

2005 年度上期
オープンソースソフトウェア活用基盤整備事業

「OSS 性能・信頼性評価 / 障害解析ツール開発」

OS 層
～ DAVL 編 ～

作成
OSS 技術開発・評価コンソーシアム

商標表記

- Alicia は、ユニアデックス株式会社の登録商標です。
- Asianux は、ミラクル・リナックス株式会社の日本における登録商標です。
- Intel、Itanium および Intel Xeon は、アメリカ合衆国およびその他の国におけるインテルコーポレーションまたはその子会社の商標または登録商標です。
- Intel は、Intel Corporation の会社名です。
- Linux は、Linus Torvalds の米国およびその他の国における登録商標あるいは商標です。
- MIRACLE LINUX は、ミラクル・リナックス株式会社が使用許諾を受けている登録商標です。
- Pentium は、Intel Corporation のアメリカ合衆国及びその他の国における登録商標です。
- Red Hat は、米国およびその他の国で Red Hat, Inc. の登録商標若しくは商標です。
- Solaris は、米国 Sun Microsystems, Inc. の米国およびその他の国における商標または登録商標です。
- SUSE は、米国 Novell, Inc.の一部門である SUSE LINUX AG.の登録商標です。
- Turbolinux は、ターボリナックス株式会社の商標または登録商標です。
- UNIX は、X/Open Company Limited が独占的にライセンスしている米国ならびに他の国における登録商標です。
- Windows は、米国およびその他の国における米国 Microsoft Corp.の登録商標です。
- その他記載の会社名、製品名は、それぞれの会社の商号、商標もしくは登録商標です。

目次

1	開発目的とディスク割り当て評価ツールの新機能概要.....	1-1
1.1	開発目的.....	1-1
1.2	DAVL（ディスク割り当て評価ツール）の新機能概要.....	1-1
2	DAVLによるフラグメント差分の評価.....	2-1
2.1	DBT-1 実行前後でのフラグメント差分の評価.....	2-1
2.1.1	評価環境.....	2-1
2.1.2	評価手順.....	2-2
2.1.3	結果.....	2-2
2.1.4	考察.....	2-7
2.2	既存デフラグツール実行前後でのフラグメント差分評価.....	2-7
2.2.1	評価環境および評価手順.....	2-8
2.2.2	結果と考察.....	2-8
3	総括.....	3-1
3.1	DAVL新機能の有効性.....	3-1
3.2	開発規模.....	3-1
3.3	コミュニティへの貢献.....	3-1
3.4	今後の課題.....	3-2

1 開発目的とディスク割り当て評価ツールの新機能概要

1.1 開発目的

エンタープライズシステムにおいてディスクアクセス性能が劣化した場合の迅速な障害解析を目的として、2004 / 下期にディスク割り当て評価ツール (DAVL) の開発を行った。
(URL : <http://sourceforge.net/projects/davl>)

DAVL は当初の目的を満足するものだったが、下記のような問題が出てきた。

- (1) 対応する GUI ライブラリは GTK+1.2 であるが、バージョンが古く最新のディストリビューションには標準で含まれていない
- (2) フラグメンテーション状態の推移が分かりづらい

DAVL では、フラグメンテーション状態をファイルに保存 (以降、スナップショットと呼ぶ) することが可能であるが、スナップショット同士を比較してフラグメンテーション状態の推移を見たい場合、diff などによるテキスト出力でしか比較結果を表示できないため、その推移が分かりづらい。

そこで次のステップとして、下記を目的として開発を行った。

- (1) 最新ディストリビューションに対応するための GTK+2 ライブラリへの対応
- (2) GUI 上でのフラグメンテーション状態の推移確認を可能とする

フラグメンテーション状態の推移を確認することで、フラグメンテーションが発生し易いディレクトリ / ファイルを判断できれば、その情報から、それらを使用するアプリケーションも特定できることとなる。即ち、フラグメンテーション状態の推移から、アクセス性能劣化を発生させる可能性が高いアプリケーションを予測できることとなる。

1.2 DAVAL (ディスク割り当て評価ツール) の新機能概要

DAVAL の新機能を以下に示す。

- (1) GTK+2 対応 (現在は、GTK+1.2 も引き続きサポート)
- (2) スナップショットの差分表示機能
- (3) フラグメントファイル一覧における項目毎のソート機能

今回開発した DAVAL では、GTK+2 をサポートすることで、最新ディストリビューション (Fedora Core 3) に対応する。また、GUI 上でスナップショットの差分表示を行うことにより、フラグメンテーション状態が変化したディスクブロックを可視化する。差分表示では、ディスクブロックの変化を表示するだけでなく、フラグメントファイルの差分について

でも一覧表示が可能である。フラグメントファイル一覧では、フラグメント率やパス名等によってソートすることが可能であり、性能劣化障害の多角的な解析を支援する。

図 1.2-1 に DAVL の構成を示す。DAVL は以前と変わらず下記の3つのプログラムおよびカーネルモジュールによって構成されており、今回開発した新機能は全てディスク割り当て可視化ツール内に実装される。

- (1) ファイルシステム情報取得ツール
- (2) ディスク割り当て可視化ツール
- (3) ファイルブロック情報取得モジュール

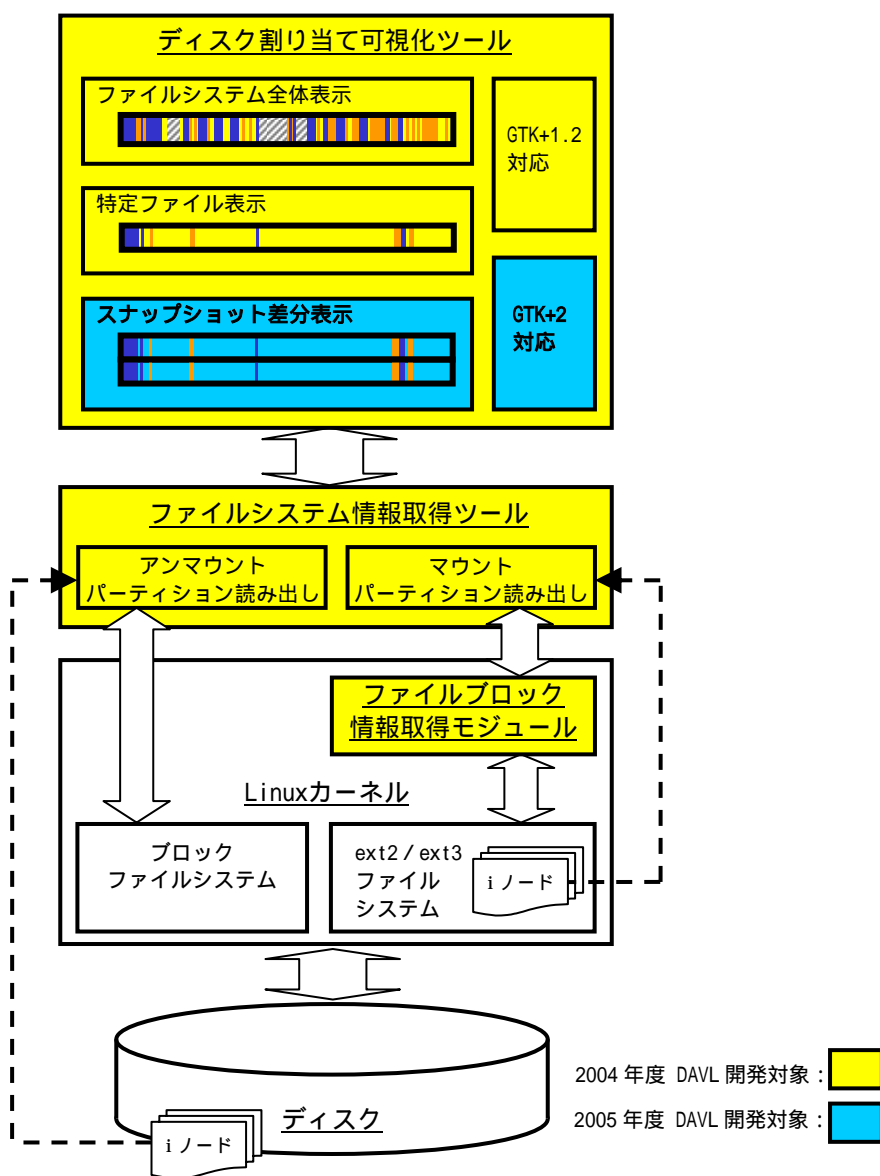


図 1.2-1 DAVL 構成

2 DAVL によるフラグメント差分の評価

2.1 DBT-1 実行前後でのフラグメント差分の評価

DAVL 新機能の有効性を確認することを目的として、DAVL によるフラグメント差分表示の評価を行った。評価対象のアプリケーションとして、本プロジェクト内でも使用されている MaxDB を用いた DBT-1 を実行し、その前後のフラグメンテーション状態を差分表示して評価を行う。

MaxDB では、各データテーブル用に使用するファイルとして巨大なファイルをいくつか確保し、そのファイル中に独自形式でデータテーブルを配置して管理 / 使用することが知られており、ディスクフラグメンテーションの発生状況が変化しづらいような仕組みとなっている。一方、DBT-1 の出力するログファイル群は通常のファイルであり、ディスクフラグメンテーションに関しては特に考慮されていないため、フラグメンテーションが発生 / 変化するはずである。

この予想が正しいかどうか、また、それらの結果を DAVL のフラグメント差分表示から読み取ることが可能かどうかを評価する。

2.1.1 評価環境

今回の評価環境を表 2.1-1 に示す。

表 2.1-1 DBT-1 実行前後でのフラグメント差分の評価環境

項番	項目	評価環境		
1	ハードウェア	CPU	Xeon 3.2GHz × 2	
2		メモリ	8GByte	
3		ハードディスク	モデル	IBM-ESXS DTN146C3UCDY10FN
4			容量	140GByte (SCSI)
5	ソフトウェア	ディストリビューション	MIRACLE LINUX V3.0	
6		カーネル	2.4.21	
7		ドライバ	Adaptec AIC79xx driver version 2.0.8	
8		ファイルシステム	ext3 (ブロックサイズ : 4KByte)	
9		負荷ツール	OSDL DBT-1	
10		DB	MaxDB-7.5.00.19(*)	

*) MaxDB

ライセンスは GPL。パラメータや設定事項が少なく、GUI が豊富で、管理が容易である。

2.1.2 評価手順

具体的な手順を以下に示す。

- (1) 前期の DB 層の評価報告書 (URL : <http://www.ipa.go.jp/software/open/forum/development/download/db.pdf>) を参照し、MaxDB データベース構築から DBT-1 の環境設定までを行う。
- (2) DBT-1 を実行する。DBT-1 実行後、ディスクアクセスが終了するまで 8 分間 (デフォルト設定時の待ち時間) 時間待ちを行い、MaxDB をオフラインにする。
- (3) DAVL を使い、ディスク全体のスナップショットを取得する。また、評価用として MaxDB 部分 (/var/opt/sdb/data/wrk/DBT1) と DBT-1 出力ログ部分 (DBT-1 実行時に引き数で指定するログディレクトリ) のディレクトリのスナップショットもそれぞれ取得する。
- (4) リポートする。
- (5) 上記手順(2)、(3)をもう一度実行する。
- (6) これまでの手順で得られたスナップショットを用いて差分表示を行う。

2.1.3 結果

DAVL を用いた MaxDB 部分のディスクブロック差分表示を図 2.1-1 に示す。

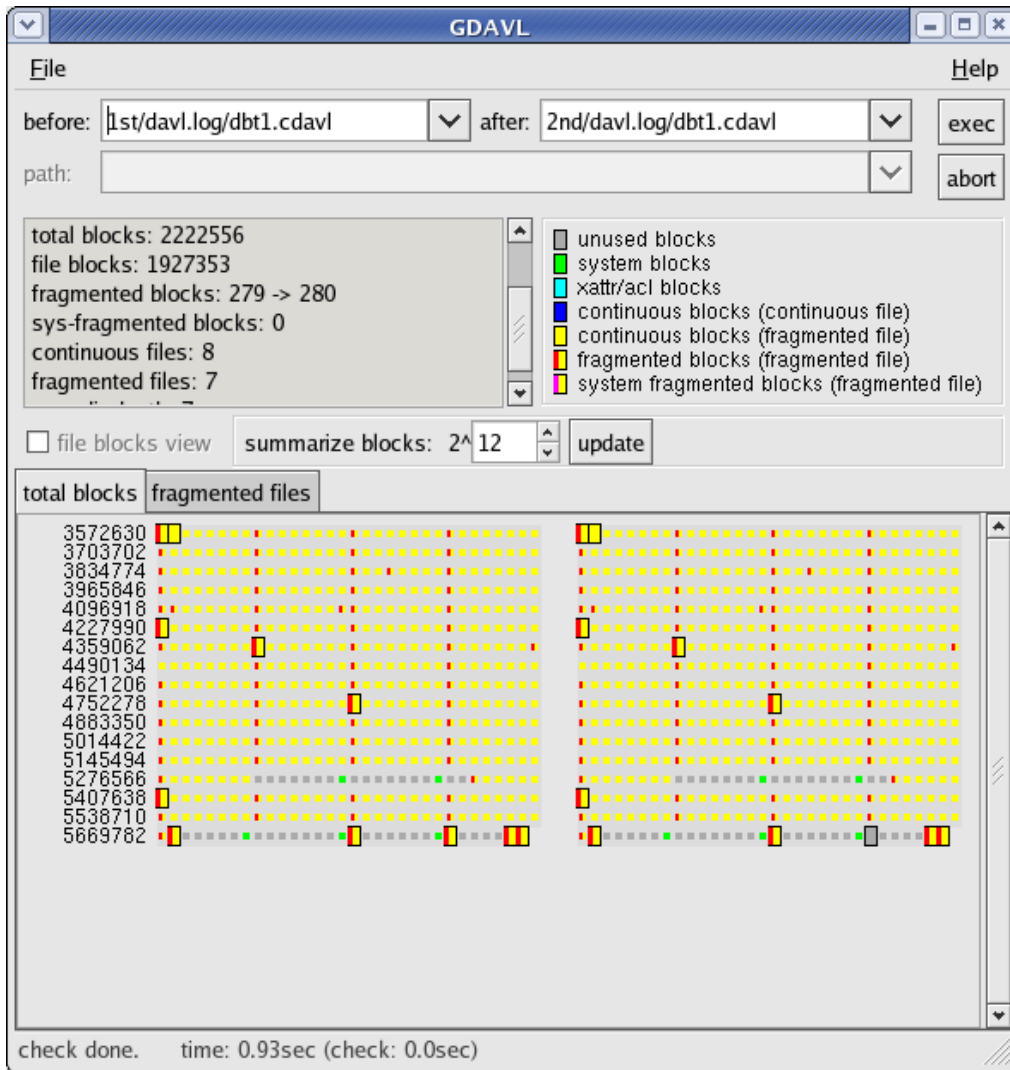


図 2.1-1 ディスクブロック差分表示 (MaxDB 部分)

この図で、左側のディスクブロックが before 覧で指定したスナップショットを示しており、右側が after 覧で指定したスナップショットを示したものである。また、変化したブロックは枠付きで表示され、変化していないブロックは枠なしで表示される。

この図より、MaxDB 部分ではほとんどディスクブロックのフラグメンテーションは変化していないことが確認できる。

次に、DBT-1 出力ログ部分のディスクブロック差分表示を図 2.1-2 に示す。

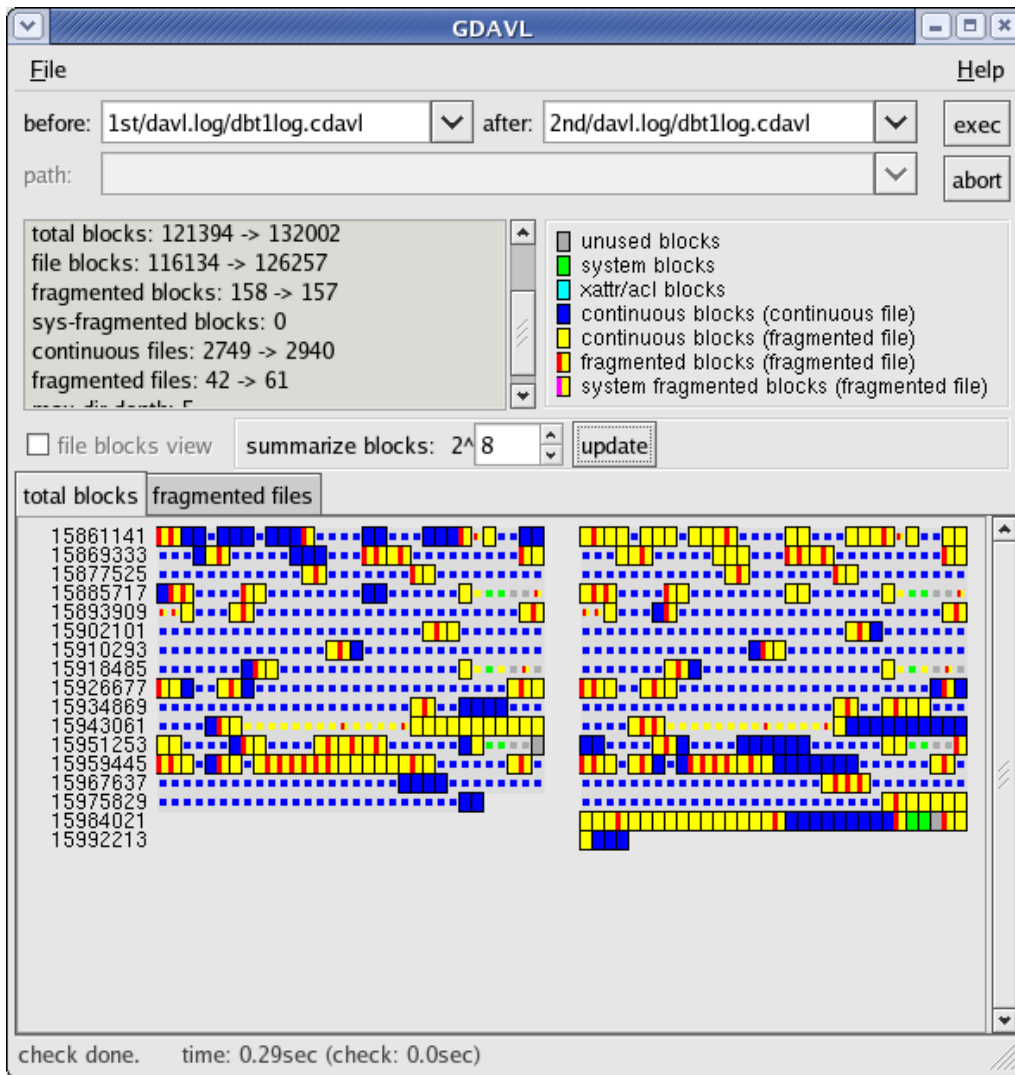


図 2.1-2 ディスクブロック差分表示 (DBT-1 ログ出力部分)

この図より、DBT-1 出力ログ部分ではディスクブロックのフラグメンテーションがかなり変化していることが確認できる。

以上により、DAVL の差分表示を用いることによって、当初の予想が正しかったことを確認できることが分かる。

今回は、評価用として特定ディレクトリのスナップショットを保存していたが、これらを保存していない場合でも、フラグメンテーションの変化が大きいディレクトリ/ファイルを特定できることを確認する。ディスク全体のディスクブロック差分表示を図 2.1-3 に示す。

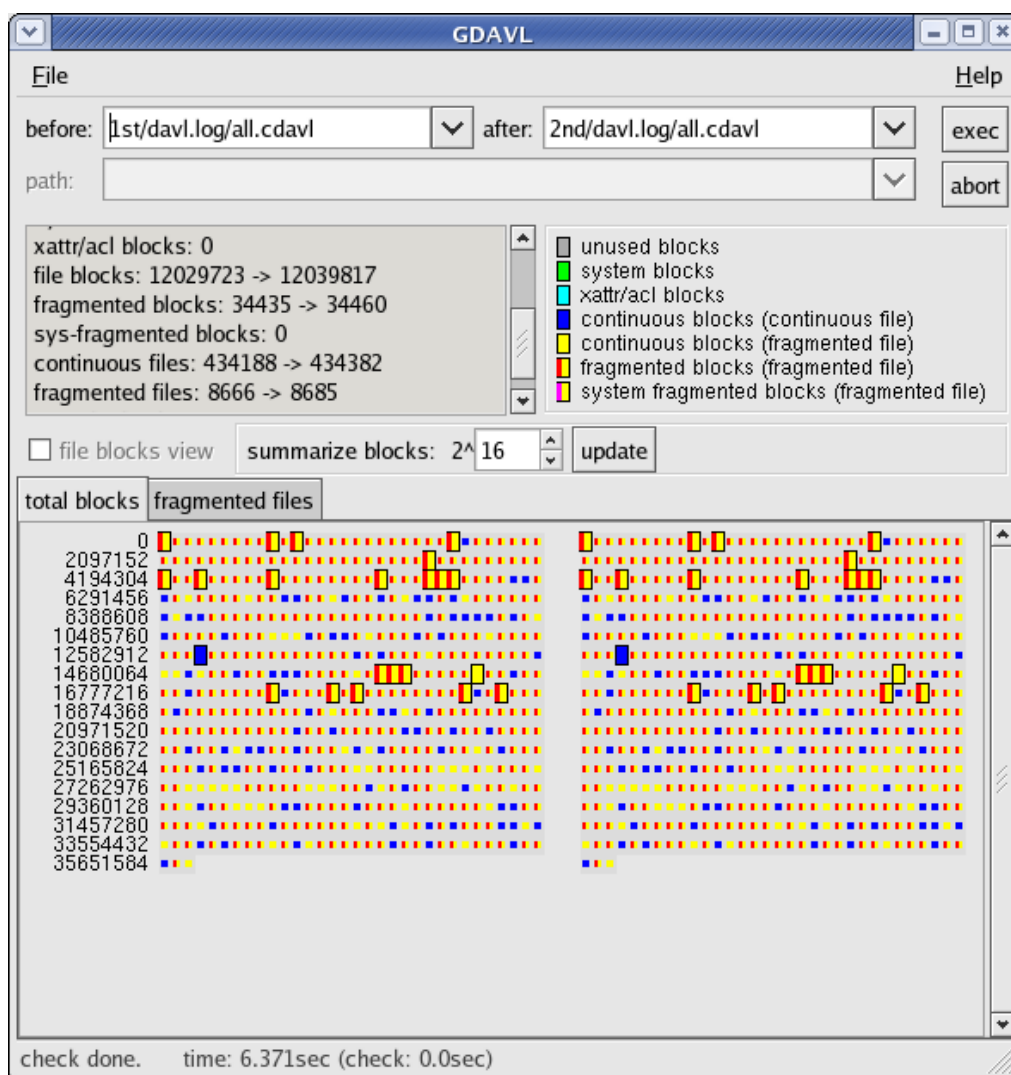


図 2.1-3 ディスクブロック差分表示 (ディスク全体)

ディスク全体のブロックを差分表示で確認した場合、様々なディレクトリ/ファイルのブロックが混在しているため、フラグメンテーションが変化している部分を特化することはできない。そこで、差分がある (ブロック数またはフラグメントが変化した) フラグメントファイルの一覧を確認する (図 2.1-4 参照)。

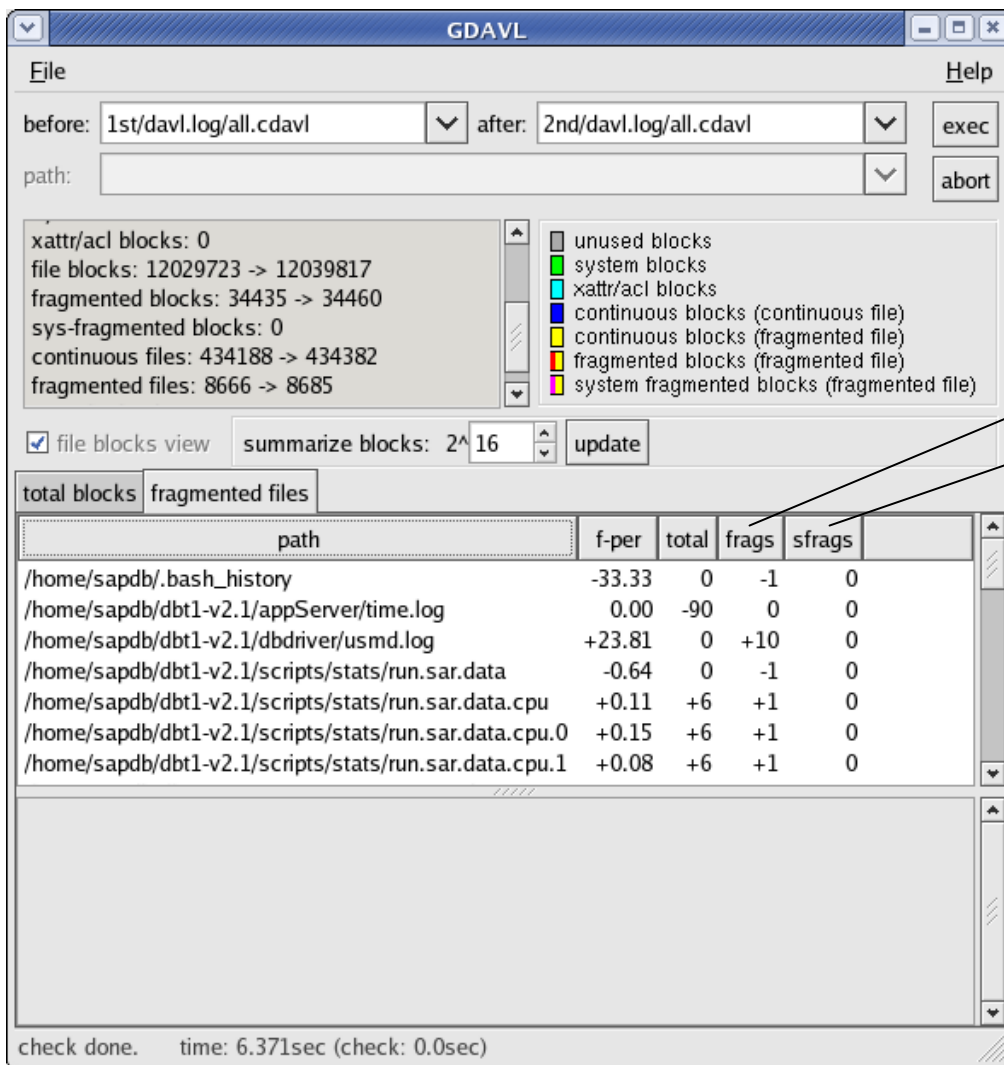


図 2.1-4 フラグメントファイル差分表示 (ディスク全体)

各ディレクトリにおけるフラグメント変化の度合を確認するため、DAVL のフラグメントファイルの差分一覧表示を参照し、ディレクトリ毎に(1)差分のあるフラグメントファイル数と(2)フラグメント数の変化 (図 2.1-4 - 、 の絶対値の合計) の総和を、分類 / 集計した表を表 2.1-2 に示す。

表 2.1-2 フラグメント率変化ディレクトリ情報

#	ディレクトリ名	分類	差分ファイル数	フラグメント数差分の絶対値の総和
1	/home/sapdb	DBT-1 ホーム	1	1
2	/home/sapdb/dbt1-v2.1/appServer	DBT-1 出力ログ	1	0
3	/home/sapdb/dbt1-v2.1/dbdriver		1	10
4	/home/sapdb/dbt1-v2.1/scripts/stats		8	8
5	/home/sapdb/fujiwara/dbt1.log		74	151
6	/var/log	システム関連	5	8
7	/var/log/sa		1	2
8	/var/log/samba		1	0
9	/var/opt/sdb/data/wrk/DBT1	MaxDB	2	9
10	/var/opt/sdb/data/wrk/root		1	0

この表から、フラグメントファイルの変化は、システム関連のディレクトリ (/var/log) 以外では、DBT-1 ホームと DBT-1 出力ログ部および、MaxDB 部で発生していることが確認できる。また、フラグメント数差分の絶対値の総和から、フラグメントの変化は DBT-1 出力ログ部分に集中していることも確認できる。

MaxDB でデータテーブル用に使用されるファイル DAT_001 および LOG_001 に関しては、差分ファイル中に含まれておらず、データテーブルではフラグメントの変化はなかったことが確認できる。

2.1.4 考察

今回の評価の結果、DAVL を使用することで、当初の予想通り MaxDB のデータテーブル用ファイルはフラグメンテーションが変化しづらく、DBT-1 の出力ログはフラグメンテーションが変化することが確認できた。

DAVL に関しては、スナップショット差分表示によって、フラグメンテーションの変化が多いディレクトリやファイルを簡単に特定することが確認できた。

2.2 既存デフラグツール実行前後でのフラグメント差分評価

前期の DAVAL 開発時に、DAVL による既存デフラグツールの効果評価を行った。この時と同じ手順によって取得したデフラグツール実行前後のスナップショットについて、差分表示を行い評価した。

2.2.1 評価環境および評価手順

使用したデフラグツールを含む評価環境は、前期の DAVL 開発時のものと同様である。下記の URL を参照のこと。

<http://www.ipa.go.jp/software/open/forum/development/download/dav-report.pdf>

評価手順に関しても、スナップショットを比較する際に DAVL の新機能を用いること以外は同様である。

2.2.2 結果と考察

DAVL によるデフラグ実行前後の差分表示を図 2.2-1 に示す。

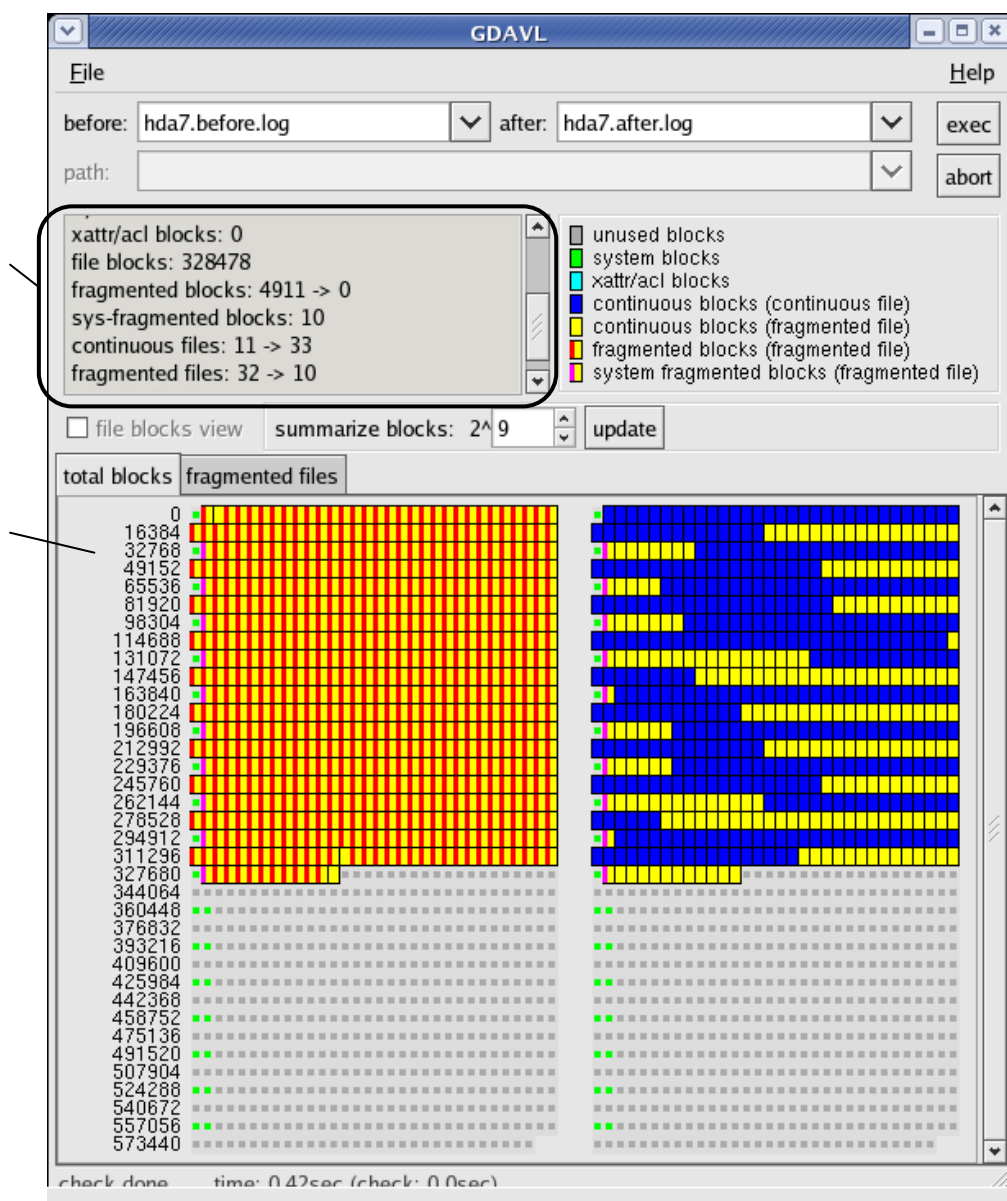


図 2.2-1 既存デフラグツール実行前後でのフラグメント差分表示

DAVL のテキスト表示部分（図中の ）から、フラグメントブロック数が 4,911 から 0 になったことや、システムブロックによるフラグメントブロック数は 10 のまま変わっていないことなどが確認できる。また、ディスクブロック表示部分（図中の ）からも、システムブロックや未使用ブロック以外のブロックは、全面的にフラグメントが変化していることが確認できる。

従来の DAVL であれば、デフラグ前後の 2 回分の GUI 表示を行い確認する必要があったが、新機能である差分表示を用いることで、1 回の GUI 表示により、デフラグ効果が一目で確認できるようになったことが分かる。

3 総括

3.1 DAVL 新機能の有効性

今回開発した DAVL の新機能を利用することで、単一のスナップショットを確認しただけでは分からなかった、フラグメンテーション状態の変化について簡単に確認することが可能になった。

また、GTK+2 に対応したことで、最新のディストリビューションでも問題なくコンパイル/動作できるようになった。

3.2 開発規模

DAVL 新機能の開発規模を表 3.2-1 に示す。

表 3.2-1 DAVL 新機能の開発規模

#	項目	開発言語	ステップ数
1	母体分	C	5,426
2	改造分	C	1,185
3	新規作成分	C	310

3.3 コミュニティへの貢献

DAVL に関して、以下の活動を行っている。

- ・ sourceforge においてソースコードを公開
- ・ メーリングリストで寄せられた質問・要望に対応 (GTK+2 対応もそのひとつ)
- ・ Linux のユーザ会において勉強会を開催
- ・ 日経 Linux 2005 年 7 月号に記事を執筆し、CD-ROM にソースコードをバンドル
- ・ LinuxWorld2005 などのカンファレンスで講演

DAVL は次期ミラクル・リナックス社製品に採用予定である。今後も、ユーザ会やシンポジウムを通じて、ユーザ、カーネル開発者、ディストリビュータ等に対し、普及活動を継続する。

3.4 今後の課題

DAVL の新機能の開発によって、当初の目的を達成することができた。今後更に次のような改善を行うことによって、対象となるシステム/ユーザが拡大し、更に活用範囲が広がると考える。

- (1) デフラグ機能の実装
- (2) XFS/JFS/ReiserFS への対応
- (3) 更なる利便性の向上...フラグメントファイルの差分一覧を CSV 形式で出力するなど