Using wkhtmltopdf and an Xvfb daemon to render HTML to PDF

WEBKIT   PDF   RENDER   XVFB

I have a lot of projects that consist of collecting information and then rendering it to PDF, initially I used tcpdf, but the rendering is VERY finicky. I was doing some additional research a few weeks ago for a personal project and found wkhtmltopdf.

I was scraping my Harvest time sheets to gather my hours for the previous two weeks, rendering to a PDF, and automatically creating a RightSignature document to be signed by my employers. If you've ever had to do this, it can be somewhat of a pain to do repeatedly, especially when you're really busy. Plus, automating things is awesome.

For that process, I used some python (with Requests), wkhtmltopdf, and Xvfb's (X virtual framebuffer) xvfb-run command. This is fine, because I run the process manually, at will.

I have some other projects that require repeated rendering on a live site, for multiple users. Running individual xvfb-run processes for each request is just a dumb idea. Your server is going to explode if you do this. Don't do it.

In order to use a single Xvfb process, I needed to set it up as a daemon that is run by the www-data user, so that the PHP wrapper I am using (PHPWkHtmlToPDF) can access the virtual frame buffer, and use the webkit rendering engine (Chrome) to save it to PDF.

I could (and probably will) setup a processing queue using rabbitmq or the likes, but for now (and for this post), this is fine.

Let's get started, shall we?

These instructions are centered around Ubuntu, so your milage may vary depending on the setup you are using.

Step 1: Install wkhtmltopdf and xvfb

```
# apt-get install wkhtmltopdf xvfb
```

Step 2: Create and enable the init script for the xvfb daemon

Note ":0″ in XVFBARGS which sets to display 0″, and the –set-guid parameter, which runs the daemon as the www-data user. If you are not running a headless server, make sure to change all references to ":0″ to ":#" (ie: :1, :2 or :22) to avoid conflicts with your existing X session(s)

Put the following in "/etc/init.d/xvfb":

```
XVFB=/usr/bin/Xvfb
XVFBARGS=":0 -screen 0 1024x768x24 -ac +extension GLX +render -noreset"
PIDFILE=/var/run/xvfb.pid
```

```
case "$1" in
  start)
    echo -n "Starting virtual X frame buffer: Xvfb"
    start-stop-daemon --chuid www-data --start --quiet --pidfile $PIDFILE --make-pidfile --
    echo "."
    ;;
  stop)
    echo -n "Stopping virtual X frame buffer: Xvfb"
    start-stop-daemon --chuid www-data --stop --quiet --pidfile $PIDFILE
    echo "."
    ;;
  restart)
    $0 stop
    $0 start
    ;;
  *)
      echo "Usage: /etc/init.d/xvfb {start|stop|restart}"
      exit 1
esac

exit 0
```

Enable your init script:

```
# update-rc.d xvfb defaults 10
```

Run your init script:

```
# /etc/init.d/xvfb start
Starting virtual X frame buffer: Xvfb.
```

Check to confirm your Xvfb is running:

```
# ps auxU www-data | grep [X]vfb
www-data 17852  0.0  0.2  54960  7904 ?        S    04:31   0:00 /usr/bin/Xvfb ...
```

Great! The daemon is up and running (hopefully).

Step 3: Modify the PHPWkHtmlToPdf Wrapper

We need to modify the wrapper's call to wkhtmltopdf so it knows we are using a specific X session (see the notes about ":0″ above)

Open up the WkHtmlToPdf.php class/file and find the "getCommand" function (around line #252 at time of writing this post) and add a line to the beginning of the command variable that exports the DISPLAY variable to match the display number you used. In my case it's ":0″, since I am running a headless server.

```
/**
 * @param string $filename the filename of the output file
 * @return string the wkhtmltopdf command string
 */
public function getCommand($filename)
```

```
    {
        $command = 'export DISPLAY=":0";';
        $command .= $this->enableEscaping ? escapeshellarg($this->bin) : $this->bin;

        $command .= $this->renderOptions($this->options);

        foreach ($this->objects as $object) {
            $command .= ' ' . $object['input'];
            unset($object['input']);
            $command .= $this->renderOptions($object);
        }

        return $command . ' ' . $filename;
    }
```

Save, and you're ready to….

Step 4: RENDER!

I use twig to fill up some html templates, but for the sake of an example, we'll use PHP to grab the Google homepage and render it to a PDF.

```
include "classes/WkHtmlToPdf.php";

// instantiate the wrapper
$pdf = new WkHtmlToPdf;

// add the content, addPage can take HTML strings, URLs, or filenames.
$pdf->addPage("http://google.com");

// for this example, we'll just use send, which sends the PDF directly to the browser
// if (!$pdf->saveAs('/tmp/google.pdf')): // this saves to a file
if (!$pdf->send()):
    throw new Exception('Could not create PDF: ' . $pdf->getError());
endif;
```

Save this, run it, and you should be presented with well rendered PDF of the Google homepage.

There are a few options to get the render a bit more accurate, but you can see those at http://mikehaertl.github.io/phpwkhtmltopdf/.

Enjoy and happy rendering!

Written by Larry Williamson

🐦 share to say thanks

Sign in or **sign up** to add your response.

8 Responses

cmcdragonkai

I'm using multiple xvfb-run that are not stopped (they run as a daemon). But I'm getting this RANDR missing extension with SlimerJS. Any ideas? Here's the command:

xvfb-run --auto-servernum --server-num=1 --server-args="-ac +extension GLX +extension RANDR +render -noreset" slimerjs ./robot.js

over 1 year ago · Like?

---

larrywilliamson

Are you using a pre-built version? Might consider building from source and making sure that all the requirements are met.

over 1 year ago · Like?

---

mikehaertl

Hey Larry,

thanks for commending my PHPWkHtml class. We're actually discussing native xvfb support. Maybe you want to join: https://github.com/mikehaertl/phpwkhtmltopdf/pull/24

over 1 year ago · Like?

---

rohitsinghal4u

I am trying to convert html to pdf using wkhtmltopdf but I am getting some funny output. When I am using through command prompt, the fonts are coming correctly but I am trying to execute that command in php using system, shell_exec and exec, its not taking correct font. What can be the issue?? One more thing I want to do is that I want to cut the pdf from the top of 20px something then make the pdf. Please help me if its possible to get by using wkhtmltopdf. I have tried many other converters which converts html to pdf but only wkhtmltopdf is giving the correct output. So please help me in only wkhtmltopdf. I can give you the link so that you can first try yourself first before reply.

My php script is the following, tell me if you can modify here itself.

$url = "http://raindrops.in/ainvvy/view/527a727a4251df44558b4567";

$blah = shellexec("xvfb-run -a -s '-screen 0 1024x768x30' wkhtmltopdf -B 0 -L 0 -R 0 -T 0 --page-height 1147.8px --page-width 1800px --dpi 300 --encoding utf-8 ".$url." /tmp/test.pdf");
$a = fileget_contents("/tmp/test.pdf");
unlink("/tmp/test.pdf");
header('Content-Type: application/pdf');
header('Cache-Control: public, must-revalidate, max-age=0'); // HTTP/1.1
header('Pragma: public');
header('Expires: Sat, 26 Jul 1997 05:00:00 GMT'); // Date in the past
header('Last-Modified: '.gmdate('D, d M Y H:i:s').' GMT');
header('Content-Length: '.strlen($a));
header('Content-Disposition: inline; filename="/tmp/test.pdf";');
echo $a;
die($a);
?>

over 1 year ago · Like?

---

larrywilliamson

@rohitsinghal4u are you using web fonts or system fonts? you have to make sure the font is installed and available, or use a webfont (my suggestion)

I'm not really sure what you are talking about with the "20px" part :P

over 1 year ago · Like?

---

mikehaertl

Since version 1.2.0 of PHPWkHtmlToPdf you don't have to modify the class file anymore.

See here: https://github.com/mikehaertl/phpwkhtmltopdf#using-the-dynamically-linked-binary

over 1 year ago · Like?

---

lecstor

Hi guys,

would love to hear your thoughts on pdfunicorn.com and what I would need to do to make it more usable for you. It's a fairly minimal API at the moment but I'm planning more features as I get a roundtuit.. It is not a HTML to PDF converter, but rather a template/API system built around the concept of pages and repeated headers and footers and based on a grid-like system.. any and all comments welcome..

jason@pdfunicorn.com

over 1 year ago · Like?

awieia

Hi guys,

Using wkhtmltopdf from Ubuntu 14.04 terminal was working but Im trying the PHP wrapper but I get this error : Fatal error: Uncaught exception 'Exception' with message 'Could not create PDF: The switch --no-outline, is not support using unpatched qt, and will be ignored.The switch --disable-smart-shrinking, is not support using unpatched qt.
Any solution?

Thanks

Jeff

over 1 year ago · Like?