

# Increasing compression performance of block based file systems

Ferenc Havasi

In embedded systems the storage device is mostly based on flash chips. The usage of these devices become more and more general - such as mobile phones, digital cameras, mp3 players, PDAs. One of the most expensive part of this machines is the flash chips, so any technique which can make the usage of the flash to be more efficient can be very important in the future.

In the open source environment there is a file system, which is widely used, and designed specially for flash - called JFFS2 (Journaling Flash File System, version 2).

JFFS2 already has a compression feature - it uses ZLIB. The main goal of our research is to improve this compressing performance as much as possible.

First idea is to replace ZLIB with a compressor framework, which is able to use an arbitrary set of compressors. You can tell this framework what is the most important for you: the compression ratio, the read or write speed. For example if you choose the compression ratio, while compressing a block it will try all compressors and will choose the one, which produces the smallest result. It is easy to extend this framework by a new compressor (only needs to implement an interface).

Another relevant limitation of the efficiency of these compressors is that JFFS2 splits the information into 4KB blocks (because of the behaviour of the flash chips) and compresses them individually. The solution to this problem can be to allow the compressors to use a model. It can be for example a dictionary, a decision tree or neural network. These compressors are called model-based compressors, and can collect information about the file structure before making the file system image, store this information into a model file, and use this model information during (de)compressing blocks.

We implemented this technique and achieved relevant speed up and/or size reduction. It will be open source after March 2004. (<http://www.inf.u-szeged.hu/jffs2/>)