

Manipulating PDF and PostScript Documents

Dipl.-Inf. Frank Hofmann

Berlin

9. Juli 2009



- Manipulating documents because ...
 - file reduction (archiving, backup)
 - printing as a draft (two pages per sheet)
 - a single page is needed
 - combine single documents into one file

- Goal:
 - show, which possibilities for manipulation exist
 - usage of the command line

Overview

- 1 Document Formats – PDF and PostScript
- 2 Tools
- 3 Conversion from PDF to PostScript
- 4 Distributing pages
- 5 Extracting single pages
- 6 Combining pages
- 7 Adjust page layout
- 8 Conclusion

Document Formats – PDF and PostScript

- PostScript
 - page description language, developed in 1972 by Adobe
 - goal: optimized printing on every printing device
 - interpreter is part of the printer firmware
 - controlling the device works perfectly

- Portable Document Format (PDF)
 - page description language, developed in 1990 by Adobe
 - goal: platform independent format
 - no dependencies for a given output device

Conversion Tools

- PostScript
 - psutils
 - poster

- PDF
 - pdf-utils
 - pdjam
 - pdftk
 - okular

All the packages are standard packages in Debian Lenny

Conversion from PostScript to PDF

- PostScript to PDF: ps2pdf*
ps2pdf, ps2pdfwr, ps2pdf12, ps2pdf13, ps2pdf14
- conversion of PostScript to PDF using ghostscript
 - ps2pdf12 creates PDF 1.2 (compatible to Acrobat 3)
 - ps2pdf13 creates PDF 1.3 (compatible to Acrobat 4)
 - ps2pdf14 and ps2pdfwr for PDF 1.4
- Example:
ps2pdf postscript.file pdf.file

Conversion of PDF to PostScript

- PDF to PostScript: `pdf2ps`, `gs`, `acroread`
- `pdf2ps` converts PDF to PostScript using `ghostscript` (`gs`)

```
pdf2ps pdf.file postscript.file
```

- using the Acrobat Reader

```
acroread -toPostScript < pdf.file > postscript.file
```

More than one page per sheet – PostScript (1)

- PostScript: `psnup`, `mpage`

creating `n` pages per sheet:

```
psnup -2 postscript-input.file postscript-output.file
```

Alternatively use:

```
mpage -2 postscript-input.file > postscript-output.file
```


More than one page per sheet – PDF (2)

- PDF: pdfnup

- Two pages per sheet, next to each other:

```
pdfnup original.pdf --nup 2x1 --outfile ausgabe.pdf
```

- Two pages per sheet, above each other:

```
pdfnup original.pdf --nup 1x2 --outfile ausgabe.pdf
```

- Four pages per sheet, in frames:

```
pdfnup original.pdf --nup 2x2 --frame true --outfile  
ausgabe.pdf
```

Extracting single pages – PostScript (1)

- PostScript: `psselect`
 - Printing page 42, only:
`psselect -q -p42 original.ps seite42.ps`
 - Saving all the odd pages as a new file (-o: odd)
`psselect -q -o original.ps ungerade.ps`
 - Saving all the even pages as a new file (-e: even)
`psselect -q -e original.ps gerade.ps`

Extracting single pages – PDF (2)

- PDF: pdftk
 - Extracting the document into single pages:
`pdftk original.pdf burst`
 - Extracting page 2:
`pdftk original.pdf cat 2 seite2.pdf`

Joining two documents

- PostScript: `psmerge`
 - `psmerge -ooutput.ps dokument1.ps dokument2.ps`
- PDF: `pdfjoin`, `pdftk`
 - `pdfjoin dokument1.pdf dokument2.pdf --outfile output.pdf`
 - `pdftk dokument1.pdf dokument2.pdf cat output output.pdf`

Creating a brochure

- PostScript: `psbook`
Sorting the single pages:
`psbook vorlage.ps broschuere.ps`
- PDF:
`pdfbook`, `okular`

Complex example: Creating a brochure

Basic format: A4, portrait, PDF

Goal: A5, PDF, printable as A4, landscape 2 pages/sheet

1 Scaling to A5

```
pdfjoin vorlage.pdf --paper a5paper --fitpaper false  
--outfile vorlage-a5.pdf
```

2 Conversion to PostScript

```
pdftops -paperw 420 -paperh 595 vorlage-a5.pdf bericht-a5.ps
```

3 Conversion into a book format

```
psbook bericht-a5.ps bericht-a5-book.ps
```

4 Conversion as two pages per sheet

```
psnup -pa4 -Pa5 -2 bericht-a5-book.ps > bericht-a5-book-2.ps
```

5 Conversion into PDF

```
ps2pdf14 bericht-a5-book-2.ps broschuere.pdf
```

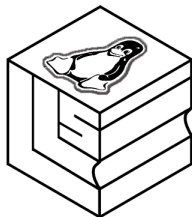
Conclusion

- free and reliable tools for PDF and PostScript manipulation exist
- many options and switches
- tools on the command line – building flexible tool chains is possible
- fulfill the requirements for printing companies

The End

Thanks for your attention :-)

YES, WE ARE OPEN!
büro2.0
Open Source Bürogemeinschaft & Expertennetzwerk



Kontakt:

Dipl.-Inf. Frank Hofmann
Hofmann EDV – Linux, Layout und Satz
c/o büro 2.0
Weigandufer 45
12059 Berlin

Email <frank.hofmann@efho.de>
web www.efho.de