PDF tips and tricks

Tips and tricks when working with the PDF[1] file format

1 Decrypting a file

When your PDF file is password-protected, you can use xpdf[2] to open it. Still it will not save it without the password which is annoying.

To do this, the solution I have found is to convert the file to postscript[3] and then back to pdf. Note however that you may loose some information in the process.

You will need poppler[4] (for `pdftops`) and ghostscript[5] (`ps2pdf`) for this to work. Assume you want to decrypt the file `secret.pdf` which has the password `foobar`, then run:

```
$ pdftops -opw foobar secret.pdf
$ ps2pdf secret.ps && rm secret.ps
```

2 Merging several files

To concatenate several files into just one, just run:

```
$ gs -dNOPAUSE -sDEVICE=pdfwrite -sOUTPUTFILE=out.pdf -dBATCH $file1 $file2 ...
```

Where `gs` is provided by ghostscript[6].

3 Extracting only some pages

To extract only from page `$start` to page `$end`, of your document `$input`, and create a new pdf `$output` with those pages, run:

```
$ gs -sDEVICE=pdfwrite -dNOPAUSE -dBATCH -dSAFER \
      -dFirstPage=$start -dLastPage=$end \
   -sOutputFile=$output $input
```

My source for this trick is this tech tip article from Linux Journal[7].

---

[1] http://en.wikipedia.org/wiki/PDF
[2] http://www.foolabs.com/xpdf/
[3] http://en.wikipedia.org/wiki/PostScript
[4] http://poppler.freedesktop.org/
[5] http://pages.cs.wisc.edu/~ghost/
[6] http://pages.cs.wisc.edu/~ghost/
[7] http://www.linuxjournal.com/content/tech-tip-extract-pages-pdf