docfs
A filesystem to provide documentation in many different formats

Winfried Trümper <winni@xpilot.org>

22. Mai 1997

Common document formats

- SGML-derivatives: HTML, linuxdoc-sgml
- Classical formats: man, info, postscript
- Commercial formats: Word, StarOffice, PDF

Todays inconveniences

Competition among formats doesn't result in enhanced features.

- Different viewers for different formats
- No references accross formats
- No global search possible
- Supporting only one format annoys people
- Supporting all formats wastes disk-space

Ideal document maintenance

- Providing every format (with some restrictions) or at least incorporate
- Backwards compatible
- Efficient use of disk-space
- Search facility
- Index files for surveys

Providing every format

- On-the-fly conversion to save diskspace
- Compressed storage without problems
- Cache techniques for speed-up

Backwards compatibility

- File-system as common denominator
- Important: `/usr/man` and `/usr/info`
- Less important: current `/usr/doc`

Searching and Surveys

- Survey results in classification of software
- Indexing documents is not trivial (need improved Metadata)

Suggestion for an implemenation

The kernel-module *userfs*

- is a mechanism by which normal user processes can be a Linux filesystem
- provides the tools and know-how to implement docfs
- is written by Jeremy Fitzhardinge <jeremy@sw.oz.au>
- can be fetched from
  `ftp://tsx-11.mit.edu/pub/linux/ALPHA/userfs`

Real world example: ftpfs

```
muserfs  ftpfs   /FTPfs      /var/cache/ftpfs
   #prg  module  mountpoint  cache-dir
cd /ftp/ftp.kernel.org/pub/linux
dir
total 0
drwxrwxrwx  2 root.root  2 Apr 13 23:58 daemons
drwxrwxrwx  2 root.root  2 Apr 13 23:58 distributions
drwxrwxrwx  2 root.root  2 Apr 13 23:58 docs
drwxrwxrwx  2 root.root  2 Apr 13 23:58 kernel
```

A block diagram of docfs components

Reasonable conversion between formats

format	man pager	info reader	web browser
man	X	X	X
info	X	X	X
misc files			X

Misc files are incorporated into html instead of conversion (if possible).

Disadvantages of *docfs*

- Man-power for implementation and document classification
- Document classification leads to a software classification
- Stability of *userfs* unclear
- Linux-centric implementation
- Long-Term goal

Advantages of *docfs*

Beside the ones described earlier:

- SGML as the document source of the future
- Dynamic document handling: cgi without server
-

Future of *docfs*

- Sleeping on my harddisk
- Advertising until developer/maintainer found
- Pushing enhanced SGML DTDs: a legal way to improve information handling