications	23
7.1	Login scenarios	23
7.1.1	The session server	23
7.1.2	Arg rewrite	24

CONTENTS	2
7.1.3 Authenticating	26
7.1.4 Database driven applications	27
7.2 Appmods	28
7.3 The opaque data	28
7.4 Customizations	29
7.4.1 404 File not found	29
7.4.2 Crash messages	30
7.5 Stream content	30
7.6 All out/1 return values	31
8 Debugging and Development	34
8.1 Logs	34
9 External scripts via CGI	35
10 FastCGI	36
10.1 The FastCGI Responder Role	36
10.2 The FastCGI Authorizer Role	37
10.3 The FastCGI Filter Role	37
10.4 FastCGI Configuration	37
11 Security	38
11.1 WWW-Authenticate	38
12 Embedded mode	40
12.1 Creating Global and Server Configurations	40
12.2 Starting Yaws in Embedded Mode	41
13 The config file - yaws.conf	42
13.1 Global Part	42
13.2 Server Part	45
13.3 Configuration Examples	54
14 WebSocket Protocol Support	57
14.1 WebSocket Callback Modules	57
14.1.1 Basic Callback Modules	57
14.1.2 Advanced Callback Modules	58

Chapter 1

Introduction



YAWS is an ERLANG web server. It's written in ERLANG and it uses ERLANG as its embedded language similar to PHP in Apache or Java in Tomcat.

The advantages of ERLANG as an embedded web page language as opposed to Java or PHP are many.

- Speed - Using ERLANG for both implementing the web server itself as well as embedded script language gives excellent dynamic page generation performance.
- Beauty - Well this is subjective
- Scalability - due to the lightweight processes of ERLANG, YAWS is able to handle a very large number of concurrent connections

YAWS has a wide feature set; it supports:

- HTTP 1.0 and HTTP 1.1
- Static content page delivery
- Dynamic content generation using embedded ERLANG code in the HTML pages
- Common Log Format traffic logs
- Virtual hosting with several servers on the same IP address
- Multiple servers on multiple IP addresses
- HTTP tracing for debugging
- An interactive interpreter environment in the Web server for use while developing and debugging a web site

- RAM caching of commonly accessed pages
- Full streaming capabilities of both upload and download of dynamically generated pages
- SSL
- Support for WWW-Authenticated pages
- Support API for cookie based sessions
- Application Modules where virtual directory hierarchies can be made
- Embedded mode
- WebSockets (RFC 6455)
- Long polling (COMET) applications
- Forward and reverse proxying

1.1 Prerequisites

This document requires that the reader:

- Is well acquainted with the ERLANG programming language.
- Understands basic Web technologies.

1.2 A tiny example

We introduce YAWS by help of a tiny example. The web server YAWS serves and delivers static content pages similar to any old web server, except that YAWS does this much faster than most web servers. It's the dynamic pages that makes YAWS interesting. Any page with the suffix “.yaws” is considered a dynamic YAWS page. A YAWS page can contain embedded ERLANG snippets that are executed while the page is being delivered to the WWW browser.

Example 1.1 is the HTML code for a small YAWS page.

It illustrates the basic idea behind YAWS. The HTML code, generally stored in a file ending with a “.yaws” suffix, can contain `<erl>` and `</erl>` tags and inside these tags an ERLANG function called `out/1` gets called and the output of that function is inserted into the HTML document, dynamically.

It is possible to have several chunks of HTML code together with several chunks of ERLANG code in the same YAWS page.

The `Arg` argument supplied to the automatically invoked `out/1` function is an ERLANG record that contains various data which is interesting when generating dynamic pages. For example the HTTP headers which were sent from the WWW client, the actual TCP/IP socket leading to the WWW client. This will be elaborated on thoroughly in later chapters.

The `out/1` function returned the tuple `{html, String}` and `String` gets inserted into the HTML output. There are number of different return values that can be returned from the `out/1` function in order to control the behavior and output from the YAWS web server.

```
<html>

<p> First paragraph

<erl>
out(Arg) ->
    {html, "<p>This string gets inserted into HTML document dynamically"}.
</erl>

<p> And here is some more HTML code

</html>
```

Figure 1.1: Example 1.1

Chapter 2

Compile, Install, Config and Run

This chapter is more of a “Getting started” guide than a full description of the YAWS configuration. YAWS is hosted on Github at <https://github.com/klacke/yaws>. This is where the source code resides in a git repository and the latest unreleased version is available via git through the following commands:

```
$ git clone https://github.com/klacke/yaws
```

Released version of YAWS are available at <http://yaws.hyber.org/download>.

2.0.1 Compile and Install

To compile and install a YAWS release one of the prerequisites is a properly installed ERLANG system. YAWS runs on ERLANG releases OTP R8 and later. Get ERLANG from <http://www.erlang.org/>

Compile and install is straight forward:

```
# cd /usr/local/src
# tar xzf yaws-X.XX.tar.gz
# cd yaws
# ./configure && make
# make install
```

The make command will compile the YAWS web server with the `erlc` compiler found by the configure script.

- `make install` - will install the executable called `yaws` in `/usr/local/bin/` and a working configuration file in `/etc/yaws.conf`
- `make local_install` - will install the executable in `$HOME/bin` and a working configuration file in `$HOME/yaws.conf`

While developing a YAWS site, it's typically most convenient to use the `local_install` and run YAWS as a non-privileged user.

2.0.2 Configure

Let's take a look at the config file that gets written to \$HOME after a local_install.

```
# first we have a set of globals

logdir = .
ebin_dir = /home/klacke/yaws/yaws/examples/ebin
include_dir = /home/klacke/yaws/yaws/examples/include

# and then a set of servers

<server localhost>
    port = 8000
    listen = 127.0.0.1
    docroot = /home/klacke/yaws/yaws/scripts/./www
</server>
```

Figure 2.1: Minimal Local Configuration

The configuration consists of an initial set of global variables that are valid for all defined servers.

The only global directive we need to care about for now is the logdir. YAWS produces a number of log files and they will—using the Configuration from Figure 2.1—end up in the current working directory. We start YAWS interactively as

```
# ~/bin/yaws -i
Erlang (BEAM) emulator version 5.1.2.b2 [source]

Eshell V5.1.2.b2 (abort with ^G)
1>
=INFO REPORT==== 30-Oct-2002::01:38:22 ===
Using config file /home/klacke/yaws.conf
=INFO REPORT==== 30-Oct-2002::01:38:22 ===
Listening to 127.0.0.1:8000 for servers ["localhost:8000"]

1>
```

By starting YAWS in interactive mode (using the command switch *-i*) we get a regular ERLANG prompt. This is most convenient when developing YAWS pages. For example we:

- Can dynamically compile and load optional helper modules we need.
- Get all the crash and error reports written directly to the terminal.

The configuration in Example 2.1 defined one HTTP server on address 127.0.0.1:8000 called "localhost". It is important to understand the difference between the name and the address of a server. The name is the expected value in the client HTTP `Host` : header. That is typically the same as the fully-qualified DNS name of the server whereas the address is the actual IP address of the server.

Since YAWS support virtual hosting with several servers on the same IP address, this matters.

Nevertheless, our server listens to *127.0.0.1:8000* and has the name "localhost", thus the correct URL for this server is `http://localhost:8000`.

The document root (docroot) for the server is set to the `www` directory in the YAWS source code distribution. This directory contains a bunch of examples and we should be able to run all those example now on the URL `http://localhost:8000`.

Instead of editing and adding files in the YAWS `www` directory, we create yet another server on the same IP address but a different port number — and in particular a different document root where we can add our own files.

```
# mkdir ~/test
# mkdir ~/test/logs
```

Now change the config so it looks like this:

```
logdir = /home/klacke/test/logs
ebin_dir = /home/klacke/test
include_dir = /home/klacke/test

<server localhost>
  port = 8000
  listen = 127.0.0.1
  docroot = /home/klacke/yaws/yaws/www
</server>

<server localhost>
  port = 8001
  listen = 127.0.0.1
  docroot = /home/klacke/test
</server>
```

We define two servers, one being the original default and a new pointing to a document root in our home directory.

We can now start to add static content in the form of HTML pages, dynamic content in the form of `.yaws` pages or ERLANG `.beam` code that can be used to generate the dynamic content.

The load path will be set so that beam code in the directory `~/test` will be automatically loaded when referenced.

It is best to run YAWS interactively while developing the site. In order to start the YAWS as a daemon, we give the flags:

```
# yaws -D --heart
```

The *-D* or *-daemon* flags instructs YAWS to run as a daemon and the *-heart* flag will start a heartbeat program called heart which restarts the daemon if it should crash or if it stops responding to a regular heartbeat. By default, heart will restart the daemon unless it has already restarted 5 times in 60 seconds or less, in which case it considers the situation fatal and refuses to restart the daemon again. The *-heart-restart=C,T* flag changes the default 5 restarts in 60 seconds to *C* restarts in *T* seconds. For infinite restarts, set both *C* and *T* to 0. This flag also enables the *-heart* flag.

Once started in daemon mode, we have very limited ways of interacting with the daemon. It is possible to query the daemon using:

```
# yaws -S
```

This command produces a simple printout of uptime and number of hits for each configured server.

If we change the configuration, we can HUP the daemon using the command:

```
# yaws -h
```

This will force the daemon to reread the configuration file.

Chapter 3

Static content

YAWS acts very much like any regular web server while delivering static pages. By default YAWS will cache static content in RAM. The caching behavior is controlled by a number of global configuration directives. Since the RAM caching occupies memory, it may be interesting to tweak the default values for the caching directives or even to turn it off completely.

The following configuration directives control the caching behavior

- *max_num_cached_files = Integer*

YAWS will cache small files such as commonly accessed GIF images in RAM. This directive sets a maximum number on the number of cached files. The default value is 400.

- *max_num_cached_bytes = Integer*

This directive controls the total amount of RAM which can maximally be used for cached RAM files. The default value is 1000000, 1 megabyte.

- *max_size_cached_file = Integer*

This directive sets a maximum size on the files that are RAM cached by YAWS. The default value is 8000 bytes, 8 batters.

It may be considered to be confusing, but the numbers specified in the above mentioned cache directives are local to each server. Thus if we have specified `max_num_cached_bytes = 1000000` and have defined 3 servers, we may actually use $3 * 1000000$ bytes.

Chapter 4

Dynamic content

Dynamic content is what YAWS is all about. Most web servers are designed with HTTP and static content in mind whereas YAWS is designed for dynamic pages from the start. Most large sites on the Web today make heavy use of dynamic pages.

4.1 Introduction

When the client GETs a page that has a “.yaws” suffix, the YAWS server will read that page from the hard disk and divide it in parts that consist of HTML code and ERLANG code. Each chunk of ERLANG code will be compiled into a module. The chunk of ERLANG code must contain a function `out/1`. If it doesn't the YAWS server will insert a proper error message into the generated HTML output.

When the YAWS server ships a .yaws page it will process it chunk by chunk through the .yaws file. If it is HTML code, the server will ship that as is, whereas if it is ERLANG code, the YAWS server will invoke the `out/1` function in that code and insert the output of that `out/1` function into the stream of HTML that is being shipped to the client.

YAWS will (of course) cache the result of the compilation and the next time a client requests the same .yaws page YAWS will be able to invoke the already-compiled modules directly.

4.2 EHTML

There are two ways to make the `out/1` function generate HTML output. The first and most easy to understand is by returning a tuple `{html, String}` where `String` then is regular HTML data (possibly as a deep list of strings and/or binaries) which will simply be inserted into the output stream. An example:

```
<html>
<h1> Example 1 </h1>

<erl>
out(A) ->
    Headers = A#arg.headers,
    {html, io_lib:format("You say that you're running ~p",
```

```

[Headers#headers.user_agent))}.

</erl>

</html>

```

The second way to generate output is by returning a tuple `{ehtml, EHTML}`. The term `EHTML` must adhere to the following structure:

$$EHTML = [EHTML]|\{TAG, Attrs, Body\}|\{TAG, Attrs\}|\{TAG\}|$$

$$\{Module, Fun, [Args]\}|fun/0|$$

$$binary()|character()$$

$TAG = atom()$

$Attrs = [\{HtmlAttribute, Value\}]$

$HtmlAttribute = atom()$

$Value = string()|atom()|\{Module, Fun, [Args]\}|fun/0$

$Body = EHTML$

We give an example to show what we mean. The tuple

```

{ehtml, {table, [{bgcolor, grey}],
  [
    {tr, [],
      [
        {td, [], "1"},
        {td, [], "2"},
        {td, [], "3"}
      ]
    },
    {tr, [],
      [{td, [{colspan, "3"}], "444"}]}}}

```

expands into the following HTML code:

```

<table bgcolor="grey">
  <tr>
    <td> 1 </td>
    <td> 2 </td>
    <td> 3 </td>
  </tr>
  <tr>
    <td colspan="3"> 444 </td>
  </tr>
</table>

```

At a first glance it may appear as if the HTML code is more beautiful than the ERLANG tuple. That may very well be the case from a purely aesthetic point of view. However the ERLANG code has the advantage of being perfectly indented by editors that have syntax support for ERLANG (read Emacs). Furthermore, the ERLANG code is easier to manipulate from an ERLANG program.

Note that ehtml supports function calls as values. Functions can return any legal ehtml value, including other function values. YAWS supports `{M,F,[Args]}` and `fun/0` function value forms.

As an example of some more interesting ehtml we could have an `out/1` function that prints some of the HTTP headers. In the `www` directory of the YAWS source code distribution we have a file called `arg.yaws`. The file demonstrates the `Arg #arg` record parameter which is passed to the `out/1` function.

But before we discuss that code, we describe the `Arg` record in detail.

Here is the `yaws_api.hrl` file which is included by default in all YAWS files. The `#arg{}` record contains many fields that are useful when processing HTTP request dynamically. We have access to basically all the information associated with the client request such as:

- The actual socket leading back to the HTTP client
- All the HTTP headers – parsed into a `#headers` record
- The HTTP request – parsed into a `#http_request` record
- `clidata` – data which is POSTed by the client
- `querydata` – this is the remainder of the URL following the first occurrence of a `'?'` character, if any.
- `docroot` – the absolute path to the docroot of the virtual server that is processing the request.

```
-record(arg, {
    clisock,           % the socket leading to the peer client
    client_ip_port,   % {ClientIp, ClientPort} tuple
    headers,          % headers
    req,              % request
    clidata,          % The client data (as a binary in POST requests)
    server_path,      % The normalized server path
                    % (pre-querystring part of URI)
    querydata,        % For URIs of the form ...?querydata
                    % equiv of cgi QUERY_STRING
    appmoddata,       % (deprecated - use pathinfo instead) the remainder
                    % of the path leading up to the query
    docroot,          % Physical base location of data for this request
    docroot_mount,    % virtual directory e.g /myapp/ that the docroot
                    % refers to.
    fullpath,         % full deep path to yaws file
    cont,             % Continuation for chunked multipart uploads
    state,            % State for use by users of the out/1 callback
    pid,              % pid of the yaws worker process
})
```

```

    opaque,          % useful to pass static data
    appmod_prepath, % (deprecated - use prepath instead) path in front
                    % of: <appmod><appmoddata>
    prepath,         % Path prior to 'dynamic' segment of URI.
                    % ie http://some.host/<prepath>/<script-point>/d/e
                    % where <script-point> is an appmod mount point,
                    % or .yaws,.php,.cgi,.fcgi etc script file.
    pathinfo         % Set to '/d/e' when calling c.yaws for the request
                    % http://some.host/a/b/c.yaws/d/e
                    % equiv of cgi PATH_INFO
  })).

-record(http_request, {method,
                      path,
                      version}).

-record(headers, {
  connection,
  accept,
  host,
  if_modified_since,
  if_match,
  if_none_match,
  if_range,
  if_unmodified_since,
  range,
  referer,
  user_agent,
  accept_ranges,
  cookie = [],
  keep_alive,
  location,
  content_length,
  content_type,
  content_encoding,
  authorization,
  transfer_encoding,
  x_forwarded_for,
  other = [] % misc other headers
}).

```

There are a number of *advanced* fields in the #arg record such as appmod and opaque that will be discussed in later chapters.

Now, we show some code which displays the content of the Arg #arg record. The code is available in `yaws/www/arg.yaws` and after a `local_install` a request to `http://localhost:8000/arg.yaws` will

run the code.

```

<html>

<h2> The Arg </h2>

<p>This page displays the Arg #argument structure
supplied to the out/1 function.

<erl>

out(A) ->
  Req = A#arg.req,
  H = yaws_api:reformat_header(A#arg.headers),
  {ehtml,
    [{h4,[], "The headers passed to us were:"},
     {hr},
     {ol, [],lists:map(fun(S) -> {li,[], {p,[],S}} end,H)},

     {h4, [], "The request"},
     {ul,[],
      [{li,[], f("method: ~s", [Req#http_request.method])},
       {li,[], f("path: ~p", [Req#http_request.path])},
       {li,[], f("version: ~p", [Req#http_request.version])}]},

     {hr},
     {h4, [], "Other items"},
     {ul,[],
      [{li,[], f("clisock from: ~p", [inet:peername(A#arg.clisock)])},
       {li,[], f("docroot: ~s", [A#arg.docroot])},
       {li,[], f("fullpath: ~s", [A#arg.fullpath])}]},
     {hr},
     {h4, [], "Parsed query data"},
     {pre,[], f("~p", [yaws_api:parse_query(A)])},
     {hr},
     {h4,[], "Parsed POST data "},
     {pre,[], f("~p", [yaws_api:parse_post(A)])}]}},

</erl>

</html>

```

The code utilizes four functions from the `yaws_api` module. The `yaws_api` module is a general purpose www API module that contains various functions that are handy while developing YAWS code. We will see many more of those functions during the examples in the following chapters.

The functions used are:

- `yaws_api:f/2` — alias for `io_lib:format/2`. The `f/2` function is automatically -included in all YAWS code.
- `yaws_api:reformat_header/1` — This function takes the `#headers` record and unparses it, that is reproduces regular text.
- `yaws_api:parse_query/1` — The topic of the next section.
- `yaws_api:parse_post/1` — Ditto.

4.3 POSTs

4.3.1 Queries

The user can supply data to the server in many ways. The most common is to give the data in the actual URL. If we invoke:

```
GET http://localhost:8000/arg.yaws?kalle=duck&goofy=unknown
```

we pass two parameters to the `arg.yaws` page. That data is URL-encoded by the browser and the server can retrieve the data by looking at the remainder of the URL following the `'?'` character. If we invoke the `arg.yaws` page with the above mentioned URL we get as the result of `yaws_parse_query/1`:

kalle = duck

goofy = unknown

In ERLANG terminology, the call `yaws_api:parse_query(Arg)` returns the list:

```
[{"kalle", "duck"}, {"goofy", "unknown"}]
```

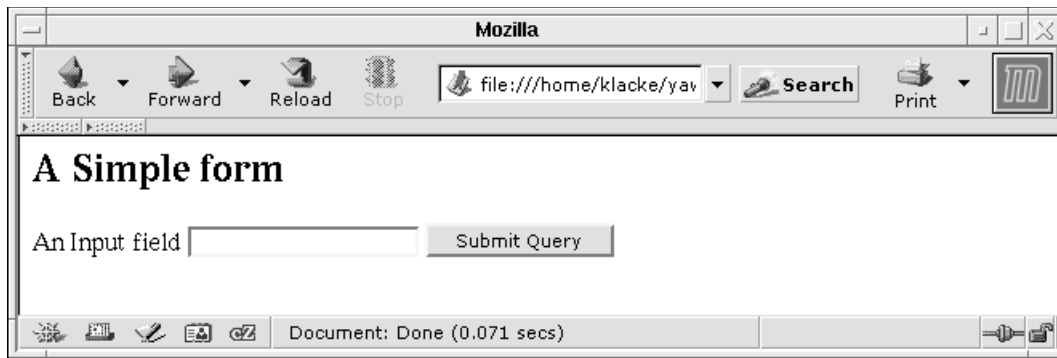
Both the key and the value are strings. Hence, a web page can contain URLs with a query and thus pass data to the web server. This scheme works both with `GET` and `POST` requests. It is the easiest way to pass data to the Web server since no form is required in the web page.

4.3.2 Forms

In order to `POST` data a form is required. Say that we have a page called `form.yaws` that contain the following code:

```
<html>
<form action="/post_form.yaws"
      method="post">

<p> A Input field
<input name="xyz" type="text">
<input type="submit">
</form>
</html>
```



This will produce a page with a simple input field and a submit button.

If we enter something—say, “Hello there”—in the input field and click the submit button the client will request the page indicated in the “action” attribute, namely `post_form.yaws`.

If that YAWS page has the following code:

```
out(A) ->
  L = yaws_api:parse_post(A),
  {html, f("~p", [L])}
```

The user will see the output

```
[{"xyz", "Hello there"}]
```

The differences between using the query part of the URL and a form are the following:

- Using the query arg only works in a GET request. We parse the query argument with the function `yaws_api:parse_query(Arg)`
- If we use a form and POST the user data the client will transmit the user data in the body of the request. That is, the client sends a request to get the page using the POST method and it then attaches the user data—encoded—into the body of the request.

A POST request can have a query part in its URL as well as user data in the body.

4.4 POSTing files

It is possible to upload files from the client to the server by means of POST. We indicate this in the form by telling the browser that we want a different encoding. Here is an example form that does this:

```
out(A) ->
  Form =
    {form, [{enctype, "multipart/form-data"},
```

```

    {method, post},
    {action, "file_upload_form.yaws"}],
    [{input, [{type, submit}, {value, "Upload"}]},
     {input, [{type, file}, {width, "50"}, {name, foo}]}]],
    {ehtml, {html, [], [{h2, [], "A simple file upload page"},
                        Form]}}}.

```

As shown in the figure, the page delivers the entire HTML page with enclosing `html` markers.



The user gets an option to browse the local host for a file or the user can explicitly fill in the file name in the input field. The file browsing part is automatically taken care of by the browser.

The action field in the form states that the client shall POST to a page called `file_upload_form.yaws`. This page will get the contents of the file in the body of the POST message. To read it, we use the `yaws_multipart` module, which provides the following capabilities:

1. It reads all parameters — files uploaded and other simple parameters.
2. It takes a few options to help file uploads. Specifically:
 - (a) `{max_file_size, MaxBytes}`: if the file size in bytes exceeds `MaxBytes`, return an error
 - (b) `no_temp_file`: read the uploaded file into memory without any temp files
 - (c) `{temp_file, FullFilePath}`: specify `FullFilePath` for the temp file; if not given, a unique file name is generated
 - (d) `{temp_dir, TempDir}`: specify `TempDir` as the directory to store the uploaded temp file; if this option is not provided, then by default an OS-specific temp directory such as `/tmp` is used
 - (e) `list`: return file data in list form; this is the default
 - (f) `binary`: return file data in binary form

Note that the `list` and `binary` options affect only file data, not filenames, headers, or other parameters associated with each file. These are always returned as strings.

Just call `yaws_multipart:read_multipart_form` from your `out/1` function and it'll return a tuple with the first element set to one of these three atoms:

- `get_more`: more data needs to be read; return this tuple directly to YAWS from your `out/1` function and it will call your `out/1` function again when it has read more POST data, at which point you must call `read_multipart_form` again
- `done`: multipart form reading is complete; a dict full of parameters is returned
- `error`: an error occurred

The dict returned with `done` allows you to query it for parameters by name. For file upload parameters, it returns one of the following lists:

```
[{filename, "name of the uploaded file as entered on the form"},
 {value, Contents_of_the_file_all_in_memory} | _T]
```

or:

```
[{filename, "name of the uploaded file as entered on the form"},
 {temp_file, "full pathname of the temp file"} | _T]
```

Some multipart/form messages also headers such as `Content-Type` and `Content-Transfer-Encoding` for different subparts of the message. If these headers are present in any subpart of a multipart/form message, they're also included in that subpart's parameter list, like this:

```
[{filename, "name of the uploaded file as entered on the form"},
 {value, Contents_of_the_file_all_in_memory},
 {content_type, "image/png"} | _T]
```

Note that for the temporary file case, it's your responsibility to delete the file when you're done with it.

Here's an example:

```
-module(my_yaws_controller).
-export([out/1]).

out(Arg) ->
    Options = [no_temp_file],
    case yaws_multipart:read_multipart_form(Arg, Options) of
        {done, Params} ->
            io:format("Params : ~p~n", [Params]),
            {ok, [{filename, FileName},{value,FileContent}|_]} =
                dict:find("my_file", Params),
            AnotherParam = dict:find("another_param", Params);
            %% do something with FileName, FileContent and AnotherParam
        {error, Reason} ->
            io:format("Error reading multipart form: ~s~n", [Reason]);
        Other -> Other
    end.
```

Here, `my_yaws_controller` is a user-defined module compiled as usual with `erlc` with the resulting `.beam` file placed in the YAWS load path. The module is then registered with YAWS as an *appmod* to allow it to receive and process requests—see section 7.2 for more details.

Chapter 5

Mode of operation

5.1 On-the-fly compilation

When the client requests a YAWS page, YAWS will look in its caches (there is one cache per virtual server) to see if it finds the requested page in the cache. If YAWS doesn't find the page in the cache, it will compile the page. This only happens the first time a page is requested. Say that the page is 400 bytes big and has the following layout:

100 bytes of HTML code
120 bytes of Erlang code
80 bytes of HTML code
60 bytes of Erlang code
140 bytes of HTML code

The YAWS server will then parse the file and produce a structure which makes it possible to readily deliver the page without parsing the next time the same page is requested.

When shipping the page it will

1. Ship the first 100 bytes from the file
2. Evaluate the first ERLANG chunk in the file and ship the output from the `out/1` function in that chunk. It will also jump ahead in the file and skip 120 bytes.
3. Ship 80 bytes of HTML code

4. Again evaluate an ERLANG chunk, this time the second and jump ahead 60 bytes in the file.
5. And finally ship 140 bytes of HTML code to the client

YAWS writes the source output of the compilation into a directory `/tmp/yaws/$UID`. The beam files are never written to a file. Sometimes it can be useful to look at the generated source code files, for example if the YAWS/ERLANG code contains a compilation error which is hard to understand.

5.2 Evaluating the Yaws Code

All client requests will execute in their own ERLANG process. For each group of virtual hosts on the same IP:PORT pair one ERLANG process listens for incoming requests.

This process spawns acceptor processes for each incoming request. Each acceptor process reads and parses all the HTTP headers from the client. It then looks at the `Host:` header to figure out which virtual server to use, i.e. which docroot to use for this particular request. If the `Host:` header doesn't match any server from *yaws.conf* with that IP:PORT pair, the first one from *yaws.conf* is chosen.

By default YAWS will not ship any data at all to the client while evaluating a YAWS page. The headers as well as the generated content are accumulated and not shipped to the client until the entire page has been processed.

Chapter 6

SSL

Secure Socket Layer (SSL) is a protocol used on the Web for delivering encrypted pages to the WWW client. SSL is widely deployed on the Internet and virtually all bank transactions as well as all online shopping today is done with SSL encryption. There are many good sources on the net that describe SSL in detail, so we will not try to do that here. See for example <http://www.tldp.org/HOWTO/SSL-Certificates-HOWTO/>, which describes how to manage certificates and keys.

In order to run an SSL server we must have a certificate. Either we can create a so-called self-signed certificate ourselves or buy a certificate from one of the many CA's (Certificate Authorities) on the net. YAWS uses the OTP interface to OpenSSL.

To setup a YAWS server with SSL we could have a *yaws.conf* file that looks like:

```
logdir = /var/log/yaws

<server www.funky.org>
    port = 443
    listen = 192.168.128.32
    docroot = /var/yaws/www.funky.org
    <ssl>
        keyfile = /etc/funky.key
        certfile = /etc/funky.cert
        password = gazonk
    </ssl>
</server>
```

This is the easiest possible SSL configuration. The configuration refers to a certificate file and a key file. The certificate file must contain the name "www.funky.org" as its "Common Name".

The keyfile is the private key file and it is encrypted using the password "gazonk".

Chapter 7

Applications

YAWS is well suited for Web applications. In this chapter we will describe a number of application templates. Code and strategies that can be used to build Web applications.

There are several ways of starting applications from YAWS.

- The first and most easy variant is to specify the `-r Module` flag to the YAWS startup script. This will apply `(Module, start, [])`
- We can also specify `runmods` in the `yaws.conf` file. It is possible to have several modules specified if want the same YAWS server to run several different applications.

```
runmod = myapp
runmod = app_number2
```

- It is also possible to do it the other way around, let the main application start YAWS. We call this embedded mode, which we will discuss in chapter 12.

7.1 Login scenarios

Many Web applications require the user to login. Once the user has logged in the server sets a cookie and then the user will be identified by help of the cookie in subsequent requests.

7.1.1 The session server

The cookie is passed in the headers and is available to the YAWS programmer in the `Arg #arg` record. The YAWS session server can help us to maintain a state for a user while the user is logged in to the application. The session server has the following 5 API functions to aid us:

1. `yaws_api:new_cookie_session(Opaque)` — This function initiates a new cookie-based session. The `Opaque` data is typically some application-specific structure which makes it possible for the application to read a user state, or it can be the actual user state itself.

2. `yaws_api:cookieval_to_opaque(Cookie)` — This function maps a cookie to a session.
3. `yaws_api:replace_cookie_session(Cookie, NewOpaque)` — Replace the opaque user state in the session server with `NewOpaque`.
4. `yaws_api:delete_cookie_session(Cookie)` — This function should typically be called when the user logs out or when our web application decides to automatically logout the user.

All cookie-based applications are different but they have some things in common. In the following examples we assume the existence of a function `myapp:auth(UserName, Passwd)` and it returns `ok` or `{error, Reason}`.

Let's assume the following record:

```
-record(session, {user,
                  passwd,
                  udata = []}).
```

The following function is a good template function to check the cookie.

```
get_cookie_val(CookieName, Arg) ->
    H = Arg#arg.headers,
    yaws_api:find_cookie_val(CookieName, H#headers.cookie).

check_cookie(A, CookieName) ->
    case get_cookie_val(CookieName, A) of
        [] ->
            {error, "not logged in"};
        Cookie ->
            yaws_api:cookieval_to_opaque(Cookie)
    end.
```

We need to check all requests and make sure the session_server has our cookie registered as an active session. Also, if a request comes in without a working cookie we want to present a login page instead of the page the user requested. Another quirky issue is that the pages necessary for display of the login page must be shipped without checking the cookie. The next sections explain how these needs can be met.

7.1.2 Arg rewrite

In this section we describe a feature whereby the user is allowed to rewrite the `Arg` at an early stage in the YAWS server. We do that by specifying an `arg_rewrite_mod` in the `yaws.conf` file.

```
arg_rewrite_mod = myapp
```

Then in the myapp module we have:

```
arg_rewrite(Arg) ->
    OurCookieName = "myapp_sid"
    case check_cookie(A, OurCookieName) of
        {error, _} ->
            do_rewrite(Arg);
        {ok, _Session} ->
            %return Arg untouched
            Arg
    end.

%% these pages must be shippable without a good cookie
login_pages() ->
    ["/banner.gif", "/login.yaws", "/post_login.yaws"].

do_rewrite(Arg) ->
    Req = Arg#arg.req,
    {abs_path, Path} = Req#http_request.path,
    case lists:member(Path, login_pages()) of
        true ->
            Arg;
        false ->
            Arg#arg{req = Req#http_request{path = {abs_path, "/login.yaws"}},
                    state = {abs_path, Path}}
    end.
```

Our arg rewrite function lets all Args go through untouched that either have a good cookie or belong to a set of predefined pages that are acceptable to get without being logged in. If we decide that the user must log in, we change the path of the request, thereby making the YAWS server ship a login page instead of the page the user requested. We also set the original path in the Arg state argument so that the login page can redirect the user to the original page once the login procedure is finished.

Within an arg rewrite function, examining and modifying HTTP headers can be achieved using the following functions from the `yaws_api` module:

- `set_header/2, set_header/3`
- `get_header/2, get_header/3`
- `delete_header/2`

All functions operate on instances of the `#headers{}` record type defined in the `yaws_api.hrl` include file.

7.1.3 Authenticating

Now we're approaching the `login.yaws` page, the page that displays the login prompt to the user. The login page consists of two parts: one part that displays the login data as a form, and one form processing page that reads the data the user entered in the login fields and performs the actual authentication.

The login page performs a tiny well-known Web trick where it passes the original URL request in a hidden field in the login page and thereby passing that information to the form processing page.

The page `login.yaws`:

```
<erl>

out(A) ->
    {ehtml,
     {html, [],
      [{h2, [], "Login page"},
       {hr},
       {form, [{action, "/login_post.yaws"},
               {method, post}],

               [{p, [], "Username"}, {input, [{type, text}, {name, uname}]},
                {p, [], "Password"}, {input, [{type, password}, {name, passwd}]},
                {input, [{type, submit}, {value, "Login"}]},
                {input, [{type, hidden}, {name, url},
                        {value, A#arg.state}]}]}]}].

</erl>
```

The form processing page which gets the `POST` data from the code above:

```
<erl>

-include("myapp.hrl").
%% we have the session record there
%% we must set the include_path in the yaws.conf file
%% in order for the compiler to find that file

kv(K,L) ->
    {value, {K, V}} = lists:keysearch(K,1,L),
    V.

out(A) ->
    L = yaws_api:parse_post(A),
    User = kv(user, L),
    Pwd = kv(passwd, L),
```

```

case myapp:auth(User, Pwd) of
  ok ->
    S = #session{user = User,
                  passwd = Pwd,
                  udata = []},
    %% Now register the session to the session server
    Cookie = yaws_api:new_cookie_session(S),
    [{redirect_local, kv(url, L)},
     yaws_api:setcookie("myapp_sid",Cookie)]
  Err ->
    {ehtml,
     {html, [],
      {p, [], f("Bad login: ~p",[Err])}}}
end.
</erl>

```

The function returns a list of two new (not previously discussed) return values: instead of returning HTML output as in `{html, Str}` or `{ehtml, Term}` we return a list of two new values. (There are many different possible return values from the `out/1` function and they will all be described later.) The two new values are:

1. The tuple `{redirect_local, Path}` makes the YAWS web server return a 302 redirect to the specified Path. Optionally a different status code can be supplied which will be used in place of 302, e.g. `{redirect_local, Path, 307}`.
2. `yaws_api:setcookie("myapp_sid",Cookie)` generates a Set-Cookie header.

Now if we put all this together we have a full-blown cookie-based login system. The last thing we did in the form processing code was to register the session with the session server thereby letting any future requests go straight through the Arg rewriter.

This way both YAWS pages as well as all or some static content is protected by the cookie login code.

7.1.4 Database driven applications

We can use code similar to the code in the previous section to associate a user session to entries in a database. Mnesia fits perfectly together with YAWS and keeping user persistent state in Mnesia is both easy and convenient.

Once the user has logged in we can typically use the user name as key into the database. We can mix `ram_tables` and `disc_tables` to our liking. The Mnesia database must be initialized via `create_table/2` before it can be used. This is typically done while installing the web application on a machine.

Another option is to let the application check that Mnesia is initialized whenever the application starts.

If we don't want or need to use Mnesia, it's of course possible to use a simple `dets` file or a text file as well.

7.2 Appmods

Appmods is mechanism to invoke different applications based upon the URL. A URL—as presented to the web server in a request—has a path part and a query part.

It is possible to install several appmods in the *yaws.conf* file as shown below:

```
appmods = foo myapp
```

Now, if the user requests a URL where any component in the directory path is an appmod, the parsing of the URL will terminate there and instead of reading the actual file from the disk, YAWS will invoke the appmod with the remainder of the path inserted into `Arg#arg.appmoddata`.

Say the user requests the URL `http://www.funky.org/myapp/xx/bar.html`. YAWS will not ship the file `bar.html` to the client, instead it will invoke `myapp:out(Arg)` with `Arg#arg.appmoddata` set to the string `xx/bar.html`. Any optional query data—that is, data that follows the first `'?'` character in the URL—is removed from the path and passed as `Arg#arg.querydata`.

Appmods can be used to run applications on a server. All requests to the server that has an appmod in the URL will be handled by that application. If the application decides that it want to ship a page from the disk to the client, it can return the tuple `{page, Path}`. This return value will make YAWS read the page from the disk, possibly add the page to its cache of commonly accessed pages and ship it back to the client.

The `{page, Path}` return value is equivalent to a redirect, but it removes an extra round trip, and is thus faster.

Appmods can also be used to fake entire directory hierarchies that don't exist on disk.

7.3 The opaque data

Sometimes an application needs application-specific data such as the location of its data files. There exists a mechanism to pass application-specific configuration data from the YAWS server to the application.

When configuring a server we have an opaque field in the configuration file that can be used for this purpose. Say we have the following fields in the config file:

```
<server foo>
  listen = 192.168.128.44
  <opaque>
    foo = bar
    somefile = /var/myapp/db
    myname = hyber
  </opaque>
</server>
```

This will create a normal server that listens to the specified IP address. An application has access to the opaque data that was specified in that particular server through `Arg#arg.opaque`.

If we have the opaque data specified above, the `Arg opaque` field will have the value:

```
[{foo, "bar"},
 {somefile, "/var/myapp/db"},
 {myname, "hyber"}
]
```

7.4 Customizations

When actually deploying an application at a live site, some of the standard YAWS behaviors are not acceptable. Many sites want to customize the web server behavior when a client requests a page that doesn't exist on the web server. The standard YAWS behavior is to reply with status code 404 and a message explaining that the page doesn't exist.

Similarly, when YAWS code crashes, the reason for the crash is displayed in the Web browser. This is very convenient while developing a site but not acceptable in production.

7.4.1 404 File not found

We can install a special handler for 404 messages. We do that by specifying a `errormod_404` in the *yaws.conf* file.

If we have:

```
<server foo>
  ..
  ..
  ..
  errormod_404 = myapp

</server>
```

When YAWS gets a request for a file that doesn't exist, it invokes the `errormod_404` module to generate both the status code as well as the content of the message.

If `Module` is specified as the `errormod_404` module, YAWS will invoke `Module:out404(Arg, GC, SC)`, passing the arguments as described below:

- `Arg` is an `#arg{}` record
- `GC` is a `#gconf{}` record (defined in *yaws.hrl*)
- `SC` is a `#sconf{}` record (defined in *yaws.hrl*)

The function can and must do the same things that a normal `out/1` function does.

7.4.2 Crash messages

We use a similar technique for generating the crash messages: we install a module in the *yaws.conf* file and let that module generate the crash message. We have:

```
errormod_crash = Module
```

The default is to display the entire formatted crash message in the browser. This is good for debugging but not good for production.

If *Module* is specified as the *errormod_crash* module, the function *Module:crashmsg(Arg, SC, Str)* will be called. The *Str* argument is the real crash message formatted as a string.

7.5 Stream content

If the *out/1* function returns the tuple *{content, MimeType, Content}* YAWS will ship that data to the Client. This way we can deliver dynamically generated content to the client which is of a different MIME type than "text/html".

If the generated response is very large and it not possible or practical to generate the whole thing, we can return the value:

```
{streamcontent, MimeType, FirstChunk}
```

which delivers data back to the client using HTTP chunked transfer (see RFC 2616 section 3.6.1) and then from a different ERLANG process deliver the remaining chunks by using the functions described below:

1. *yaws_api:stream_chunk_deliver(YawsPid, Data)* where the *YawsPid* is the process id of the YAWS worker process. That pid is available in *Arg#arg.pid*.
2. *stream_chunk_end(YawsPid)* which must be called to indicate the end of the stream.

A streaming alternative is also available for applications that need a more direct way to deliver data to clients, such as those dealing with data too large to buffer in memory but not wishing to use chunked transfer, or applications that use long-polling (Comet) techniques that require them to hold client connections open for extended periods. For these situations we can return the value:

```
{streamcontent_from_pid, MimeType, Pid}
```

to tell YAWS we wish to deliver data of mime type *MimeType* to the client from process *Pid*. In this case, YAWS will prepare the socket for delivery from *Pid* and then send one of the following messages to *Pid*:

- *{ok, YawsPid}* tells *Pid* that it is now OK to proceed with sending data back to the client using the socket. The socket is accessible as *Arg#arg.clisock*.
- *{discard, YawsPid}* informs *Pid* that it should not attempt to use the socket, typically because the requested HTTP method requires no response body.

We call one of the following functions to send data:

- `yaws_api:stream_process_deliver(Socket, IoList)` sends data `IoList` using socket `Socket` without chunking the data. To ensure chunking is not in effect, return

```
{header, {transfer_encoding, erase}}
```

in a list with the `streamcontent_from_pid` directive (which must always be last in such a list).

- `yaws_api:stream_process_deliver_chunk(Socket, IoList)` sends data `IoList` using socket `Socket` but converts the data into chunked transfer form before sending it.

Pids using chunked transfer must indicate the end of their transfer by calling the following function:

- `yaws_api:stream_process_deliver_final_chunk(Socket, IoList)`

which delivers a special HTTP chunk to mark the end of the data transfer to the client.

Finally, `Pid` must always call `yaws_api:stream_process_end(Socket, YawsPid)` when it finishes sending data or when it receives the `{discard, YawsPid}` message from YAWS — this is required to inform YAWS that `Pid` has finished with the socket and will not use it directly anymore. If the application has to close the socket while it's in control of it, though, it must pass the atom `closed` as the first argument to `yaws_api:stream_process_end` in place of the socket to inform YAWS that the socket has been closed and it should no longer attempt to use it.

Applications that use the `streamcontent_from_pid` directive that also want to avoid chunked transfer encoding for their streams should be sure to include a setting for the `Content-Length` header in their `out/1` return value. YAWS automatically sets the `Transfer-Encoding` header to `chunked` if it does not detect a `Content-Length` header.

7.6 All out/1 return values

- `{html, DeepList}` This assumes that `DeepList` is formatted HTML code. The code will be inserted in the page.
- `{ehhtml, Term}` This will transform the ERLANG term `Term` into a stream of HTML content.
- `{content, MimeType, Content}` This function will make the web server generate different content than HTML. This return value is only allowed in a YAWS file which has only one `<erl> </erl>` part and no `html` parts at all.
- `{streamcontent, MimeType, FirstChunk}` This return value plays the same role as the `content` return value above. However it makes it possible to stream data to the client using HTTP chunked transfer if the YAWS code doesn't have access to all the data in one go. (Typically if a file is very large or if data arrives from back end servers on the network.)
- `{streamcontent_with_timeout, MimeType, FirstChunk, Timeout}` Similar to the above, but with an explicit timeout. The default timeout is 30 secs, i.e. if the application fails to deliver data to the YAWS process, the streaming will stop. This is often not the desired behaviour in Comet/Ajax applications. It's possible to provide 'infinity' as the timeout.

- `{streamcontent_from_pid, MimeType, Pid}` This return value is similar to the `streamcontent` return value above. However it makes it possible to stream data to the client directly from an application process to the socket. This approach can be useful for applications that employ long-polling (Comet) techniques, for example, and for applications wanting to avoid buffering data or avoid HTTP chunked mode transfer for streamed data.
- `{streamcontent_with_size, Sz, MimeType, Pid}` This is similar to the `streamcontent` return value above. However it makes it possible to stream data to the client by setting the content length of the response. As the opposite of other ways to stream data, in this case, the response is not chunked encoded.
- `{header, H}` Accumulates a HTTP header. Used by for example the `yaws_api:setcookie/2-6` function.
- `{header, {H, erase}}` A specific case of the previous directive; use this to remove a specific header from a response. For example, streaming applications and applications using server-sent events (see <http://www.w3.org/TR/eventsource/>) should use `{header, {transfer_encoding, erase}}` to turn off chunked encoding for their responses.
- `{allheaders, HeaderList}` Will clear all previously accumulated headers and replace them.
- `{status, Code}` Sets the response HTTP status code to `Code`.
- `break` Will stop processing of any consecutive chunks of `erl` or HTML code in the YAWS file.
- `ok` Do nothing.
- `flush` Flush remaining data sent by the client.
- `{redirect, Url}` Erase all previous headers and accumulate a single HTTP Location header. Set the status code to 302.
- `{redirect, Url, Status}` Same as `redirect` above with the additional option of supplying the status code. The default for a redirect is 302 but 301, 303 and 307 are also valid redirect status codes.
- `{redirect_local, Path}` Does a redirect to the same `Scheme://Host:Port/Path` in which we are currently executing. `Path` can be either be the path directly (equivalent to `abs_path`), or one of `{{abs_path, Path}}` or `{{rel_path, RelativePath}}`
- `{redirect_local, Path, Status}` Same as `redirect_local` above with the additional option of supplying the status code. The default for a redirect is 302 but 301, 303 and 307 are also valid redirect status codes.
- `{get_more, Cont, State}` When we are receiving large POSTs we can return this value and be invoked again when more data arrives.
- `{page, Page}` Make YAWS return a different local page than the one being requested.
- `{page, {Options, Page}}` Like the above, but supplying an additional deep list of options. For now, the only type of option is `{header, H}` with the effect of accumulating the HTTP header `H` for page `Page`.
- `{websocket, CallbackModule, Options}` Tell YAWS to use `CallbackModule` as a WebSockets Protocol handler for traffic on the client socket. See chapter 14 for more details.

- [ListOfValues] It is possible to return a list of the above defined return values. Any occurrence of the atoms `streamcontent`, `streamcontent_with_timeout`, `streamcontent_with_size`, `streamcontent_from_pid`, `get_more`, `page` or `break` in this list is legal only if it is the last position of the list. If not, remaining values in the list are ignored.

Chapter 8

Debugging and Development

YAWS has excellent debugging capabilities. First and foremost we have the ability to run the web server in interactive mode by means of the command line switch `-i`, which gives us a regular ERLANG command line prompt we can use to compile helper code or reload helper code. Furthermore all error messages are displayed there. If a `.yaws` page produces any regular ERLANG I/O, that output will be displayed at the ERLANG prompt, assuming we are running in interactive mode.

If we give the command line switch `-d` we get some additional error messages. Also YAWS does some additional checking of user supplied data such as headers.

8.1 Logs

YAWS produces various logs. All log files are written into the YAWS logdir directory. This directory is specified in the config file.

We have the following log files:

- The access log. Access logging is turned on or off per server in the *yaws.conf* file. If `access_log` is turned on for a server, YAWS will produce a log in Common Access Log Format called *HostName:PortNumber.access*
- The auth log. Auth logging is turned on or off per server in the *yaws.conf* file. If `auth_log` is turned on for a server, YAWS will produce a log called *HostName:PortNumber.auth* which contains all HTTP auth-related messages.
- *report.log* This file contains all error and crash messages for all virtual servers in the same file.
- Trace files. The two command line flags `-t` and `-T` tells YAWS to trace all traffic or just all HTTP messages and write them to files.

Chapter 9

External scripts via CGI

YAWS can also interface to external programs generating dynamic content via the Common Gateway Interface (CGI). This has to be explicitly enabled for a virtual host by listing `cgi` in the `allowed_scripts` line in the configuration file. Any request for a page ending in `.cgi` (or `.CGI`) will then result in trying to execute the corresponding file.

If you have a PHP executable compiled for using CGI in the `PATH` of the YAWS server, you can enable PHP support by adding `php` to `allowed_scripts`. Requests for pages ending in `.php` will then result in YAWS executing `php` (configurable via `php_handler`) and passing the name of the corresponding file to it via the appropriate environment variable.

These ways of calling CGI scripts are also available to `.yaws` scripts and `appmods` via the functions `yaws_api:call_cgi/2` and `yaws_api:call_cgi/3`. This makes it possible to write wrappers for CGI programs, irrespective of the value of `allowed_scripts`.

The author of this YAWS feature uses it for self-written CGI programs as well as for using a standard CGI package. You should not be surprised however, should some scripts not work as expected due to an incomplete or incorrect implementation of certain CGI meta-variables. The author of this feature is interested in hearing about your experiences with it. He can be contacted at `carsten@codimi.de`.

Chapter 10

FastCGI

YAWS supports the responder role and the authorizer role of the FastCGI protocol. See www.fastcgi.com for details on the FastCGI protocol.

The benefits of using FastCGI include:

1. Unlike CGI, it is not necessary to spawn a new process for every request; the application server can handle multiple requests in a single process.
2. The fact that the application server can run on a different computer benefits scalability and security.
3. The application server can be written in any language for which a FastCGI library is available. Existing applications which have been written for other web servers can be used with YAWS.
4. FastCGI can also be used to implement external authentication servers (in addition to generating dynamic content).

Support for FastCGI was added to YAWS by Bruno Rijsman (brunorijsman@hotmail.com).

10.1 The FastCGI Responder Role

The FastCGI responder role allows YAWS to communicate with an application server running on a different (or on the same) computer to generate dynamic content.

The FastCGI protocol (which runs over TCP) is used to send the request information from YAWS to the application server and to send the response information (e.g. the generated dynamic content) from the application server back to YAWS.

FastCGI responders can be invoked in two ways:

1. By including `fcgi` in the `allowed_scripts` line in the configuration file (note that the default value for `allowed_scripts` includes `fcgi`).

In this case a request for any resource with the `.fcgi` extension will result in a FastCGI call to the application server to dynamically generate the content.

Note: the YAWS server will only call the application server if a file corresponding to the resource name (i.e. a file with the `.fcgi` extension) exists locally on the YAWS server. The contents of that file are not relevant.

2. By creating an appmod which calls `yaws_api:call_fcgi_responder`. See the `yaws_api(5)` man page for details.

10.2 The FastCGI Authorizer Role

The FastCGI authorizer role allows YAWS to communicate with an authentication server to authenticate requests.

The FastCGI protocol is used to send the request information from YAWS to the authentication server and the authentication response back from the authentication server to YAWS.

If access is allowed, YAWS proceeds to process the request normally.

If access is denied, the authentication server provides the response which is sent back to the client. This is typically a “not authorized” response or a redirect to a login page.

FastCGI authorizers are invoked by creating an appmod which calls `yaws_api:call_fcgi_authorizer`. See the `yaws_api(5)` man page for details.

10.3 The FastCGI Filter Role

FastCGI defines a third role, the filter role, which YAWS does not currently support.

10.4 FastCGI Configuration

The following commands in the `yaws.conf` file control the operation of FastCGI.

If you use FastCGI, you *must* include the `fcgi_app_server` setting in the configuration file to specify the host name (or IP address) and TCP port of the FastCGI application server.

You may include the `fcgi_trace_protocol` setting to enable or disable tracing of FastCGI protocol messages. This is useful for debugging.

You may include the `fcgi_log_app_error` setting to enable or disable logging application errors (any output to stderr and non-zero exit codes).

You may include the `extra_cgi_vars` command to pass additional environment variables to the application.

Chapter 11

Security

YAWS is of course susceptible to intrusions. YAWS has the ability to run under a different user than root, even if we need to listen to privileged port numbers. Running as root is generally a bad idea.

Intrusions can happen basically at all places in YAWS code where the YAWS code calls either the BIF `open_port` or when YAWS code calls `os:cmd/1`. Both `open_port` and `os:cmd/1` invoke the `/bin/sh` interpreter to execute its commands. If the commands are nastily crafted bad things can easily happen.

All data that is passed to these two function must be carefully checked.

Since YAWS is written in ERLANG a large class of cracks are eliminated since it is not possible to perform any buffer overrun cracks on a YAWS server. This is very good.

Another possible point of entry to the system is by providing a URL which takes the client out from the docroot. This should not be possible – and the impossibility relies on the correctness of the URL parsing code in YAWS.

11.1 WWW-Authenticate

YAWS has support for WWW-Authentication. WWW-Authenticate is a standard HTTP scheme for the basic protection of files with a username and password. When a client browser wants a protected file, it must send a `Authenticate: username:password` header in the request. Note that this is plain text. If there is no such header or the username and password is invalid the server will respond with status code 401 and the realm. Browsers will then tell the user that a username and password is needed for “realm”, and will resend the request after the user enters the information.

WWW-Authentication is configured in the *yaws.conf* file, in as many `<auth>` directives as you desire:

```
<server foo>
  docroot = /var/yaws/www/

..
..

<auth>
  realm = secretpage
```

```
dir    = /protected
dir    = /anotherdir
user   = klacke:gazonk
user   = jonny:xyz
user   = ronny:12r8uyp09jksfdge4
</auth>
</server>
```

YAWS will require one of the given username:password pairs for all files in the */protected* and */anotherdir* directories. Note that these directories are specified as a server path, that is, the filesystem path that is actually protected here is */var/yaws/www/protected* .

Chapter 12

Embedded mode

YAWS is a normal OTP application. It is possible to integrate YAWS into another larger application. The YAWS source tree must be integrated into the larger application's build environment. YAWS is then simply started by `application:start()` from the larger application's boot script, or the YAWS components needed for the larger application can be started individually under the application's supervisor(s).

By default YAWS reads its configuration data from a config file, the default is `/usr/local/etc/yaws/yaws.conf`. If YAWS is integrated into a larger application, however, that application typically has its configuration data kept at some other centralized place. Sometimes we may not even have a file system to read the configuration from if we run a small embedded system.

YAWS reads its application environment. If the environment key `embedded` is set to `true`, YAWS starts in embedded mode. Once started it must be fed a configuration, and that can be done after YAWS has started by means of the function `yaws_api:setconf/2`.

It is possible to call `setconf/2` several times to force YAWS to reread the configuration.

12.1 Creating Global and Server Configurations

The `yaws_api:setconf/2` function mentioned in the previous section takes two arguments:

- a `#gconf` record instance, specifying global YAWS configuration
- a list of lists of `#sconf` record instances, each specifying configuration for a particular server instance

These record types are specified in `yaws.hrl`, which is not normally intended for inclusion by applications. Instead, YAWS provides the `yaws_api:embedded_start_conf/1,2,3,4` functions that allow embedded mode applications to specify configuration data using property lists (lists of `{key, value}` pairs).

The `yaws_api:embedded_start_conf` functions all return a tuple containing the following four items:

- the atom `ok`.
- a list of lists of `#sconf` record instances. This variable is intended to be passed directly to `yaws_api:setconf/2` as its second argument.

- a `#gconf` record instance. This variable is intended to be passed directly to `yaws_api:setconf/2` as its first argument.
- a list of supervisor child specification for the YAWS components the embedded mode application's configuration specified should be started. This allows embedded mode applications to start YAWS under its own supervisors.

Note that `yaws_api:embedded_start_conf` does not actually start any servers, but rather it only returns the configuration information and child specifications needed for the embedded mode application to start and configure YAWS itself.

If you have difficulty figuring out how to set up your `#gconf` or `#sconf` records for embedded mode, you should first consider getting something running in non-embedded mode using a `yaws.conf` file. Once you're satisfied with the setup and YAWS is running, execute the following command:

```
yaws --running-config
```

This command will show the YAWS configuration from your `yaws.conf` file in terms of `#gconf` and `#sconf` records, thus showing you how to set up those records for embedded mode.

12.2 Starting Yaws in Embedded Mode

An embedded mode application can start YAWS in one of two ways:

- It can call `yaws_api:embedded_start_conf` to obtain configuration and YAWS startup information as described in the previous section, start YAWS under its own supervisors, and then pass the global and server configuration settings to `yaws_api:setconf/2`.
- It can call `yaws:start_embedded/1,2,3,4`, each of which takes exactly the same arguments as the corresponding `yaws_api:embedded_start_conf/1,2,3,4` function. Instead of just returning start and configuration information, however, `yaws:start_embedded` also starts and configures YAWS, which can be more convenient but does not allow the embedded mode application any supervision control over YAWS.

Both of these functions take care of setting the environment key `embedded` to `true`. Neither approach requires any special settings in the embedded mode application's `.app` file nor any special command-line switches to the ERLANG runtime.

For an example of how to use `yaws_api:embedded_start_conf` along with `yaws_api:setconf`, please see the files `www/ybed_sup.erl` and `www/ybed.erl` in the YAWS distribution.

Chapter 13

The config file - yaws.conf

In this section we provide a complete listing of all possible configuration file options. The configuration contains two distinct parts: a global part which affects all the virtual hosts and a server part where options for each virtual host is supplied.

13.1 Global Part

- `logdir = Directory` — All YAWS logs will be written to files in this directory. There are several different log files written by YAWS.
 - `report.log` — this is a text file that contains all error logger printouts from YAWS.
 - `<Host>.access` — for each virtual host served by YAWS, a file `<Host>.access` will be written which contains an access log in Common Log Format.
 - `<Host>.auth` — for each virtual host served by YAWS, a file `<Host>.auth` will be written which contains all HTTP auth related messages.
 - `trace_<YYYYMMDD_hhmmss>` - Trace files are written in this subdirectory, suffixed by the creation date.
 - * `trace.<Pid>.http` - this file contains the HTTP trace if that is enabled, where `Pid` is the process id handling the TCP connection.
 - * `trace.<Pid>.traffic` - this file contains the traffic trace if that is enabled, where `Pid` is the process id handling the TCP connection.

Note that `<Host>.access` and `<Host>.auth` files will be used only if the directive `logger_mod` is not set or set to `yaws_log`.

The default value for `logdir` is `"."`

- `ebin_dir = Directory` — This directive adds `Directory` to the `ERLANG` search path. It is possible to have several of these command in the configuration file.
- `id = String` — It is possible run multiple Yaws servers on the same machine. We use the id of a Yaws server to control it using the different control commands such as:

```
# /usr/local/bin/yaws --id foobar --stop
```

To stop the Yaws server with id "foobar". Each YAWS server will write its internals data into a file called `$HOME/.yaws/yaws/ID` where ID is the identity of the server. Yaws also creates a file called `$VARDIR/run/yaws/ctl-$ID` which contain the port number where the server is listening for control commands. The default id is "default".

- `server_signature = String` — This directive sets the "Server: " output header to the custom value. The default value is "yaws/VSN, Yet Another Web Server".
- `include_dir = Directory` — This directive adds Directory to the path of directories where the ERLANG compiler searches for include files. We need to use this if we want to include `.hrl` files in our YAWS ERLANG code.
- `max_num_cached_files = Integer` — YAWS will cache small files such as commonly accessed GIF images in RAM. This directive sets a maximum number on the number of cached files. The default value is 400.
- `max_num_cached_bytes = Integer` — This directive controls the total amount of RAM which can maximally be used for cached RAM files. The default value is 1000000, 1 megabyte.
- `max_size_cached_file = Integer` — This directive sets a maximum size on the files that are RAM cached by YAWS. The default value is 8000 bytes.
- `cache_refresh_secs = Integer` — The RAM cache is used to serve pages that sit in the cache. An entry sits in cache at most `cache_refresh_secs` number of seconds. The default is 30. This means that when the content is updated under the docroot, that change doesn't show until 30 seconds have passed. While developing a YAWS site, it may be convenient to set this value to 0. If the debug flag (`-d`) is passed to the YAWS start script, this value is automatically set to 0.
- `trace = traffic | http` — This enables traffic or HTTP tracing. Tracing is also possible to enable with a command line flag to YAWS.
- `use_old_ssl = true | false` — This re-enables the old OTP ssl implementation. By default we use the new ssl implementation.
- `auth_log = true | false` — *Deprecated and ignored. Now, this target must be set in server part.*
- `max_connections = nolimit | Integer` — This value controls the maximum number of connections from HTTP clients into the server. This is implemented by closing the last socket if the threshold is reached.
- `keepalive_maxuses = nolimit | Integer` — Normally, YAWS does not restrict the number of times a connection is kept alive using keepalive. Setting this parameter to an integer X will ensure that connections are closed once they have been used X times. This can be a useful to guard against long-running connections collecting too much garbage in the ERLANG VM.
- `process_options = undefined | Proplist` — Set process spawn options for client acceptor processes. Options must be specified as a quoted string of either the atom `undefined` or as a proplist of valid process options. The supported options are `fullsweep_after`, `min_heap_size`, and `min_bin_vheap_size`, each taking an associated integer value. Other process options are ignored. The proplist may also be empty. See `erlang:spawn_opt/4` for details on these options.

- `acceptor_pool_size = Integer` — Set the size of the pool of cached acceptor processes. The specified value must be greater than or equal to 0. The default value is 8. Specifying a value of 0 effectively disables the process pool.
- `log_wrap_size = Integer` — The logs written by YAWS are all wrap logs, the default value at the size where they wrap around and the original gets renamed to `File.old` is 1000000, 1 megabyte. This value can be changed.

If we set the value to 0 the logs will never wrap. If we want to use YAWS in combination with a more traditional log wrapper such as `logrotate`, set the size to 0 and YAWS will reopen the logfiles once they have been renamed/removed.

- `log_resolve_hostname = true | false` — By default the client host IP is not resolved in the access logs.
- `fail_on_bind_err = true | false` — Fail completely or not if YAWS fails to bind a listen socket. Default is `true`.
- `soap_srv_mods = ListOfModuleSetting` — If `enable_soap` is `true`, a startup YAWS will invoke `yaws_soap_srv:setup()` to setup modules set here. `ModuleSetting` is either a triad like `<Mod, HandlerFun, WsdlFile>` or a tetrad like `<Mod, HandlerFun, WsdlFile, Prefix>` which specifies the prefix. A prefix will be used as argument of `yaws_soap_lib:initModel()` and then be used as a XML namespace prefix. Note, the `WsdlFile` here should be an absolute-path file in local file systems.

For example, we can specify

```
soap_srv_mods=<Mod1,Handler,WsdlFile1> <Mod2,Handler,Wsdl2,SpecifiedPrefix>
```

- `php_exe_path = Path` — *this target is deprecated and useless. use 'php_handler' target in server part instead.*

The name of (and possibly path to) the php executable used to interpret php scripts (if allowed). Default is `php-cgi`.

- `copy_error_log = true | false` — Enable or disable copying of the error log. When we run in embedded mode, there may very well be some other systems process that is responsible for writing the errorlog to a file whereas when we run in normal standalone mode, we typically want the Erlang errorlog written to a `report.log` file. Default value is `true`.
- `ysession_mod = Module` — Allows to specify a different YAWS session storage mechanism instead of an ETS table. One of the drawbacks of the default `yaws_session_server` implementation is that server side cookies are lost when the server restarts. Specifying a different module here will pass all writes/read operations to this module (it must implement appropriate callbacks).
- `runmod = ModuleName` — At startup YAWS will invoke `ModuleName:start()` in a separate process. It is possible to have several runmods. This is useful if we want to reuse the Yaws startup shell script for our own application.
- `pick_first_virthost_on_nomatch = true | false` — When YAWS gets a request, it extracts the `Host:` header from the client request to choose a virtual server amongst all servers with the same IP/Port pair. This configuration parameter decides whether Yaws should pick the first (as defined in the `yaws.conf` file) if no name match or not. In real live hosting scenarios we typically want this to be `false` whereas in testing/development scenarios it may be convenient to set it to `true`. Default is `true`.

- `keepalive_timeout = Integer | infinity` — If the HTTP session will be kept alive (i.e., not immediately closed) it will close after the specified number of milliseconds unless a new request is received in that time. The default value is 30000. The value `infinity` is legal but not recommended.
- `subconfig = File` — Load specified config file.
- `subconfigdir = Directory` — Load all config files in specified directory.
- `x_forwarded_for_log_proxy_whitelist = ListOfUpstreamProxyServerIps` — *This target is deprecated and will be ignored.*

13.2 Server Part

YAWS can virthost several web servers on the same IP address as well as several web servers on different IP addresses. The only limitation here is that there can be only one server with SSL enabled per each individual IP address. Each virtual host is defined within a matching pair of `<server ServerName>` and `</server>`. The `ServerName` will be the name of the web server.

The following directives are allowed inside a server definition.

- `port = Port` — This makes the server listen on `Port`. Default is 8000.
- `listen = IpAddress` — This makes the server listen on `IpAddress` when virthosting several servers on the same IP/port address, if the browser doesn't send a `Host:` field, YAWS will pick the first server specified in the config file. Multiple `listen` directives may be used to specify several addresses to listen on.
- `listen_backlog = Integer` — This sets the TCP listen backlog for the server to define the maximum length the queue of pending connections may grow to. The default is the same as the default provided by `gen_tcp:listen/2`, which is 5.
- `server_signature = String` — This directive sets the "Server: " output header to the custom value and overloads the global one for this virtual server.
- `rhost = Host[:Port]` — This forces all local redirects issued by the server to go to `Host`. This is useful when YAWS listens to a port which is different from the port that the user connects to. For example, running YAWS as a non-privileged user makes it impossible to listen to port 80, since that port can only be opened by a privileged user. Instead YAWS listens to a high port number port, 8000, and iptables are used to redirect traffic to port 80 to port 8000 (most NAT:ing firewalls will also do this for you).
- `rscheme = http | https` — This forces all local redirects issued by the server to use this method. This is useful when an SSL off-loader, or stunnel, is used in front of YAWS.
- `auth_log = true | false` — Enable or disable the auth log for this virtual server. Default is true.
- `access_log = true | false` — Setting this directive to false turns off traffic logging for this virtual server. The default value is true.

- `logger_mod = Module` — It is possible to set a special module that handles access and auth logging. The default is to log all web server traffic to `<Host>.access` and `<Host>.auth` files in the configured or default `logdir`.

This module must implement the behaviour `yaws_logger`. Default value is `yaws_log`.

The following functions should be exported:

- `Module:open_log(ServerName, Type, LogDir)` — When YAWS is started, this function is called for this virtual server. If the initialization is successful, the function must return `{true, State}` and if an error occurred, it must return `false`.
- `Module:close_log(ServerName, Type, State)` — This function is called for this virtual server when YAWS is stopped.
- `Module:wrap_log(ServerName, Type, State, LogWrapSize)` — This function is used to rotate log files. It is regularly called by YAWS and must return the possibly updated internal `NewState`.
- `Module:write_log(ServerName, Type, State, Infos)` — When it needs to log a message, YAWS will call this function. The parameter `Infos` is `{Ip, Req, InHdrs, OutHdrs, Time}` for an access log and `{Ip, Path, Item}` for an auth log, where:
 - * `Ip` — IP address of the accessing client (as a tuple).
 - * `Req` — The HTTP method, URI path, and HTTP version of the request (as an `#http_request{}` record).
 - * `InHdrs` — The HTTP headers which were sent from the WWW client (as a `#headers{}` record).
 - * `OutHdrs` — The HTTP headers sent to the WWW client (as a `#outh{}` record).
 - * `Path` — The URI path of the request (as a string).
 - * `Item` — The result of an authentication request. May be `{ok, User}`, `403` or `{401, Realm}`.
 - * `Time` — The time taken to serve the request, in microseconds.

For all of these callbacks, `ServerName` is the virtual server's name, `Type` is the atom `access` or `auth` and `State` is the internal state of the logger.

- `shaper = Module` — Defines a module to control access to this virtual server.

Access can be controlled based on the IP address of the client. It is also possible to throttle HTTP requests based on the client's download rate. This module must implement the behaviour `yaws_shaper`.

There is no such module configured by default.

- `dir_listings = true | true_nozip | false` — Setting this directive to `false` disallows the automatic dir listing feature of YAWS. A status code `403 Forbidden` will be sent. Set to `true_nozip` to avoid the auto-generated `all.zip` entries. Default is `false`.
- `extra_cgi_vars =` — Add additional CGI or FastCGI variables. For example:

```
<extra_cgi_vars dir='/path/to/some/scripts'>
  var = val
  ...
</extra_cgi_vars>
```

- `statistics = true | false` — Turns on/off statistics gathering for a virtual server. Default is `false`.
- `fcgi_app_server = Host:Port` — The hostname and TCP port number of a FastCGI application server. The TCP port number is not optional. There is no default value.
- `fcgi_trace_protocol = true | false` — Enable or disable tracing of FastCGI protocol messages as info log messages. Disabled by default.
- `fcgi_log_app_error = true | false` — Enable or disable logging of application error messages (output to `stderr` and non-zero exit value). Disabled by default.
- `deflate = true | false` — Turns on or off deflate compression for a server. Default is `false`.
- `<deflate> ... </deflate>` — This begins and ends the deflate compression configuration for this server. The following items are allowed within a matching pair of `<deflate>` and `</deflate>` delimiters.
 - `min_compress_size = nolimit | Integer` — Defines the smallest response size that will be compressed. If `nolimit` is not used, the specified value must be strictly positive. The default value is `nolimit`.
 - `compress_level = none | default | best_compression | best_speed | 0..9` — Defines the compression level to be used. 0 (`none`), gives no compression at all, 1 (`best_speed`) gives best speed and 9 (`best_compression`) gives best compression. The default value is `default`.
 - `window_size = 9..15` — Specifies the zlib compression window size. It should be in the range 9 through 15. Larger values of this parameter result in better compression at the expense of memory usage. The default value is 15.
 - `mem_level = 1..9` — Specifies how much memory should be allocated for the internal compression state. `mem_level=1` uses minimum memory but is slow and reduces compression ratio; `mem_level=9` uses maximum memory for optimal speed. The default value is 8.
 - `strategy = default | filtered | huffman_only` — This parameter is used to tune the compression algorithm. See `zlib(3erl)` for more details on the `strategy` parameter. The default value is `default`.
 - `use_gzip_static = true | false` — If true, YAWS will try to serve precompressed versions of static files. It will look for precompressed files in the same location as original files that end in `".gz"`. Only files that do not fit in the cache are concerned. The default value is `false`.
 - `mime_types = ListOfTypes | defaults | all` — Restricts the deflate compression to particular MIME types. The special value `all` enable it for all types (It is a synonym of `/*/*`). MIME types into `ListOfTypes` must have the form `'type/subtype'` or `'type/*'` (indicating all subtypes of that type). Here is an example:

```
mime_types = default image/*
mime_types = application/xml application/xhtml+xml application/rss+xml
```

By default, the following MIME types are compressed (if `deflate` is set to `true`):

```
* text/*
* application/rtf
* application/msword
* application/pdf
```



```
* application/x-dvi
* application/javascript
* application/x-javascript
```

Multiple `mime_types` directive can be used.

- `docroot = Directory ...` — This makes the server serve all its content from `Directory`.
It is possible to pass a space-separated list of directories as `docroot`. If this is the case, the various directories will be searched in order for the requested file. This also works with the `ssi` and `yssi` constructs where the full list of directories will be searched for files to `ssi/yssi` include. Multiple `docroot` directives can be used. You need at least one valid `docroot`, other invalid `docroots` are skipped with their associated `auth` structures.
- `auth_skip_docroot = true | false` — If true, the `docroot` will not be searched for `.yaws_auth` files. This is useful when the `docroot` is quite large and the time to search it is prohibitive when YAWS starts up. Defaults to false.
- `partial_post_size = Integer` — When a YAWS file receives large POSTs, the amount of data received in each chunk is determined by the this parameter. The default value is 10240.
- `dav = true | false` — Turns on the DAV protocol for this server. The `dav` support in YAWS is highly limited. If `dav` is turned on, `.yaws` processing of `.yaws` pages is turned off. Default is false. Setting it to `nolimit` is potentially dangerous. The socket read timeout is supplied by the `keepalive_timeout` setting. If the read is not done within the timeout, the POST will fail.
- `tilde_expand = true|false` — If this value is set to false YAWS will never do tilde expansion. Tilde expansion takes a URL of the form `http://www.foo.com/~username` and changes it into a request where the `docroot` for that particular request is set to the directory `~username/public_html/`. The default value is false.
- `allowed_scripts = ListOfSuffixes` — The allowed script types for this server. Recognized are `yaws`, `cgi`, `fcgi`, `php`. Default is `allowed_scripts = yaws php cgi fcgi`.
- `tilde_allowed_scripts = ListOfSuffixes` — The allowed script types for this server when executing files in a users `public_html` folder Recognized are `yaws`, `cgi`, `fcgi`, `php`. Default is `tilde_allowed_scripts = (i.e., empty)`.
- `appmods = ListOfModuleNames` — If any the names in `ListOfModuleNames` appear as components in the path for a request, the path request parsing will terminate and that module will be called. There is also an alternate syntax for specifying the `appmods` if we don't want our internal erlang module names to be exposed in the URL paths. We can specify

```
appmods = <Path1, Module1> <Path2, Modules2> ...
```

Assume for example that we have the URL `http://www.hyber.org/myapp/foo/bar?user=joe` while we have the module `foo` defined as an `appmod`, the function `foo:out(Arg)` will be invoked instead of searching the filesystems below the point `foo`.

The `Arg` argument will have the missing path part supplied in its `appmoddata` field.

It is also possible to exclude certain directories from `appmod` processing. This is particularly interesting for `'/'` `appmods`. Here is an example:

```
appmods = </, myapp exclude_paths icons js top/static>
```

The above configuration will invoke the `myapp erlang` module on everything except any file found in directories `icons`, `js` and `top/static` relative to the docroot.

- `errormod_404 = Module` — It is possible to set a special module that handles 404 Not Found messages. The function

`Module:out404(Arg, GC, SC)` will be invoked. The arguments are

- `Arg` — a `#arg{}` record
- `GC` — a `#gconf{}` record (defined in `yaws.hrl`)
- `SC` — a `#sconf{}` record (defined in `yaws.hrl`)

The function can and must do the same things that a normal `out/1` does.

- `errormod_401 = Module` — It is possible to set a special module that handles 401 Unauthorized messages. This can for example be used to display a login page instead. The function

`Module:out401(Arg, Auth, Realm)` will be invoked. The arguments are

- `Arg` — a `#arg{}` record
- `Auth` — a `#auth{}` record
- `Realm` — a string

The function can and must do the same things that a normal `out/1` does.

- `errormod_crash = Module` — It is possible to set a special module that handles the HTML generation of server crash messages. The default is to display the entire formatted crash message in the browser. This is good for debugging but not in production.

The function `Module:crashmsg(Arg, SC, Str)` will be called. The `Str` is the real crash message formatted as a string.

The function must return, `{content,MimeType,Cont}` or `{html, Str}` or `{ehtml, Term}`. That data will be shipped to the client.

- `expires = ListOfExpires` — Controls the setting of the Expires HTTP header and the max-age directive of the Cache-Control HTTP header in server responses for specific MIME types. The expiration date can set to be relative to either the time the source file was last modified, or to the time of the client access. `ListOfExpires` is defined as follows:

```
expires = <MimeType1, access+Seconds> <MimeType2, modify+Seconds> ...
```

These HTTP headers are an instruction to the client about the document's validity and persistence. If cached, the document may be fetched from the cache rather than from the source until this time has passed. After that, the cache copy is considered "expired" and invalid, and a new copy must be obtained from the source. Here is an example:

```
expires = <image/gif, access+2592000> <image/png, access+2592000>
expires = <image/jpeg, access+2592000> <text/css, access+2592000>
```

- `arg_rewrite_mod = Module` — It is possible to install a module that rewrites all the `Arg #arg{}` records at an early stage in the YAWS server. This can be used to do various things such as checking a cookie, rewriting paths etc.

The module `yaws_vdir` can be used in case you want to serve static content that is not located in your docroot. See the example at the bottom of this man page for how to use the `opaque + vdir` elements to instruct the `yaws_vdir` module what paths to rewrite.

- `start_mod = Module` — Defines a user-provided callback module. At startup of the server, `Module:start/1` will be called. The `#sconf{}` record (defined in `yaws.hrl`) will be used as the input argument. This makes it possible for a user application to synchronize the startup with the YAWS server as well as getting hold of user specific configuration data, see the explanation for the `<opaque>` context.
- `revproxy = Prefix Url [intercept_mod Module]` — Make YAWS a reverse proxy. `Prefix` is a path inside our own docroot and `Url` is a URL pointing to a website we want to "mount" under the `Prefix` path. For example:

```
revproxy = /tmp/foo http://yaws.hyber.org
```

This makes the hyber website appear under `/tmp/foo`.

It is possible to have multiple reverse proxies inside the same server by supplying multiple `revproxy` directives.

If the optional keyword `intercept_mod` and `Module` are supplied in the `revproxy` directive, then `Module` indicates a proxy request and response interception module. When the reverse proxy receives a request from a client, it will first pass `#http_request{}` and `#headers{}` records representing the client request to the intercept module's `rewrite_request/2` function. The function must return a 3-tuple of the following form:

```
{ok, #http_request{}, #headers{}}
```

where the record instances can be the original values passed in or new values. The reverse proxy will use these record instances when passing the request to the backend server.

Similarly, when the backend server returns a response, the reverse proxy will call the intercept module's `rewrite_response/2` function, passing it an `#http_response{}` record and a `#headers{}` record. The function must return a 3-tuple of the following form:

```
{ok, #http_response{}, #headers{}}
```

where the record instances can be the original values passed in or new values. The reverse proxy will use these record instances when returning the response to the client.

Any failure within an intercept module function results in HTTP status code 500 (Internal Server Error) being returned to the client.

Intercept modules can use the following functions from the `yaws_api` module to retrieve, set, or delete HTTP headers from `#headers{}` records:

- `set_header/2`, `set_header/3`
- `get_header/2`, `get_header/3`

- delete_header/2

- fwdproxy = true|false — Make YAWS a forward proxy. By enabling this option you can use YAWS as a proxy for outgoing web traffic, typically by configuring the proxy settings in a web-browser to explicitly target YAWS as its proxy server.
- servername = Name — If we're virthosting several servers and want to force a server to match specific Host: headers we can do this with the servername directive. This name doesn't necessarily have to be the same as the name inside <server Name> in certain NAT scenarios. Rarely used feature.
- php_handler — Set handler to interpret .php files. It can be one of the following definitions:

- php_handler = <cgi, Filename> — The name of (and possibly path to) the PHP executable used to interpret PHP scripts (if allowed).
- php_handler = <fcgi, Host:Port> — Use the specified FastCGI server to interpret .php files (if allowed).

YAWS does not start the PHP interpreter in FastCGI mode for you. To run PHP in FastCGI mode, call it with the *-b* option. For example:

```
php5-cgi -b '127.0.0.1:54321'
```

This starts PHP5 in FastCGI mode listening on the local network interface. To make use of this PHP server from YAWS, specify:

```
php_handler = <fcgi, 127.0.0.1:54321>
```

The PHP interpreter needs read access to the files it is to serve. Thus, if you run it in a different security context than YAWS itself, make sure it has access to the .php files.

Please note that anyone who is able to connect to the PHP FastCGI server directly can use it to read any file to which it has read access. You should consider this when setting up a system with several mutually untrusted instances of PHP.

- php_handler = <extern, Module:Function | Node:Module:Function> — Use an external handler, possibly on another node, to interpret .php files (if allowed).

To interpret a .php file, the function Module:Function(Arg) will be invoked (evaluated inside an rpc call if a Node is specified), where Arg is a #arg{} record.

The function must do the same things that a normal out/1 does.

Default value is <cgi, "/usr/bin/php-cgi">.

- phpfcgi = HostPortSpec — *this target is deprecated. use php_handler target in server part instead.*

Use this directive is same as: php_handler = <fcgi, HostPortSpec>.

- <ssl> </ssl> This begins and ends an SSL configuration for this server. It's possible to virthost several SSL servers on the same IP given that they all share the same certificate configuration. In general it is complicated to virthost several SSL servers on the same IP address since the certificate is typically bound to a domainname in the common name part of the certificate. One solution (the only?) to this problem is to have a certificate with multiple subjectAltNames. See http://wiki.cacert.org/VhostTaskForce#Interoperability_Test

- keyfile = File — Specifies which file contains the private key for the certificate.

- `certfile = File` — Specifies which file contains the certificate for the server.
- `cacertfile = File` — A file containing trusted certificates to use during client authentication and to use when attempting to build the server certificate chain. The list is also used in the list of acceptable client CAs passed to the client when a certificate is requested.
- `verify = 0 | 1 | 2 | verify_none | verify_peer` — Specifies the level of verification the server does on client certs. 0 means that the server will not ask for a cert (`verify_none`), 1 means that the server will ask the client for a cert but not fail if the client does not supply a client cert (`verify_peer`, `fail_if_no_peer_cert = false`), 2 means that the server requires the client to supply a client cert (`verify_peer`, `fail_if_no_peer_cert = true`).

Setting `verify_none` means that the x509 validation will be skipped (no certificate request is sent to the client), `verify_peer` means that a certificate request is sent to the client (x509 validation is performed).

You might want to use `fail_if_no_peer_cert` in combination with `verify_peer`.

- `fail_if_no_peer_cert = true | false` — If `verify` is set to `verify_peer` and set to `true` the connection will fail if the client does not send a certificate (i.e. an empty certificate). If set to `false` the server will fail only if an invalid certificate is supplied (an empty certificate is considered valid).
- `depth = Int` — Specifies the depth of certificate chains the server is prepared to follow when verifying client certs. For the OTP new ssl implementation it is also used to specify how far the server, i.e. we, shall follow the SSL certificates we present to the clients. Hence, using self signed certs, we typically need to set this to 0.
- `password = String` — If the private key is encrypted on disk, this password is the 3des key to decrypt it.
- `ciphers = String` — This string specifies the SSL cipher string. The syntax of the SSL cipher string is a little horrible sub-language of its own. It is documented in the SSL man page for "ciphers".

- `<redirect> ... </redirect>` — Defines a redirect mapping. The following items are allowed within a matching pair of `<redirect>` and `</redirect>` delimiters.

We can have a series of `Path = URL` or `Path = file`

All accesses to `Path` will be redirected to `URL/Path` or to `scheme:host:port/file/Path` if a file is used. Note that the original path is appended to the redirected url. So if we for example have:

```
<redirect>
  /foo = http://www.mysite.org/zapp
  /bar = /tomato.html
</redirect>
```

Assuming this config resides on a site called `http://abc.com`, We have the following redirects:

```
http://abc.com/foo -> http://www.mysite.org/zapp/foo
```

```
http://abc.com/foo/test -> http://www.mysite.org/zapp/foo/test
```

```
http://abc.com/bar -> http://abc.com/bar
```

```
http://abc.com/bar/x/y/z -> http://abc.com/bar/x/y/z
```

Sometimes we do not want to have the original path appended to the redirected path. To get that behaviour we specify the config with '==' instead of '='.

```
<redirect>
  /foo == http://www.mysite.org/zapp
  /bar = /tomato.html
</redirect>
```

With this configuration, a request for `http://abc.com/foo/x/y/z` simply gets redirected to `http://www.mysite.org/zapp`. This is typically used when we simply want a static redirect at some place in the docroot.

When we specify a file as target for the redirect, the `redir` will be to the current `http(s)` server.

- `<auth> ... </auth>` — Defines an auth structure. The following items are allowed within a matching pair of `<auth>` and `</auth>` delimiters.

- `docroot = Docroot` — If a `Docroot` is defined, this auth structure will be tested only for requests in the specified docroot. No docroot configured means all docroots. If two auth structures are defined, one with a docroot and one with no docroot, the first of both overrides the second one for requests in the configured docroot.
- `dir = Dir` — Makes `Dir` to be controlled by `WWWAuthenticate` headers. In order for a user to have access to `WWWAuthenticate` controlled directory, the user must supply a password. The `Dir` must be specified relative to the docroot. Multiple `dir` can be used. If no `dir` is set, the default value, `"/`", will be used.
- `realm = Realm` — In the directory defined here, the `WWW-Authenticate Realm` is set to this value.
- `authmod = AuthMod` — If an `AuthMod` is defined then `AuthMod:auth(Arg, Auth)` will be called for all access to the directory. The `auth/2` function should return one of: `true`, `false`, `{false, Realm}`, `{appmod, Mod}`. If `{appmod, Mod}` is returned then a call to `Mod:out401(Arg, Auth, Realm)` will be used to deliver the content. If `errormod_401` is defined, the call to `Mod` will be ignored. (`Mod:out(Arg)` is deprecated).

This can, for example, be used to implement cookie authentication. The `auth()` callback would check if a valid cookie header is present, if not it would return `{appmod, ?MODULE}` and the `out401/1` function in the same module would return `{redirect_local, "/login.html"}`.

- `user = User:Password` — Inside this directory, the user `User` has access if the user supplies the password `Password` in the popup dialogue presented by the browser. We can obviously have several of these value inside a single `<auth> </auth>` pair.

The usage of `User:Password` in the actual config file is deprecated as of release 1.51. It is preferred to have the users in a file called `.yaws_auth` in the actual directory. The `.yaws_auth` file has to be file parseable by `file:consult/1`.

Each row of the file must contain terms on the form.

- `pam service = pam-service` — If the item `pam` is part of the auth structure, Yaws will also try to authenticate the user using "pam" using the `pam service` indicated. Usual services are typically found under `/etc/pam.d`. Usual values are "system-auth" etc.

`pam` authentication is performed by an Erlang port program which is typically installed as `suid root` by the Yaws install script.

- `allow = all | ListOfHost` — The `allow` directive affects which hosts can access an area of the server. Access can be controlled by IP address or IP address range. If `all` is specified, then all hosts are allowed access, subject to the configuration of the `deny` and `order` directives. To allow only particular hosts or groups of hosts to access the server, the host can be specified in any of the following formats:

A full IP address

```
allow = 10.1.2.3
allow = 192.168.1.104, 192.168.1.205
```

network/netmask pair

```
allow = 10.1.0.0/255.255.0.0
```

network/nnn CIDR specification

```
allow = 10.1.0.0/16
```

- `deny = all | ListOfHost` — This directive allows access to the server to be restricted based on IP address. The arguments for the `deny` directive are identical to the arguments for the `allow` directive.
- `order = Ordering` — The `order` directive, along with `allow` and `deny` directives, controls a three-pass access control system. The first pass processes either all `allow` or all `deny` directives, as specified by the `order` directive. The second pass parses the rest of the directives (`deny` or `allow`). The third pass applies to all requests which do not match either of the first two.

Ordering is one of (Default value is `deny, allow`):

- * `allow, deny` — First, all `allow` directives are evaluated; at least one must match, or the request is rejected. Next, `deny` directives are evaluated. If any matches, the request is rejected. Last, any requests which do not match an `allow` or a `deny` directive are denied by default.
 - * `deny, allow` — First, all `deny` directives are evaluated; if any match, the request is denied unless it also matches an `allow` directive. Any requests which do not match any `allow` or `deny` directives are permitted.
- `<opaque> ... </opaque>` — This begins and ends an opaque configuration context for this server, where `'Key = Value'` directives can be specified. These directives are ignored by YAWS (hence the name `opaque`), but can be accessed as a list of tuples `{Key, Value}` stored in the `#sconf.opaque` record entry. See also the description of the `start_mod` directive.

This mechanism can be used to pass data from a surrounding application into the individual `.yaws` pages.

13.3 Configuration Examples

The following example defines a single server on port 80.

```
logdir = /var/log/yaws
<server www.mydomain.org>
    port = 80
```

```

        listen = 192.168.128.31
        docroot = /var/yaws/www
    </server>

```

And this example shows a similar setup but two web servers on the same IP address:

```

logdir = /var/log/yaws
<server www.mydomain.org>
    port = 80
    listen = 192.168.128.31
    docroot = /var/yaws/www
</server>

<server www.funky.org>
    port = 80
    listen = 192.168.128.31
    docroot = /var/yaws/www_funky_org
</server>

```

When there are several virtual hosts defined for the same IP number and port, and an HTTP request arrives with a `Host` field that does not match any defined virtual host, then the one which defined “first” in the file is chosen.

An example with `www-authenticate` and no access logging at all.

```

logdir = /var/log/yaws
<server www.mydomain.org>
    port = 80
    listen = 192.168.128.31
    docroot = /var/yaws/www
    access_log = false
    <auth>
        dir = /var/yaws/www/secret
        realm = foobar
        user = jonny:verysecretpwd
        user = benny:thequestion
        user = ronny:havinganamethatendswithy
    </auth>

</server>

```

And finally a slightly more complex example with two servers on the same IP, and one SSL server on a different IP.

The `is_default` is used to select the funky server if someone types in for example `urlhttp://192.168.128.31/` in his/her browser.


```
logdir = /var/log/yaws
max_num_cached_files = 8000
max_num_cached_bytes = 6000000

<server www.mydomain.org>
    port = 80
    listen = 192.168.128.31
    docroot = /var/yaws/www
</server>

<server www.funky.org>
    port = 80
    is_default = true
    listen = 192.168.128.31
    docroot = /var/yaws/www_funky_org
</server>

<server www.funky.org>
    port = 443
    listen = 192.168.128.32
    docroot = /var/yaws/www_funky_org
    <ssl>
        keyfile = /etc/funky.key
        certfile = /etc/funky.cert
        password = gazonk
    </ssl>
</server>
```

Chapter 14

WebSocket Protocol Support

YAWS supports the WebSocket Protocol (RFC 6455), which enables two-way communication between clients and web servers. YAWS also provides support for working drafts of the WebSocket protocol, specifically drafts 10 to 17 of the hybi working group. No support for other drafts, such as those from the hixie working group, is provided.

You can find example usage of the WebSocket Protocol in the file `www/websockets_example.yaws`. This example, intended for use with any browser supporting RFC 6455, returns HTML and JavaScript that allow the client to establish a WebSocket connection to the server. These connections are handled by the code in `www/websockets_example_endpoint.yaws`, which when invoked simply establishes `examples/src/basic_echo_callback.erl` as the WebSocket callback module for the connection.

14.1 WebSocket Callback Modules

A WebSocket callback module implements either the `handle_message/1` callback function or the `handle_message/2` callback function, depending on whether it's a basic or advanced callback module.

14.1.1 Basic Callback Modules

The argument passed to `handle_message/1` callback function takes one of the following forms:

- `{text, Text}` — the callback receives an unfragmented text message.
- `{binary, Message}` — the callback receives an unfragmented binary message.
- `{close, Status, Reason}` — the callback receives a notification that the client has closed the socket. `Status` is the numerical status code sent by the client or the value 1000 (as suggested by RFC 6455 section 7.4.1) if the client sent no status code. For an abnormal client closure, the status code is 1006 (as specified by RFC 6455 section 7.1.5). `Reason` is a binary containing any text the client sent to indicate the reason for closing the socket; this binary may be empty.

The `handle_message/1` callback function supplies one of the following as a return value:

- `noreply` — do nothing, just wait for the next message.

- `{reply, {Type, Data}}` — reply to the message. `Type` must be either `text` or `binary` to indicate the type of data in the reply message, and `Data` is the reply message itself.
- `{close, Reason}` — close the connection and exit the handling process with `Reason`. For a regular non-error close, `Reason` should be the atom `normal`. `Reason` may alternatively be any allowed numerical value specified for `close` frames in section 7.4 of RFC 6455. Callback handlers for `close` messages from the client must always return `{close, Reason}`.

To inform YAWS of the details of the callback module, return `{websocket, CallbackModule, Options}` from your `out/1` function, where `CallbackModule` is the name of your callback module and `Options` is a list of options. The following options are available:

- `{callback, CallbackType}` — supply this atom to indicate the type of the callback module. `CallbackType` can be either of the following:
 - `basic` — specify this to indicate your callback module is the basic type. This is the default.
 - `{advanced, InitialState}` — specify this to indicate your callback module is an advanced callback module. Here, `InitialState` is the callback's initial state for handling this client. See 14.1.2 for more details.
- `{origin, Origin}` — specify the `Origin` URL from which messages will be accepted. This is useful for protecting against cross-site attacks. This option defaults to `any`, meaning calls will be accepted from any origin.

14.1.2 Advanced Callback Modules

Advanced callback modules—those that want to supply their own initial state and are prepared to handle fragmented messages themselves—supply a `handle_message/2` callback function.

To indicate an advanced callback module, include `{callback, {advanced, InitialState}}` in the `Options` list when you return `{websocket, CallbackModule, Options}` from your `out/1` function, as described above.

The arguments to the `handle_message/2` callback function are as follows:

- `#ws_frame_info` — this record, defined in `include/yaws_api.hrl`, provides all details of a frame section. See section 5 of RFC 6455 for details.
- `State` — this is the callback module's current state. The initial state is supplied when you return `{callback, {advanced, InitialState}}` as part of the options list you returned from your `out/1` function to establish the WebSocket callback module.

The return values for the `handle_message/2` callback function can be any of the following:

- `{noreply, State}` — do nothing, just wait for the next message. `State` is the (possibly updated) state for the callback module.
- `{reply, Reply, State}` — reply to the received message with `Reply`, which is either `{text, Data}` or `{binary, Data}`. `State` is the (possibly updated) state for the callback module.

- `{close, Reason}` — close the connection and exit the handling process with `Reason`. For a regular non-error close, `Reason` should be the atom `normal`. `Reason` may alternatively be any allowed numerical value specified for `close` frames in section 7.4 of RFC 6455. Callback handlers for `close` frames from the client must always return `{close, Reason}`.