# NILFS: 連続スナップショッティングを実現したファイルシステム
# NILFS: Filesystem gains a continuous snapshotting

天海良治† 森田和孝† 小西隆介†
佐藤孝治† 木原誠司† 盛合敏†

## 1. はじめに

NILFS[1)~3)](the New Implementation of a Log-Structured File System) は Linux Kernel 2.6 用のローカルファイルシステムで，ログ構造化ファイルシステム[4)] (Log Structured FS: LFS) を現代の技術で実装したものである．現在第 2 版を GPL で公開している☆．LFS はディスクブロックを上書きせず，データをログのように追記していく．上書きがないので，ファイルシステムの首尾一貫性が常に保たれる．さらに，ファイルの書出しや変更の履歴がすべて残る．つまり，連続的かつ自動的にファイルシステムのスナップショット（ファイルシステムのある瞬間の状態）が残っていく．これにより，ファイルの配置は，現在時間のファイルを保持するディレクトリツリーに加え，過去を遡る時間軸を獲得した．

利用者の時間軸方向のナビゲーションのため，NILFS はカメラ画像や種々のセンサ情報など，時刻のタグといえる情報を保持し管理する機能を備えている．Linux カーネルの奥底でユーザデータを守っていたファイルシステムは，いまや実世界とのインタラクション能力を得て，積極的に利用者を支援する役割を果たす．

## 2. 大容量ディスクの有効利用

ファイルシステムには，信頼性，可用性，耐故障性などの確実にデータを保持する基本的機能がまず求められ，その上で性能や運用の容易さといった機能が考慮される．

しかし，これらは過去の比較的小さなディスクドライブを利用していたときから重要であり続けた性質である．最近のテラバイト級の大容量ドライブを従来のファイルシステムで利用したなら，ディスクの容量の増加は単に保持できるデータの量が増えるだけで，利用者に利便性の質的向上をもたらさない．例えば，未使用領域が大量に残っているのに，今消してしまったファイルの復活ができないファイルシステムが現在でも常用されている．ディスク使用効率のため，消去されたファイルが占めていたディスク領域を直ちに再利用しているためだ．他にも，使用効率のために，ユーザデータ以外の情報はファイルのブロック管理用のデータだけといったシステムが一般的である．この点でも，大容量ドライブを活かしてメタデータをより多く保持し，質的向上を狙うアプローチが考えられてよい．

## 3. NILFS の概要

Linux ファイルシステムのさらなる信頼性，可用性の確保，そして，何より大容量ドライブの恩恵を利用者の使いやすさやオペレータの操作心理負担の軽減に活かすため，われわれは LFS に注目した．

NILFS は，LFS の性質を活かしてディスク容量の全体を使ったオーバーヘッドのない連続スナップショッティングを実現した．取得したスナップショット上のファイルにもオーバーヘッドなくアクセスできる．これにより，利用者が誤操作をしたり，設定ミスなどでファイルを失った場合，事前のバックアップ作成や明示的なスナップショット取得指示をしていなくても誤操作前の首尾一貫したデータが取り戻せる．また，すべての過去を保持しているとディスク容量が枯渇するので，NILFS 第 2 版では過去のデータを選別して消去するクリーナ機能を実装した．

ここで用語を整理する．NILFS では，ファイルブロックの書出しやディレクトリ操作の区切りでファイルシステムが首尾一貫した状態となったときチェック

† 日本電信電話株式会社 NTT サイバースペース研究所
NTT Cyber Space Laboratories, NTT Corporation
☆ URL: http://www.nilfs.org/ (ダウンロードもここから)

ポイント（以下 CP）をディスクに書出し，そのときのファイルシステムの世代を確定する．CP は過去の状態をアクセスするポイントだが，利用者は CP から直接過去をアクセスすることはできない．CP を NILFS のスナップショット（以下 SS）として印付けすると，SS をファイルシステムとして読出し専用でマウントできるようになる．CP に属するディスクブロックは，のちにクリーナによって領域を回収，再利用されるが，SS は回収されない．当然であるが，ある CP がクリーナによって回収されても，他の SS や現在時刻のファイルシステムからディスクブロックが共有されているときは，そのブロックは回収されない．途中の世代を回収してもその前の世代が SS ならアクセス可能なまま残る．SS は任意個数作成することができる．SS を CP に戻すことも可能である．

## 4. NILFS スケッチ

ファイルの配置が時間軸方向へ広がったが，利用者にとって過去のバージョンのファイルが容易に見つからなければ無意味だ．NILFS は利用者の過去へのナビゲーションのため，時間軸へのタグ付け機能を備えている．NILFS スケッチと名付けたこの機能は，次の特徴がある．1. CP にユーザデータ（スケッチデータ）を保持する．2. スケッチデータのアクセスメソッドはファイルシステムが提供する．3. スケッチデータの内容についてはファイルシステムは関知しない．

スケッチデータとして考えられるのは，デスクトップ画面のサムネイル，作業時のカメラ画像，その時のニュースのヘッドライン，予定表の写し，閲覧していた WEB ページの RSS データ，センサ情報など，その時に書出したファイル内容や作業を人間が思い出すヒントになるような時刻に結び付いた情報である．どんな情報を保持すれば過去を思い出すきっかけになるかは，ファイルシステムとは別の次元で議論されるべき話題であるが，NILFS はシステムとして，タグ情報を保持し，維持するメカニズムを提供する．

## 5. デスクトップ検索との連係

ファイル検索の効率化に大きく貢献しているいわゆるデスクトップ検索技術は，過去のファイルの探索に自然に拡張できる．検索のためのインデックスを，過去のファイルについても保持すること，検索できたファイルを実際に取り出すことの 2 つの要素で，過去の検索が可能である．消してしまったファイルはもちろん，ファイルを編集して消したり上書きした文章からも検索可能となる．スケッチデータも検索対象として，スケッチの時刻タグと AND 検索することで，目的の過去のファイルをすばやく捜し出すことも考えられる．ただし，クリーナによって消去された CP に含まれるインデックスは消去しなくてはいけない．現在，インデックス管理とクリーナとの連携部分を実装中である．

## 6. 外界とのインタラクション

NILFS で自動的に取得されるファイルシステムの世代のうち，SS として残したい世代とは，例えば完成品の一式，一日の終りのもの，といろいろ考えられるが，ぜひほしいのは，なんらかのミスをしたその直前の世代である．CP は大量に作成されるので，SS 化指示にもシステムサポートを提供している．例えば，マイクで利用者の音声をひろい，大声を出したら，その直前の CP を SS 化する．「しまった!!」といった言葉を認識できれば，さらによい．席を立ったら，暗くなったら，コーヒーを飲んでいたら，と，なんらかの事故やミスが起きそうなときに SS 化するのが有望であろう．なお，クリーナは設定された一定時間を経過した CP だけを回収の対象とする．失敗に気づいたらあわてずに SS 化し，その直前の状態を復帰すればよい．

最近は種々のセンサが安価に入手できる．ファイルシステムもカーネルの奥底で地味にデータを片付けているだけでなく，センサを通して外界とインタラクションをもち，積極的に利用者を支援すべきだ．NILFS はそういう人に優しいファイルシステムである．

## 7. 終 り に

NILFS はオープンソースで公開している開発中のソフトウェアである．皆様のコメントや貢献を歓迎する．

## 参 考 文 献

1) 天海良治ほか．Linux 用ログ構造化ファイルシステム nilfs の設計と実装．情報処理学会研究報告 2005-OS-99, pp. 61–68, 2005.
2) Ryusuke Konishi, et al. The Linux implementation of a log-structured file system. *SIGOPS Operating System Review*, Vol. 40, No. 3, pp. 102–107, 2006.
3) 久保類ほか．NILFS: 世界を忘れないファイルシステム．第 48 回プログラミング・シンポジウム報告集, pp. 119–126. プログラミング・シンポジウム委員会, 1 2007.
4) Mendel Rosenblum and John K. Ousterhout. The design and implementation of a log-structured file system. *ACM Transactions on Computer Systems*, Vol. 10, No. 1, pp. 26–52, 1992.