Table Of Contents

Home > Faq > BASH Shell

pdfimages: Extract and Save Images From A Portable Document Format (PDF) File

Posted by Vivek Gite <vivek@nixcraft.com>

Q. How do I extract images from a PDF file under Linux / UNIX shell account?

A. pdfimages works as Portable Document Format (PDF) image extractor under Linux / UNIX operating systems. It saves images from a PDF file as Portable Pixmap (PPM), Portable Bitmap (PBM), or JPEG files. Pdfimages reads the PDF file PDF-file, scans one or more pages, and writes one PPM, PBM, or JPEG file for each image, image-root-nnn.xxx, where nnn is the image number and xxx is the image type (.ppm, .pbm, .jpg).

[1]

pdfimages is installed using poppler-utils package under various Linux distributions:

```
# yum install poppler-utils
```

OR

```
# apt-get install poppler-utils
```

pdfimages syntax

pdfimages /path/to/file.pdf /path/to/output/dir
Extract the PDF file called bar.pdf and save every image as image-00{1,2,3..N}.ppm, enter:

```
$ pdfimages bar.pdf /tmp/image
$ ls /tmp/image*
```

Sample output:

```
image-000.ppm     image-1025.ppm    image-1140.ppm    image-1256.ppm    image-247.ppm    image-374.pp
image-001.ppm     image-1026.ppm    image-1141.ppm    image-1257.ppm    image-248.ppm    image-375.pp
image-002.ppm     image-1027.ppm    image-1142.ppm    image-1258.ppm    image-249.ppm    image-376.pp
```

Normally, all images are written as PBM (for monochrome images) or PPM (for non-monochrome images) files. With the -j option, images in DCT format are saved as JPEG files. All non-DCT images are saved in PBM/PPM format as usual:

```
$ pdfimages -j bar.pdf /tmp/image
```

The -f option Specifies the first page to scan. To scan first 5 pages, enter:

```
$ pdfimages -j -f 5 bar.pdf /tmp/image
```

The -l option specifies the last page to scan. To scan last 5 pages, enter:

```
$ pdfimages -j -l 5 bar.pdf /tmp/image
```

4000+ howtos and counting! Want to read more Linux / UNIX howtos, tips and tricks? Subscribe to our daily email newsletter or weekly newsletter to make sure you don't miss a single tip/tricks. Alternatively, subscribe via RSS/XML feed.

URL to article: **http://www.cyberciti.biz/faq/easily-extract-images-from-pdf-file/**

URLs in this post:

[1] Image: **http://www.cyberciti.biz/faq/category/linux/**