

巨大ブロックサイズファイルシステムによる
セキュアなコンテンツ共有・流通基盤の大規模 I/O 性能の向上
I/O Performance Improvement on Secure Data Sharing and Distribution Platform
with Large Block Sized Filesystem

近 丈 一 郎[†] 中 島 健 司[†] 山 口 実 靖[†]
Joichiro Kon Kenji Nakashima Saneyasu Yamaguchi

1. はじめに

インターネットの普及に伴い世界で扱われるデータが増加し、それらのデータを効率的に利活用していくことに注目が集まっている。また、センシティブなデータを安心して扱うことのできる様にするには、セキュアなコンテンツ共有・流通基盤の構築が重要であると考えられる。このようなコンテンツ共有・流通を行うためには、暗号技術の併用が欠かせない。しかし、大規模なデータを全て暗号化するとデータ量が大幅に増加することがあり、実用的な処理時間を実現するためには大規模データの I/O 処理速度の向上が重要となる。

本稿では、巨大な I/O に対して有効である大きなブロックサイズでファイルの管理を行うファイルシステム [1] のブロックサイズと性能に関する考察を行う。そして、ブロックサイズのさらなる拡大による I/O 性能の向上する手法を提案し、性能評価によりその有効性を示す。

2. 既存研究

本章では、既存研究で用いられたゲノム配列の秘匿検索アプリケーションと巨大ブロックサイズファイルシステムについて述べる。

2.1 ゲノム配列の秘匿検索アプリケーション

バイオインフォマティクスの分野の研究では、ゲノム配列の検索が頻繁に行われる。しかし、個人ゲノムの利用はプライバシー保護への対策が必須であり、暗号化したままゲノム配列の検索処理を行う必要がある。これを実現する手法の一つとして、PBWT 離散データ構造と完全準同型暗号を用いたゲノム秘匿検索手法 [2] がある。図 1 にゲノム配列検索の概略図を示す。

当該手法を実装したアプリケーションのモデルは、クライアント(検索者)はサーバ(データ保有者)に対してクエリ内容を秘匿し、サーバはクライアントに対してクエリの検索結果以外の情報を秘匿する。当アプリケーションは、1000ゲノムプロジェクト [3] における SNP 配列 2、184 サンプルを持つデータベースを使用して性能評価が行われた。

2.2 巨大ブロックサイズファイルシステム

既存研究 [1] において、ゲノム配列の秘匿検索アプリケーション実行中のディスク I/O を解析し、巨大なシーケンシャルアクセスが行われていることを発見している。そして、このアプリケーションのために、巨大なブロックサイ

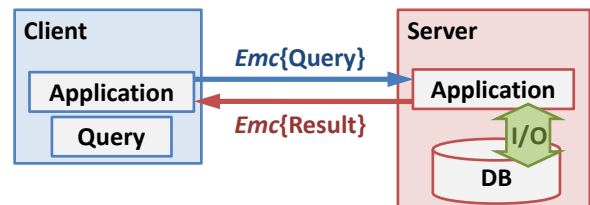


図 1 ゲノム配列検索の概略図

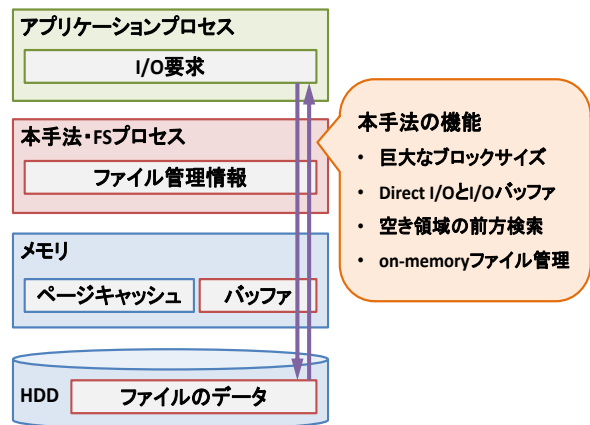


図 2 巨大ブロックサイズファイルシステムの概略図

ズでファイルの管理を行うファイルシステムを構築している。

本節にて、巨大ブロックファイルシステムの概要と実装について述べる。

2.2.1 巨大ブロックサイズファイルシステムの概要

本ファイルシステムは、定記録密度方式の HDD に対してシーケンシャルアクセスを行う場合、ディスク外周側の連続領域に対して行う方が内周側に対して行うよりも高い I/O 性能が得られるという知見に基づいている。

そして、連続性の高い I/O 要求を HDD の持つ高いシーケンシャルアクセス性能を活かして処理するために、巨大なブロックサイズでファイルの管理を行っている。

2.2.2 巨大ブロックサイズファイルシステムの実装

本ファイルシステムの実装は、FUSE (Filesystem in Userspace) を用いて行われている。図 2 に概略図を示す。

ユーザ空間上での実装であり、ディスクへの I/O 処理はデバイススペシャルファイルに対して直接行い、各ファイルの名前や格納位置などのメタ情報はメモリ上で管理している。メタ情報をメモリ上で管理することにより、これら

[†] 工学院大学大学院 工学研究科 電気・電子工学専攻,
Electrical Engineering and Electronics, Kogakuin University
Graduate School

がディスク上に存在しデータアクセスの連続性が損なわれる可能性を排除している。ファイルシステムのブロックサイズは 256 MB に設定されており、ディスク領域はこのブロック単位で分割されている。そして、前述の様にディスク領域は全てファイルのデータ領域として使用することにより、連続性の高いディスクアクセスを実現している。シーケンシャルアクセス性能の高いディスク外周部を積極的に使用するために、新規ファイルの格納位置検索方法は、ディスクアドレスの低い方から順に空きブロックの検索を行う仕様となっている。また、シーケンシャルアクセスに対して効果が低いとされるページキャッシュを無効化するために、ディスクアクセスには Direct I/O を使用している。このディスクアクセスをブロックサイズ単位で行うために、ブロックサイズ単位のバッファを設けている。

3. 改善手法

前章の巨大ブロックサイズファイルシステムのブロックサイズは 256 MB であった。本稿では、このブロックサイズを拡大することによりさらなる I/O 性能の向上を実現する手法について考察する。

4. 性能評価

4.1 測定方法および環境

ゲノム配列検索アプリケーションのデータベースを格納するファイルシステムを XFS または既存手法 (ブロックサイズ 256 MB) または改善手法 (ブロックサイズ 2 GB) としたときのデータベース読込時間を比較することにより性能評価を行った。アプリケーションはそれぞれ 10 回ずつ実行し、平均読込時間を比較した。測定環境は、物理計算機 1 台でサーバとクライアントを並列に動作させ、データベースの容量は 2 GB とした。測定に使用した物理計算機の仕様は表 1 の通りであり、データベースを格納する HDD の仕様は表 2 の通りである。

4.2 測定結果

データベース平均読込時間を図 3 に示す。ファイルシステムに XFS を使用した場合の平均読込時間は、11.63 秒、

表 1 測定用物理計算機の仕様

CPU	Intel Core i5-4670 3.40 GHz
OS	CentOS 7 x86_64 minimal
Kernel	Linux-3.10.103
SSD	512 GB (ext4)
HDD	1 TB
Main Memory	16 GB

表 2 測定用 HDD の仕様

型番	WD1003FZEX
インターフェース	SATA 3.0 (6.0 Gbps)
容量	1 TB
バッファ容量	64MB
回転数	7,200 rpm

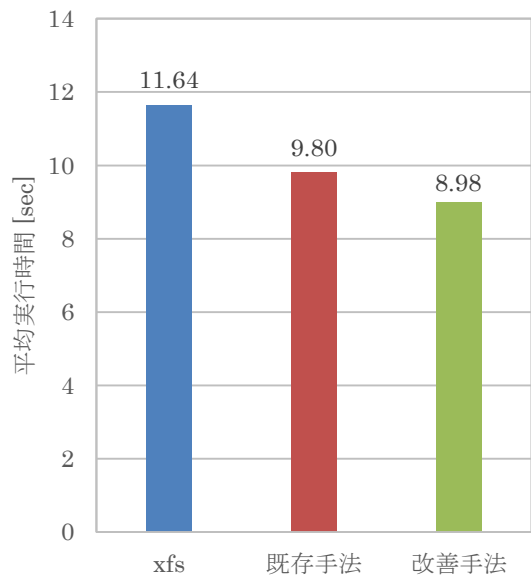


図 3 データベースの平均読込時間

既存手法を使用した場合は 9.80 秒となっている。そして、改善手法を使用した場合の平均読込時間は 8.98 秒となり、XFS と比較して 22.7%、既存手法と比較して 8.3% の性能向上が確認された。

ブロックサイズをデータベースと同じ 2 GB に設定したことで、既存手法においては 8 回行われたディスクアクセスが 1 回となり、性能が向上したと考えられる。

5. おわりに

本稿では、セキュアなコンテンツ共有・流通基盤用に作成されたゲノム配列の秘匿検索アプリケーションの I/O 性能を向上させる巨大なブロックサイズでファイル管理を行うファイルシステムに着目し、ブロックサイズのさらなる拡大による I/O 性能の向上について考察した。

今後は、書込処理についての考察を行っていく予定である。

謝辞

本研究は、JST、CREST JPMJCR1503 の支援を受けたものである。

本研究は JSPS 科研費 26730040, 15H02696, 17K00109 の助成を受けたものである。

参考文献

- [1] 藤島 永太, 中島 健司, 近 丈一郎, 山口 実靖, “ビッグデータ統合利用基盤における大規模 I/O 性能の向上”, DEIM2017 第 9 回データ工学と情報マネジメントに関するフォーラム, H5-3, (2017).
- [2] 石巻 優, 清水 佳奈, 縫田 光司, 山名 早人, “完全準同型暗号を用いた高速なゲノム秘匿検索”, 2016 Symposium on Cryptography and Information Security (SCIS2016), (2016.1).
- [3] The 1000 Genome Project Consortium: “An integrated map of genetic variation from 1,092 human genomes”, Nature, Vol.491, pp.56-65, (2012).