

D-046

## インターネット上で分散可能なビッグデータファイル

古庄 晋二<sup>†a)</sup>

### Distributable Big Data File over the Internet

Shinji FURUSHO<sup>†a)</sup>

#### あらまし

新開発インデックスにより、インターネット上で分散可能なビッグデータファイルができた。

- 1) 新開発インデックスにより
    - ・低レイテンシなのでインターネット上でも動作可能
    - ・インデックスはすべての項目に張られている
    - ・複数のテーブルのインデックスを瞬時に結合できる
  - 2) インターネット上でこのファイルを仮想的に結合し、ビッグデータを検索可能
  - 3) クライアント側 CPU で検索でき、サーバレスにでき大幅コストダウン
  - 4) ファイルを守るだけなのでセキュリティ対策がしやすい
- キーワード データベース、インデックス、ビッグデータ、インターネット、ファイル、分散データベース

## 1. まえがき

インターネットの普及と高速化、ビッグデータの時代となってビッグデータをインターネット上で自由に共有し、自由に利用するというニーズが出てきた。しかし、以下のハードルが残されていた。

課題1. (インターネット回線の) 低レイテンシ対応

課題2. サーバを使用しないビッグデータの利用方法の開発

課題3. 世界中に分散したビッグデータの結合方法の開発

課題4. ビッグデータのセキュリティ確保 (特に改竄阻止)

上記の諸課題を解決したシステムを開発し、政府系の機関で1年以上の使用実績を重ねた。非常に複雑な検索において特に高速で旧システムの概ね 100 倍程度の検索性能を得ることができた。

## 2. システムのイメージ

使用方法はビッグデータを1テーブル1ファイルとし、このファイルをNAS(Network Accessible Storage)などのハードディスク等に格納し、インターネット上の各地に格納、クライアント側でこれをファイルとして読み込み、自由に結合し、検索などを行う。Fig 1 にシステムの実現イメージを示す。

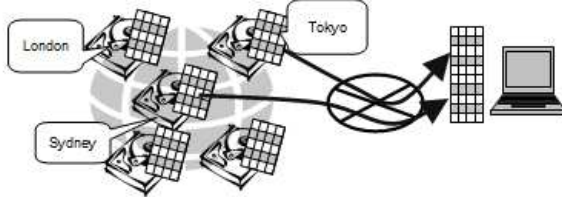


Fig. 1 各地に分散したビッグデータを結合

### <Writer>

このファイルを D5A ファイルと呼ぶ。D5A ファイルは PDF ファイルのように作成ソフトウェアと読み出しソフトウェアが完全に分離している。

ビッグデータを公開したい人は Writer を使って D5A ファイルを作成し、インターネットからアクセスできるディスクに置き、その URL をホームページ等に公開すれば良い。高価でランニングコストもかかるビッグデータのサーバは無くてもよく、普通の人がビッグデータを容易に公開できる。(Fig. 2)

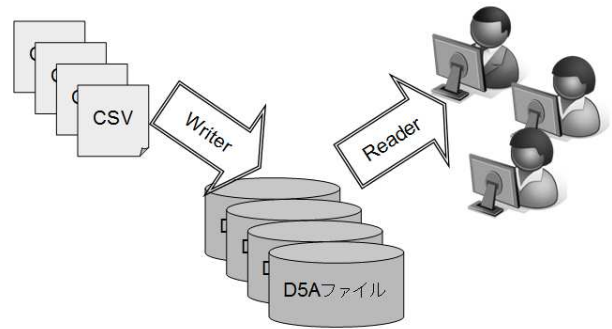


Fig. 2 Writer で作成, Reader で検索

### <Reader>

Reader を使うとクライアント側は乏しいリソースでもビッグデータを検索することができる。読み出しソフトウェア(Reader)はクライアント側の CPU のみを使ってインターネット上のファイルをルックアップ、高速な検索を行うことができる。(Fig. 3)

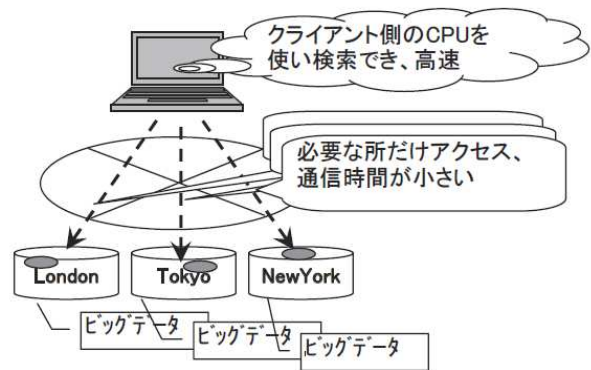


Fig. 3 Reader から各地のデータを自由にアクセス

## 3. 課題の解決

### 3.1 回線の低レイテンシ・低バンド幅の対策

地球の裏側までデータのリードリクエストを出し、回答が返ってくるまで、光の速さでも 0.134 秒かかる。(Fig. 3)



Fig. 3 インターネット上のレイテンシ

したがって従来の B-Tree などのテクノロジーよりも往復回数

†株式会社ターボデータラボラトリー, 神奈川県  
Turbo Data Laboratories, Inc., 3-21-2 Motohama-chou, Naka-ku,  
Yokohamashi Kanagawa, 221-0005 Japan  
a) E-mail: s-furusuo@turbo-data.co.jp

を抑えるような方式が必要であった。B-Tree では値とポインタを一体に格納している(Fig. 4) が、新インデックスではそれらの構成要素を分離し、さらに別のいくつかの要素を加え、1 回のアクセスで検索終了までのステップが大きく進むようにした。

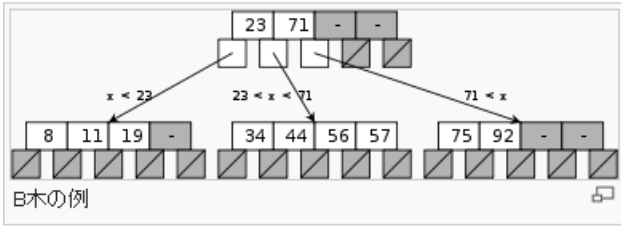


Fig. 4 B-Tree の説明(Wikipedia より)

定量的な測定は完成していないが、上海、インドのバンガロールから横浜のファイルにアクセス、結合して動かすことができた。特に上海は十分に滑らかに動かすことができた。バンガロールはホテルの回線を使ったためかぎこちなさの残る動きであった。

### 3.2 コストの大幅削減

ソフトウェアの価格は政策的なものであり、オープンソース無償配布であれば実質価格は 0 円にできる。また、インターネット回線は従量制でない限り実質 0 円であり、クライアント側設備は他にも種々使用しているので実質 0 円と考えても良い。しかし、サーバはビッグデータの処理のために専有されたり、それに近い状態で使われている。今回の提案のようにこれを無くし、NAS だけで運用すれば購入コストは数十分の一、消費電力は 1/10 以下になる。(Fig. 5)



Fig. 5 サーバとハードディスクのコスト比較

### 3.3 分散したビッグデータの結合

インターネット上のビッグデータの結合を行うにはビッグデータの論理的な順序を保存しつつデータの結合を行い、かつインデックスを結合することが必要。前者は簡単に行えるが、後者はパフォーマンスが大きく削がれてしまう。新インデックスはこれを解決している。(Fig. 6)。現在は、40 億行、25 ファイルで 1000 億行まで実験済み。ハードディスクはアドレスが離れるほどシーク時間がかかるが、現在のハードディスクは 100 億行あたりで特性の劣化が大きくなるので、100 億行 100 ファイルで 1 兆行程度までは目指せると思われる。

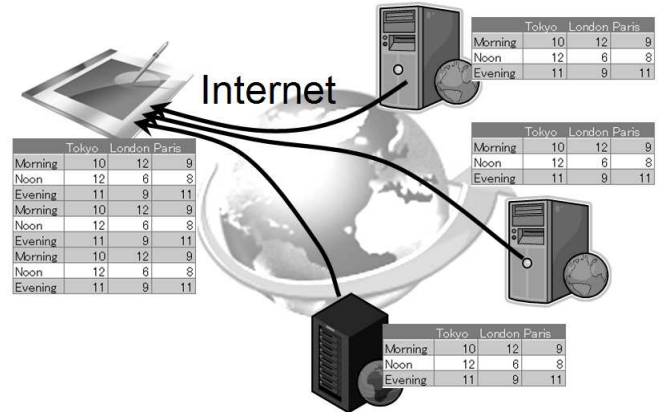


Fig. 6 分散したビッグデータの結合

### 3.4 セキュリティの課題解決

セキュリティの確保は大きな問題である。特にインターネットに常時さらされている場合は難しい。しかし、本方式では多くの攻撃方法が存在するサーバが無くて良い点でセキュリティ上有利である。ファイルが攻撃された場合でも、①ファイルが暗号化されていること、②その暗号を破ってもデータ構造を理解していないと意味のある改ざんは出来ないこと、③改ざんしても MD5 などでファイル整合性をチェックできることなどでプロテクトしやすい。

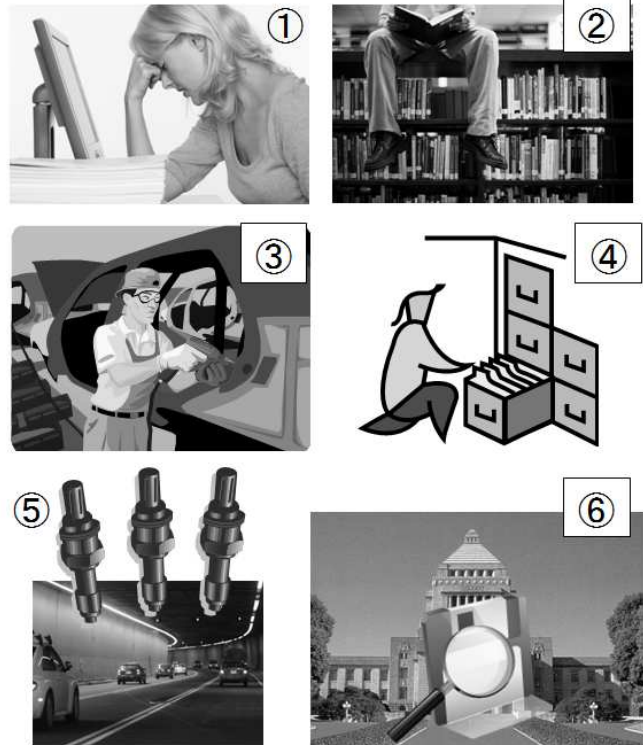
## 4. まとめ

### 4.1 開発したシステムの機能

- ① Writer を使って D5A ファイルを作成し、インターネットからアクセスできるディスクにこれを置く。次に Reader を使ってこのディスクに置かれた D5A ファイルにアクセスし、検索を行うことができる。
- ② Reader はインターネット上の複数の場所にある D5A ファイルを仮想的に 1 つに結合して使用することができる。
- ③ D5A ファイルは暗号化されており改ざんしにくい。改ざんされた場合でも改ざんを検出できる。また、サーバレスにできるので脆弱な点が少ない。

4.2 新たに開発された技術

- ① 低レイテンシのインターネットでも使用可能にするため、既存のインデックスに比べ大幅に少ない通信回数で検索を終了できるインデックス。
- ② 複数のインデックスを仮想的に瞬時にマージできるインデックス技術。



文 献

古庄 晋二 ISBN4-902721-01-5 汎用超高速データベース処理技術

5. 付録. 想定される応用

5.1 応用が期待できる分野

- ① オフィスの作業 (社内外に分散したデータを瞬時に手元に取り込む),
- ② 研究開発の現場 (世界中に分散したデータを簡単に手元に持ってくる),
- ③ 製造現場 (パーツなどの状況確認, トレーサビリティ),
- ④ セキュリティ (ログデータなどの記録),
- ⑤ センサーデータなどによる設備監視 (トンネルなど),
- ⑥ オープンなビッグデータの開示

(Fig. 7)

Fig. 7 応用が期待される分野

5.2 分散データの社内利用 (分散性の利用) Fig. 7 ①, ②

多数の拠点で保存されているデータ, 例えば POS データは本社に集めると重いので保存期間を長く取りにくい. このようなとき, 長期の保管は各ストアに任せておくことが可能になる. (Fig. 8-1)

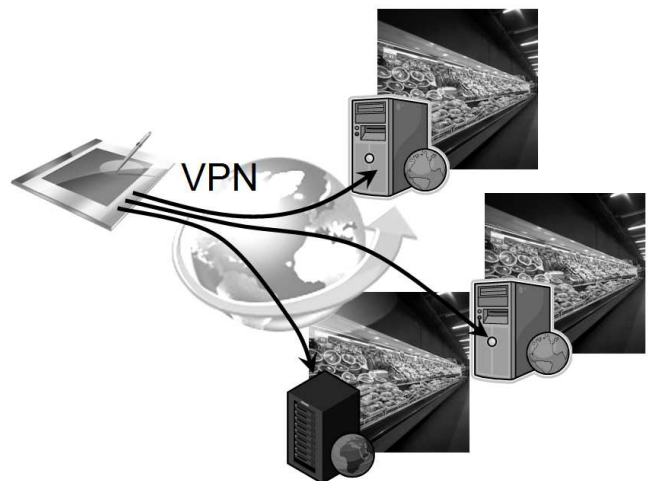


Fig. 8-1 ビッグデータを分散保持

5.3 低価格トレーサビリティの例 (低コスト性の利用)

農家や消費者は高価なトレーサビリティシステムを使用しにくい. D5A ファイルを用いるとコストが 1/100 位なのでこの状

況を解消できる。Fig. 8-2 はメロンのトレーサビリティを示す。もちろんメロンにかぎらず、米や家電でも良い。このようなトレーサビリティは①ビッグデータの保持及び②その柔軟かつ高速な検索機能の確保に技術的な困難さがあったが、本件で示した新技術でこれらが解決されるので実現が容易になる。設備も簡単になるので低コストになる。

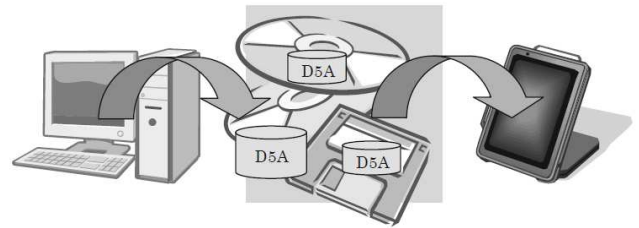


Fig. 9 汎用アーカイブ（ビッグデータのポータビリティ）

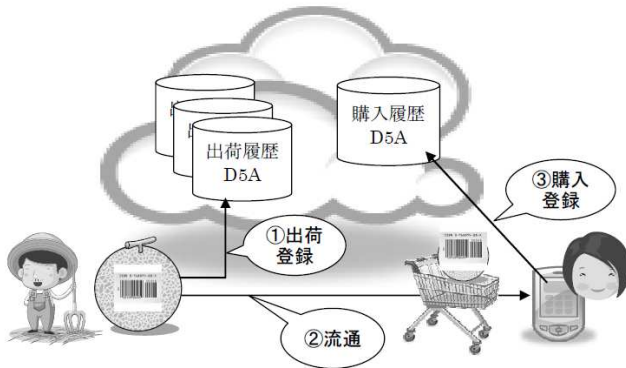


Fig. 8-2 トレーサビリティ

#### 説明

メロンは、2010年の国内生産量 188,100 トンで1個 1kg とすると、約 188,100,000 個が生産されていることになる。このデータを 10 年蓄積すれば、約 1,881,000,000 個となる。このデータサイズで多数の生産者、消費者を相互に結びつけようとする、従来技術ではコストがかかる。

図 1 では、①生産者が出荷をする際に出荷登録を行い、②流通経路を経て消費者に届き、③消費者は購入登録を行う。

このようなシステムを組むとこれまで高価だったシステムの価格を大幅に引き下げられる。

#### 5.4 オープンビッグデータ（分散性の利用） Fig. 7 ⑥

ビッグデータをホームページ同様、気軽に公開できるようになる。ホームページ同様、ホームオープンビッグデータが可能になる。

#### 5.5 汎用アーカイブ（ビッグデータのポータビリティ）

現実問題として多くのビッグデータのアーカイブはソフトウェアとセットになっており、そのソフトウェアは特定の OS の上でしか動かない。そのためアーカイブが再生できる環境を保持することは手間がかかる。本テクノロジーによる製品は様々な環境にインストールされ、過去のデータを容易に再生できる。また、ビッグデータのポータビリティにも利用できる。

(Fig. 9)



古庄 晋二

昭和 57 年 東京大学工学部産業機械工学科卒

(株)富士通ファナックを経て昭和 61 年 東京大学工学系研究科機械工学専門課程（修士）卒

平成 12 年 株式会社ターボデータラボラトリーを設立、代表取締役、現在に至る

Abstract

Distributable Big Data File over the Internet;  
Newly developed index enables operations over low latency media.  
Possible to operate over the Internet.

key words

database, index, big data, Internet, file, Internet file, distributed  
database