

インタビュー / NTTサイバースペース研究所

いつでも過去の時点が復元できるLinux用の
新ファイルシステム「NILFS」を開発・公開

はじめに、NTTサイバースペース研究所におけるオープンソース・ソフトウェア（OSS）の取組みについてお話ししてください。

南部 NTTサイバースペース研究所では、OSSを積極的に活用し、企業の基幹となるシステムへの適用に向けて、LinuxなどのOS（オペレーティングシステム）やデータベース・マネージメント・システムなどのミドルウェアの信頼性、機能、性能の強化に取組み、検証、解析、システム構築などのノウハウを蓄積しながら、コミュニティと連携し、OSおよびミドルウェアの開発、保守を推進しています。また2004年には、OSDL（Open Source Development Labs）に加盟し、技術仕様の策定とその実現に向けた研究開発を行っています。

OSDLでの取組みについてお話ししてください。

盛合 OSDLが2005年2月に発表した、企業サーバ向けのLinuxの要求定義仕様「OSDLデータセンタLinux仕様1.1版」に、当研究所が中心となって提案した「高信頼ファイルシステム」の機能要件が採用されました。この機能要件を満たすだけでなく、オペレーションに係るコストを削減するための新たなファイルシステムとして、「NILFS（ニルフス）」を提案し、開発に取り組んでいます。NILFSの全てのソースコードは、OSSのライセンスであるGPL（GNU General Public License）にて、2005年9月26日から公開^{（*1）}されています。

*1：「NILFS」プロジェクトのホームページ：<http://www.nilfs.org/>

NILFSの開発の経緯をお話ししてください。

天海 これまでLinuxでは、コンピュータが異常停止すると、その条件によってはデータの同期書き込みの完



日本電信電話㈱
NTTサイバースペース研究所
所長
南部 明氏



日本電信電話㈱
NTTサイバースペース研究所
OSS コンピューティングプロジェクト
主幹研究員 工学博士
盛合 敏氏



日本電信電話㈱
NTTサイバースペース研究所
OSS コンピューティングプロジェクト
主任研究員
天海 良治氏

了が保証されず、データを喪失する恐れがありました。データが喪失した場合、停止直前の状態に復元し、システムを再開させるためには、多大な労力と時間が必要でした。今回開発したNILFSは、Linuxを利用したシステムにおいて、このような労力と時間をなくすことを目的としています。ファイルの世代管理が必要な企業内の業務システムなどへ適用されることにより、さらなるLinuxの利用拡大が期待されます。

従来は、どのようなファイルシステムが使用されていたのでしょうか。

盛合 5年程前までは、EXT2というファイルシステムが主流でしたが、その後、ジャーナリング技術を採用したファイルシステムが登場しました。このファイルシス

ログ構造型のディスクフォーマットを採用

- ・ファイルの管理情報やデータを固定位置に書き込むのではなく、それらの変更部分のみをログとして次々とディスク上に追記
- ・データ損傷の防止、高速な書き込み

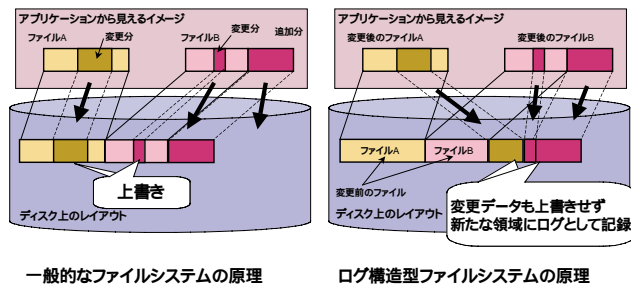


図1 NILFSの特徴

テムは、コンピュータの異常停止に備え、データを別の場所（ジャーナルファイル）へ一旦書き出し、それが完全に完了した後に、改めて本来の場所（ユーザーデータ領域）に書き込む方式を採用しています。これにより、ファイルシステムの整合性がいつでも保たれるようになります。Linuxではカーネル2.4.15からサポートされています。しかし、書き出すデータ量が増加するとともに、ジャーナルファイルとユーザーデータの間でシークが発生し、性能が低下してしまうなど、いくつかの問題点がありました。このジャーナルファイルシステムを超える高い信頼性を実現するファイルシステムとして、デスクブロックの上書きを極力排除したログ構造化ファイルシステム（Log-structured File System）に着目し、新たに開発したものがNILFS（The New Implementation of a Log-structured File System）です。

NILFSにはどのような特徴があるのでしょうか。

天海 記録済みデータへの上書きを行わず、データの追加や変更分をディスク上の空き領域にチェックサム（*2）付きで追記することで、データが破壊される確率を格段に低くしています。加えて、書き込み性能が高いため、サーバのログの出力のような連続して大量に吐き出されるデータの保存に最適です（図1参照）。

一連の書き込みが完了した時点で、時刻情報付きのスナップショット（ファイルシステムの瞬間的な状態）を生成するので、一貫性検査が不要であり、障害時の迅速な復旧が可能です。また、ディスクの容量が許す限り、

- ・一貫性の取れた時点でチェックサムと時刻情報付きのスナップショットを生成
- ・スナップショットへの瞬時リカバリが可能、過去のデータ復元が容易
- ・サービス停止不要の自動的なスナップショット作成（操作不要）

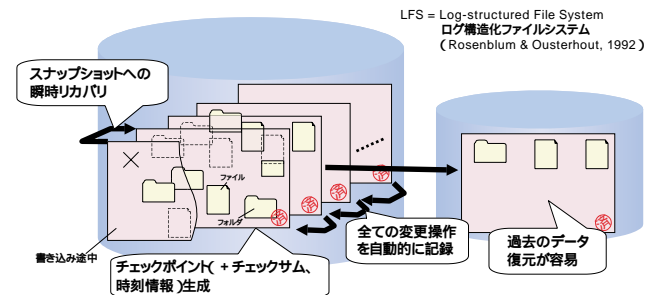


図2 NILFSの特徴

いつでも過去の時点の状態を復元することが可能です（図2参照）。さらに、他のファイルシステムで見られるようなスナップショットを取得するためのサービス停止が不要です。そして、ファイルシステムの管理にB木（*3）を採用することで、大きなファイルや多数のファイルの処理の効率化を実現しています。

*2：データの誤りを検出する方法の一つで、データの全ての値を合計する等の簡単な計算をしたもの。

*3：B-tree。動的に変化する情報を高速に検索する手法の一つ。

NILFSの適用領域と導入のメリットについてお話ししてください。

盛合 適用領域は主に中小規模エンタープライズサーバですが、デスクトップLinuxでの利用も可能です。導入のメリットとしては、高価な商用製品が不要になることとストレージ障害の頻度を格段に低下させることがあげられます。さらに、バックアップやリカバリーのオペレーションを容易にし、システム全体の信頼性を向上させることができます。

最後に、今後の展開についてお話しください。

南部 このNILFSの開発と並行して、Linuxの標準ファイルシステムの信頼性を検証し、発見した問題を解消するためのパッチをLinuxカーネルコミュニティに提供しています。今後は、引き続き開発を行いながら、入出力性能とディスク利用効率の向上による最適化やユーザーフレンドリーな管理運用ツールなどを提供し、さらなる機能、性能、品質の向上を図り、Linuxの利用拡大を促進していきます。