

# Linux Commando

Initially a Linux command-line interface blog, it has evolved to cover increasingly more GUI app topics. Instead of just giving you information like some man page, I illustrate their usage in real-life scenarios.

ⓘ ✕

Trusted by millions of students, facult and professionals worldwide. Try now

- Home
- Resources
- Contact Me / Advertising

Search

Saturday, February 23, 2013

## Splitting up is easy for a PDF file

Occasionally, I needed to extract some pages from a multi-page pdf document. Suppose you have a 6-page pdf document named *myoldfile.pdf*. You want to extract into a new pdf file *mynewfile.pdf* containing only pages 1 and 2, 4 and 5 from *myoldfile.pdf*.

I did exactly that using *pdftk*, a command-line tool.

If *pdftk* is not already installed, install it like this on a Debian or Ubuntu-based computer.

```
$ sudo apt-get update
$ sudo apt-get install pdftk
```

Then, to make a new pdf with just pages 1, 2, 4, and 5 from the old pdf, do this:

```
$ pdftk myoldfile.pdf cat 1 2 4 5 output mynewfile.pdf
```

Note that *cat* and *output* are special *pdftk* keywords. *cat* specifies the operation to perform on the input file. *output* signals that what follows is the name of the output pdf file.

You can specify page ranges like this:

```
$ pdftk myoldfile.pdf cat 1-2 4-5 output mynewfile.pdf
```

*pdftk* has a few more tricks in its back pocket. For example, you can specify a *burst* operation to split each page in the input file into a separate output file.

```
$ pdftk myoldfile.pdf burst
```

By default, the output files are named *pg\_0001.pdf*, *pg\_0002.pdf*, etc.

*pdftk* is also capable of merging multiple pdf files into one pdf.

```
$ pdftk pg_0001.pdf pg_0002.pdf pg_0004.pdf pg_0005.pdf output mynewfile.pdf
```

That would merge the files corresponding to the first, second, fourth and fifth pages into a single output pdf.

If you know of another easy way to split up pages from a pdf file, please tell us in a comment. Much appreciated.

Two updates (part 2, part 3) are available for this post.

Posted by Peter Leung

✉

📧

📧

📧

📧

📧

22 comments:

Chillar Anand said...



Followers

infolinks

Adobe

See what's new in Creative Cloud. Now.

Buy now



▲

▼

\$559

\$6,499

\$800

\$630

\$1,490

done

May 28, 2015 at 6:01 AM

Unknown said...

Thank you a lot for sharing this.Besides, I found this [PDF split resource](#), I'm not sure whether it supports Linux?

July 26, 2015 at 7:55 PM

JRCP said...

hi! i'm Jose, from Spain

i have tried the Nazim Aghabayov script, but it's like there is a bug...

i saved the script as cortar.sh, and this is what is shown

cortar.sh: 18: cortar.sh: let: not found

cortar.sh: 20: [: -eq: argument expected

cortar.sh: 40: cortar.sh: let: not found

as far i can know, the message of line 18 is about

let modl=currentp%\$ppd

and the message of line 20 is indeed about \$modl

can anybody see where the bug is, if any?

thanks a lot, guys

October 11, 2015 at 4:33 AM

Anonymous said...

very useful for breaking up pdf books, thanks!

January 11, 2016 at 8:38 AM

monarch a sadist said...

thanks man helped a lot... i owe u atleast a thaks

June 9, 2016 at 5:31 AM

Unknown said...

Here is the link for Split pdf document. Hope this gives you a start for you file pdf program on rasteredge page <http://www.rasteredge.com/how-to/csharp-imaging/pdf-split/>

July 5, 2016 at 11:29 PM

Anonymous said...

JUST realized that closing the left side pane containing the thumbnails of each page in the PDF allows for the file to scroll 98-99% smoothly.

Stumbled upon the solution as I was printing PDF files with regards to page ranges and chapters in order to split the book up into smaller file sizes, which was working very goooooo too by the way. But simply closing the left side thumb-nails is a lot less work :)

July 28, 2016 at 8:17 PM

Akom said...

I had to write a script to split the original PDF into pages in order to allow tesseract and imagemagick to handle it without running out of memory, and to overcome the TIFF with alpha channel issues (spp not in set {1,3,4})

Script and write-up are here: <http://tech.akom.net/archives/126-OCR-on-a-large-PDF-using-tesseract-and-pdfik.html>

Thanks for the starting point!

January 19, 2017 at 12:59 PM

Unknown said...

Im linux not able to find my pdf do i have to paste in at specific place

April 21, 2017 at 8:19 PM

Robi said...

Found one really easy way to split and merge pdfs here, worked for me

infolinks

https://linuxcommando.blogspot.com/2013/02/splitting-up-is-easy-for-pdf-file.html

3/4

▲

▼

\$559

\$6,499

\$800

\$630

\$1,490

February 9, 2018 at 4:37 AM

Anonymous said...

pdftk seems not be free anymore

December 10, 2018 at 3:27 AM

pranjal said...

So helpful thanks..👍

March 4, 2019 at 5:26 AM

Anonymous said...

please use pdfseparate on linux/ubuntu  
pdfinfo  
pdfunite and others

March 13, 2019 at 10:31 AM

Post a Comment

Newer Post

Home

Older Post

Subscribe to: [Post Comments \(Atom\)](#)