

攝津 敦[†] 菅井 尚人[†] 畑田 貴史[‡] 山口 義一[†]

[†]三菱電機 (株) 情報技術総合研究所

[‡]三菱電機 (株) 電力・産業システム事業所

1. はじめに

近年、UNIX 計算機は、高機能/高信頼の要求から産業用システムに適用されてきている。産業用システムにおいて、屋外設置される制御装置などでは、高い耐環境性を満足するため、可動部分がなくかつ大容量/低コストであるフラッシュ EEPROM を補助記憶装置として使用するものがある。

前回の報告では、フラッシュ EEPROM を UNIX ファイルシステムのブロックデバイスとして扱うことができるフラッシュ FS デバイスドライバを設計した[1]。本稿では、フラッシュ FS デバイスドライバの実装と性能評価結果について述べる。

2. 実装方式

フラッシュ EEPROM をブロックデバイスに適用する場合、データの消失を防ぐことと、更新単位が大きいことによる書き込み性能の劣化を防ぐことの2つが課題となる。そこで今回、図1のような実装方式を採用した。

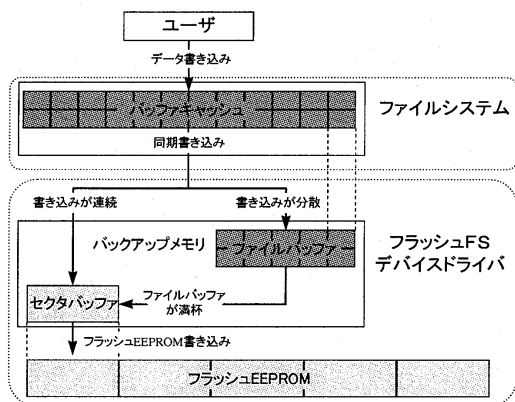


図1. 実装方式

ファイルシステムでは、不意の電源断によるデータの消失を防ぐために、バッファキャッシュを同期書き込みに行っている。

ドライバでは、フラッシュ EEPROM の更新単位 (256KB。以下、セクタ) と同じサイズのバッファ (セクタバッファ) と、バッファキャッシュのブロック単

位 (4KB) と同じサイズのバッファ (ファイルバッファ) を設ける。各バッファは電源断時のデータ消失を防ぐために電源断が発生してもデータが消去されないバックアップメモリに設置される。セクタバッファおよびファイルバッファの個数は、バッテリバックアップメモリの容量および書き込み性能の試算結果から、セクタバッファ 1 個とファイルバッファ 14 個という構成にした。この構成により、連続したブロックへの書き込み要求をフラッシュ EEPROM の更新単位 (256KB) と同じサイズのセクタバッファで吸収し、かつ分散したブロックへの書き込み要求はバッファキャッシュのブロック単位と同じサイズのファイルバッファで吸収することで書き込み性能の劣化を防ぐ。

書き込み要求はセクタバッファおよびファイルバッファで吸収するが、セクタバッファおよびファイルバッファに空きが存在しない場合は、以下のアルゴリズムにてファイルバッファに空きを作る。

- ① セクタバッファの内容をフラッシュ EEPROM に書き込む
- ② 書き込み要求ブロックおよびファイルバッファ内のブロックの内、1セクタ内に収まるブロックの数が一番多いセクタ番号を算出し、そのセクタをセクタバッファに読み込む。
- ③ 書き込み要求ブロックおよびファイルバッファ内のブロックで、②によりセクタバッファにコピーされたセクタに属するブロックをセクタバッファに移動する。(移動されたファイルブロック内のブロックは空になる)
- ④ ③で書き込み要求ブロックがセクタバッファに移動しなかった場合、③によって空きとなったファイルバッファに要求ブロックを書き込む。
- ⑤ ②～④の操作によって空きとなるファイルバッファの数が5個未満の場合、ファイルバッファの内容を強制的にフラッシュ EEPROM に書き込む。

⑤は、②～④の操作を続けると、ファイルバッファ内に孤立ブロック (そのブロックが属するセクタに含まれるブロックが他に存在しない) が増加するため、その対処として強制的に孤立ブロックをフラッシュ EEPROM に書き出している。

An evaluation of Unix File System using Flash EEPROM.

Atsushi SETTSU, Naoto SUGAI,
Takashi HATADA, Yoshikazu YAMAGUCHI
Mitsubishi Electric Corp.

4. 性能評価

4.1 コマンド実行性能

まず、今回想定したシステムの要求性能を満足しているかを評価した。今回想定したシステムでは、ファイルシステムへの書き込みは制御装置の保守作業時に実施される。この時行われる作業はインストールや設定ファイルの更新等であり、使用されるコマンドは限られる。これら保守作業時に使用されるコマンドの実行時間について測定した(表 1)。コマンドの実行時間については、文献[1]で示した本実装方式の試算値および、書き込み要求を逐次フラッシュ EEPROM に書き込んだ場合の試算値(表 1 中では参考値)も併記する。

表 1. コマンド実行性能

	参考値	試算値	実測値
ファイルコピー(64K)	3分22秒	3.61秒(17.7KB/秒)	2.45秒(26.1KB/秒)
ファイルコピー(32K)	1分44秒	0.26(123.1)	0.26(123.1)
5MBアーカイブ展開	18時間54分	127.04(40.3)	146.48(35.0)
mks	126.15秒	91.06	82.28

(10MB のファイルシステムでの試算および測定)

コマンドの実行時間については、ほぼ試算値と同程度になっており、想定したシステムでの要求性能を満たすことが確認できた。

4.2 書き込み性能の特性

次に今回検討した方式に対し、ファイルシステムの書き込み性能の特性を評価するため、通常ファイルへの1回の書き込みに対する性能測定を行った。測定では、ファイルサイズを4KB~256KBに変化させた場合の性能を測定した(図 2)。

この結果より、本方式での書き込み特性として以下のことが判明した。

- 4~32KB まではフラッシュ EEPROM の書き込み性能よりも性能が良い(260~370KB/秒)。これは、ファイルバッファおよびセクタバッファによるバッファリングの効果があり、フラッシュ EEPROM への書き込みが発生しないためと考えられる。また、この区間での書き込み性能は、ファイルバッファおよびセクタバッファが存在するバックアップメモリの性能に支配されるが、通常ファイルの書き込みではスーパーブロックやinode領域に複数回の書き込みが発生するため、バックアップメモリの性能(約3MB/秒)の1/10程度となっている。
- 32~64KB で性能は低下する。
通常ファイルの書き込みの場合、セクタバッファにスーパーブロックおよびinode領域が保存された後、ファイルブロックデータの書き込み要求が発生するので、データはファイルバッファ

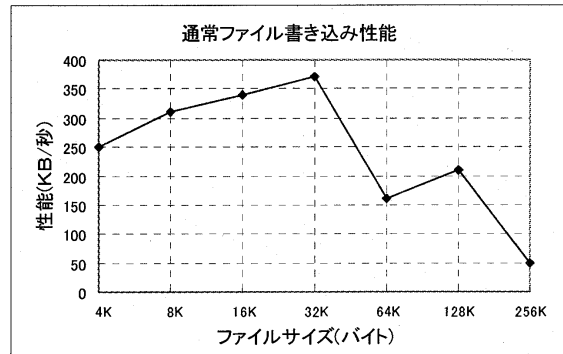


図 2. 1 回の書き込みに対する書き込み性能

に溜まることになる。従って、データがファイルバッファの最大サイズ(14個×4KB=56KB)よりも大きくなると、フラッシュ EEPROM への書き込みが発生するため、消去時間を含めたフラッシュ EEPROM の書き込み性能(約130KB/秒)により性能が低下していると考えられる。

- 64~128KB で一旦性能が向上する。
これは、64KB 付近でフラッシュ EEPROM への書き込みが発生することにより、ファイルバッファに空きができ、1.と同様なバッファリングの効果が現れているためと考えられる。
- 128~256KB でまた性能が低下する。
本実装方式では、フラッシュ EEPROM のセクタに管理ブロックを設けており、実際に書き込めるデータは256KBよりも小さい。そのため、256KB のデータ書き込みでは最低1回の書き込みが発生するため、ここで性能が低下していると考えられる。

結果、本実装方式では64KB付近におけるファイルバッファの最大サイズが原因による性能低下と、256KB 付近におけるフラッシュ EEPROM のセクタサイズが原因による性能低下が現れることがわかる。

5. おわりに

本稿では、フラッシュ FS デバイスドライバの構造と性能評価について以下の内容を述べた。

- セクタバッファとファイルバッファという2つのバッファを用いた方式を実装した。
- 適用システムを想定した性能測定の結果、要求性能を満たすことが確認できた。
- 本実装方式では、64KB 付近と 256KB 付近の2箇所性能低下が現れることが判明した。

参考文献

- [1] 攝津 敦、菅井 尚人、山口 義一、フラッシュ EEPROM を用いた UNIX ファイルシステムの実現、情報処理学会第 58 回全国大会予稿集, 1F-2, 1999