FileSystem選択法

奥山　健一

2004/08/29

Linuxで使えるファイルシステム

- ext2/ext3
- ReiserFS
- XFS
- JFS
- VFAT,NTFS,等々も使えるが、起動fsにするには…

ext2/ext3

- 2nd Extended fs/ 3rd Extended fs
  (第2版拡張fs / 第3版拡張fs) の略
- MINIX fs が「機能拡張前の版」
  - 最大64Mbyteのパーティションサイズ
  - 32文字までのファイル名
- Ext は MINIX fs を拡張したもの
  - 最大2Gbyte のパーティション
  - 255文字までのファイル名

ext2/ext3 -続き-

- Ext2:
  - 16Tbyte partition size
  - 4Tbyte file size
  - BSD Fast File Systemを意識した作り
- Ext3:
  - Ext2 に journaling 機能を追加

基本的にこの２つは
同じフォーマットを用いている

ext2/ext3 のメリット

- 実はあまりない
  - BSD のFFSを参考にして作られている
  - 他のFSはFFSとの比較でメリットがでるようにデザインされているので、自動的にext2/ext3と比較しても効率がよくなる
- …昔から使われているので、Linux関係のツールでの親和性がよい事ぐらい?
- initrdファイルのファイルイメージフォーマットにも使っているけど…

ext2/ext3 の弱点

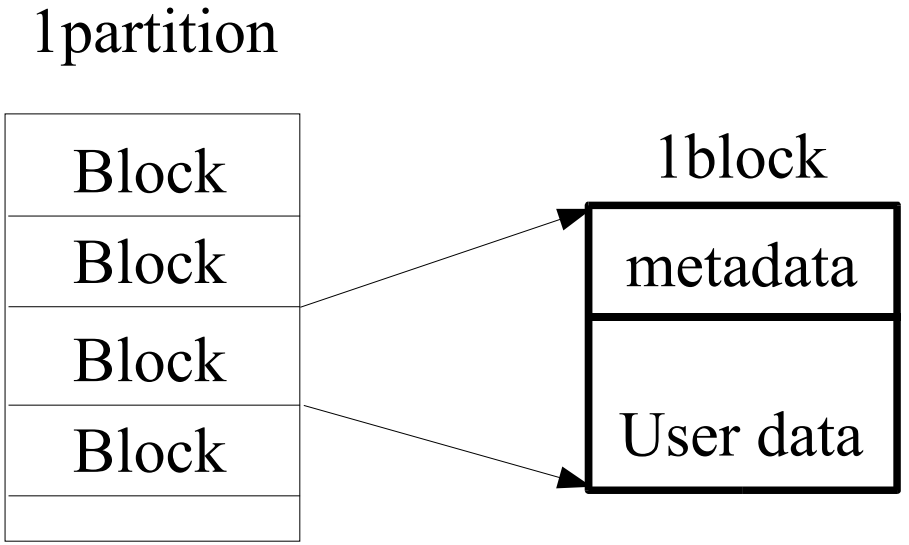
- Inode数が初期化段階で決まってしまう
- 大きなファイルだと確実にフラグメントを起こす
- 小さなファイルだと無駄が多い
- 壊れた場合、ユーザデータが回復しない
  - ext3がjournal機能で被害を限定的にしているがそれでもユーザデータは回復しない

inode数が初期化段階で決まる

- unixのファイルシステムは、inodeでファイルを管理している
  - 1ファイル1 inodeを消費
  - Directoryも1 inodeを消費する
  - Inodeがなくなったらもうファイルは作れない
- Inode数が固定だということは
  - 多く作りすぎると無駄がでる
  - 少なすぎるとdiskが空いてるのに使えない

大きなファイルだとfragmentを起こす

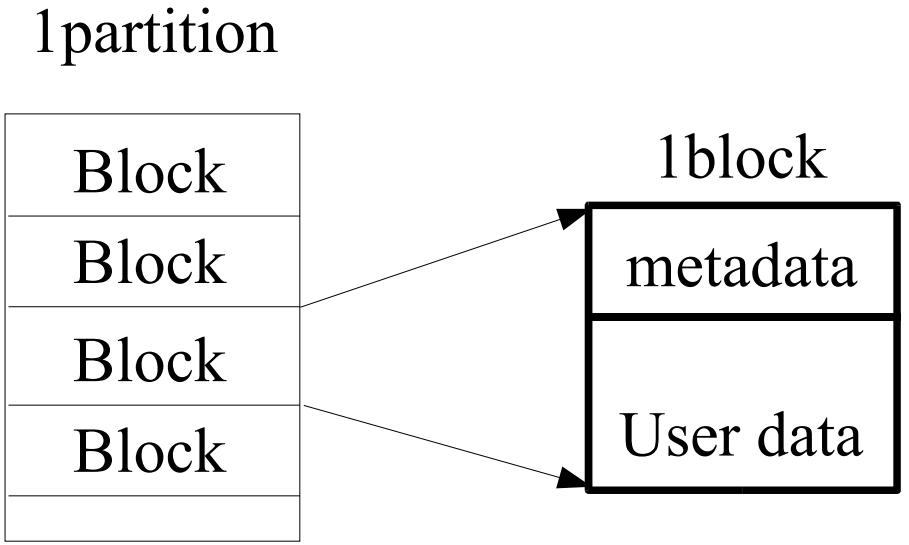
- ext2/3はpartitionを小さなblockに分割、それぞれにmetadata(inodeとかを入れる場所)とuser data(ファイルの中身を入れる場所)に分割
- 基本的に1block内に全部納めようとする



大きなファイルだと fragmentを起こす

- 1blockに収まらないほどデータが大きい場合は隣のblockを使い始める
- 当然その間にはmetadata領域がある

ここでfragmentが確実に起こる



小さなファイルだと無駄が多い

- ext2/3はどんなファイルでも、4kbyte単位でdiskを消費していく

壊れた場合
ユーザデータが回復しない

- ファイルシステムがHDDにデータを書くときに、一定の規則を守る必要がある
- 守らないと、一瞬壊れた状態になる
- 壊れた状態になった瞬間に電源が落ちると回復できない
  - Ext3はjournal機能があるのである程度回復する
  - が、ユーザデータは回復しない

ReiserFS

- Hans Raiser博士による新しいFS
- 小さなファイルに対する扱いが極めて上手い disk利用効率が高い
- inodeを使わないので、inode不足にならない。Inodeが余る事もない
- Atomic Writeをサポート(ユーザデータは、書き換えたか、全く書き変わらないかどちらかで、中途半端な状態にならない)
- 大きなファイルに対する性能が悪い(fragmentを起こしやすい)

XFS

- SGIがCGデータ用に作り上げたファイルシステム
- 大きなファイルの扱いが上手い(fragmentを起こしにくい)
- 小さなファイルに対するdisk利用効率は悪い
- inodeの数を自動調整する
- ユーザデータを更新中に電源が落ちるとユーザデータはほぼ確実に壊れる

JFS

- IBMがOS/2用を作り替えたファイルシステム

性質:中庸

- ReiserFSほど小さいファイルの扱いが上手くない/XFSより上手
- XFSほど大きいファイルの扱いが上手くない/ReiserFSより上手
- Inodeの数は自動調整
- ユーザデータを更新中に電源が落ちるとユーザデータはほぼ確実に壊れる

どのFSを使うべきか?

	RaiserFS	XFS	JFS	Ext3
NotePC	◎	△	○	×
Desktop	○	◎	○	×
Server	△	◎	○	×
Database	×	◎	△	×

欲しいfsが使えるとは限らない