

# デジタル・トラストを支えるデータ中心 アーキテクチャーDataffinic Computing

## Data-Centric Architecture Dataffinic Computing to Support Digital Trust

田村 雅寿      吉田 英司      山田 幸二

---

### あらまし

近年、様々な現場で発生する大量データを蓄積して利活用することで、ビジネスの革新やイノベーションの創出を行う動きが本格化してきている。従来のデータベースで扱う構造化データに加えて、映像データをはじめとする非構造化データの利用が進んでいるが、非構造化データを効率的に利活用するためには、データ処理性能、コスト性能比、およびデータ管理の各要件が必要となる。富士通研究所では、大量データの利活用を支えるデータ中心アーキテクチャーとして、分散ストレージを用いて大量データの高速処理を実現するDataffinic Computingの研究開発を進めている。

本稿では、Dataffinic Computingの要素技術であるデータ近傍処理、大容量メモリ技術、高速シンクライアント技術を概説する。更に適用事例として、映像監視システムへの取り組みを紹介する。

### Abstract

Recently, there is a movement gaining momentum to create business reforms and innovations by storing and utilizing large volumes of data generated on various sites. In addition to the structured data handled in conventional databases, advancements are being made in the use of video and other unstructured. There are three requirements for efficient utilization of unstructured data: data processing performance, cost-performance ratio, and data management. Fujitsu Laboratories is working on the R&D of “Dataffinic Computing,” which uses distributed storage systems to achieve high-speed processing of large volumes of data, as a data-centric architecture that supports the utilization of massive volumes of data. This paper outlines data neighborhood processing, large-volume memory technology, and high-speed thin client technology, which are elemental technologies of Dataffinic Computing. It also presents an approach to video monitoring systems as an application example.

---

## 1. まえがき

近年、様々な現場で発生する大量データを利活用することで、ビジネスの革新やイノベーションの創出を行う動きが本格化してきている。これに伴って、従来のデータベースで管理されてきた顧客データやPOS（Point of Sale）データなどの構造化データに加えて、急速に増加している映像データやログデータなどの非構造化データの利用が進んでいる。非構造化データの利活用においては、AI（人工知能）分析などのデータ処理によって得られた情報を用いることで、効率的な分析が可能となる。

増加し続ける大量データを効率的に利活用するためには、以下の三つの要件が必要となる。

一つ目がデータ処理性能である。従来のデータセンターアーキテクチャーでは、データの蓄積部と処理部が分離され、その間をネットワークでつなぐ構成となっていた。そのため、処理するデータ量の増加に伴い、処理部へのデータ移動が処理性能のボトルネックとなってしまう。これを解決するためには、データの増加に対してスケラブルに高速処理を実現する必要がある。二つ目がコスト性能比である。データ量に応じて膨大な蓄積コストが必要となるが、高性能なストレージになるほどコストが高くなる。そのため、用途に合わせて性能やコストが異なるストレージを使いこなすことが重要となる。三つ目がデータ管理である。データ利活用においては、データそのものが他社との重要な差別化要素となる。そのため、データを流出させることなく、効率的にデータ処理を行うことが必要となる。

これらの要件を満たすために、富士通研究所では、大量データの利活用を支えるデータ中心アーキテクチャーとして、分散ストレージを用いたDataffinic Computingの研究開発を進めている。

本稿では、大量データの安心・安全かつ効率的な利活用のためのデータ処理基盤を実現するデータ近傍処理技術、大容量メモリ技術、および高速シンククライアント技術について概説する。更に、Dataffinic Computingの適用事例として、映像監視システムへの取り組みを紹介する。

## 2. Dataffinic Computing

富士通研究所で研究開発を進めているDataffinic Computingとは、「データ（data）」と「密接に結びついた（affinity）」という意味の造語である。Dataffinic Computingでは、前章で述べた大量データの利活用における要件に対して、蓄積されたデータを高速かつスケラブルに処理するためのデータ近傍処理技術、高いコスト性能比を実現するための大容量メモリ技術、およびリモートから効率的にデータ処理を行うための高速シンククライアント技術を用いている。これによって、大量データの安心・安全かつ効率的な利活用を可能とするデータ処理基盤を実現している（図-1）。本章では、これらの技術について述べる。

### 2.1 データ近傍処理技術

本節では、分散ストレージ上で大量データを高速に処理する、データ近傍処理技術について述べる。

非構造化データを利活用するためには、AI分析などのデータ処理を行う必要がある。しかし、データ量の増加に伴い、処理部へのデータの移動がボトルネックとなってくる。その一方で、ストレージ上でデータを移動させずに処理することで高速処理が可

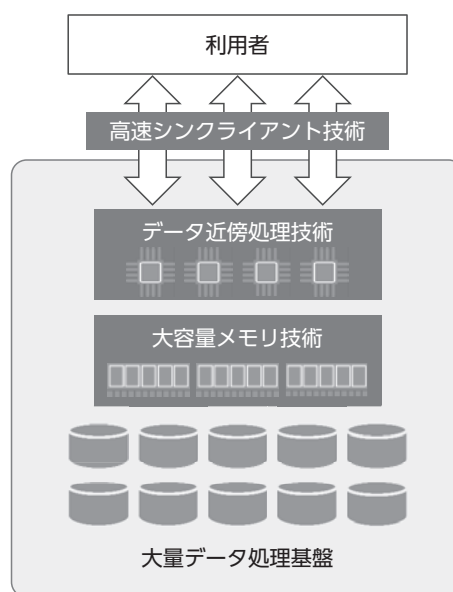


図-1 Dataffinic Computingの構成図

能となるが、データ処理の影響によるストレージ性能の低下が問題となる。非構造データの利活用において、これらの機能の両立が難しいことが課題となっている。

富士通研究所ではこの課題を解決するために、分散ストレージ上で割り当てるリソースを制御しながらデータを処理することで、高速なデータ処理とストレージの安定稼働を両立させるデータ近傍処理の研究を進めている。<sup>(1)</sup> 分散ストレージのノード上にデータ処理を可能とする実行基盤を構築し、データが存在するノード上でのデータ処理を実現する。

データ近傍処理の実行に当たっては、まず分散ストレージが持つデータ管理情報から、処理対象となるデータを蓄積しているストレージノードを算出する。次に、当該ノードにデータ処理アプリを配備して処理を実行し、結果をディスクに格納する(図-2)。一般に、処理対象となるデータと比較して、データ処理アプリおよび結果のデータ量は小さいため、データ移動のコストを抑えつつデータ処理を実現できる。

また、データ処理がストレージノード内に閉じて行われるため、ストレージノードが追加されて容量が増加してもノード間通信は発生しない。そのため、データ量の増加に応じて処理性能をスケールさ

せることが可能となる。更にストレージノード上では、本来のストレージ処理および各データ近傍処理が使用しているCPUやメモリなどのリソース量を常時監視している。ストレージ処理に割り当てるリソースが不足すると、データ近傍処理の一部を停止させてリソース開放することで、本来のストレージ処理の安定稼働を実現する。

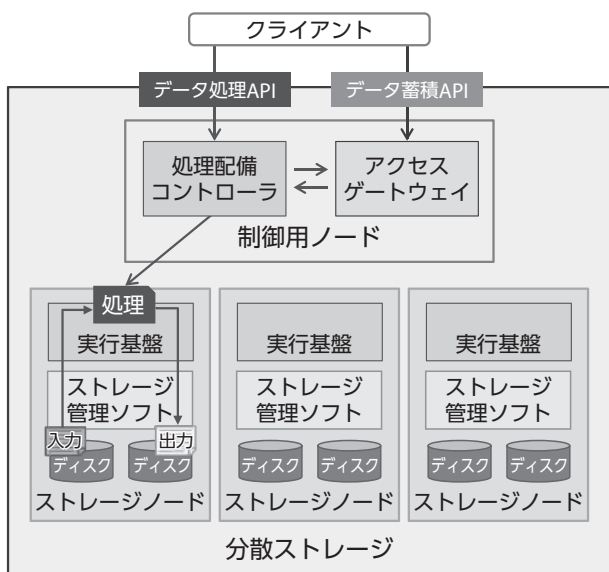
## 2.2 大容量メモリ技術

本節では高性能ストレージを実現する大容量メモリ技術について述べる。

解析すべきデータ量の増大に伴い、CPUやGPU (Graphics Processing Unit) でのデータ処理速度だけでなく、データ入出力にも更なる高速性が求められるようになってきている。このような課題に対してはインメモリ処理が有効である。インメモリ処理は、サーバ上のメインメモリに全てのデータを格納して処理を実行する方式であり、ディスクの入出力オーバーヘッドが削減されるため、桁違いに高速な処理速度を実現できる。しかし、従来メインメモリに用いられていたDRAM (Dynamic Random Access Memory) は高価であり、大量データの処理に必要な大容量のメモリを用意することがコスト的に困難であるという問題があった。

これを解決するために、富士通研究所では容量単価の安い不揮発性メモリをメインメモリとして扱うことで、インメモリ処理を高いコスト性能比で実現する方式を研究している。<sup>(2)</sup> 不揮発性メモリは、DRAMと比べて低コスト・大容量であるが速度は劣るため、その活用には速度劣化を抑える技術が重要となる。そのため、処理内容に応じてアクセス頻度の高いデータを積極的にDRAMに配置する技術や、不揮発性メモリの特性を考慮して性能を引き出すデータアクセス技術の研究を進めている。

これらの技術を用いることで、速度劣化なくメモリ容量を10倍に拡大できるため、より巨大なデータを扱う処理を高速化できるようになる。また、分散ストレージのノードに大容量メモリ技術を適用することによって、データ処理時にはアプリ処理用の大容量メモリとして、ストレージアクセス時には高速なディスクとして利用できる。これによって、ストレージと処理の効率的な一体化を実現する。



API : Application Programming Interface

図-2 データ近傍処理システム構成図

## 2.3 高速シンククライアント技術

本節では、クラウドにあるデータに対してリモートからの操作を可能とする、高速シンククライアント技術について述べる。

近年のデータ中心社会では、AI技術の発展やコンピューティング性能の向上によって、大量に蓄積されたデータを活用して新たな知見を取得し、新規ビジネスやイノベーションを創出する動きが活発になってきている。そのため、蓄積された業務データ自体が他社との重要な差別化要素となる。そこで、業務データの漏えい抑制や災害時の保全を図るために、業務データをクラウド上のサーバに集約し、クライアントからアクセスするだけで処理を行える、シンククライアント環境を導入する企業が多くなってきている。

シンククライアント環境では、手元にあるクライアントの操作情報のみをサーバに転送し、クラウドにあるサーバでデータ処理を行い、サーバのデスクトップ画像をクライアントへ転送している。しかし、動画再生やCAD設計業務では、デスクトップ画像の更新面積が大きく、更新頻度も高い。そのため、CPU負荷の増大によるデスクトップ画像の転送フレームレートの低下や、トラフィックの増加が発生し、十分な操作性を確保することが難しいといった問題があった。

この問題を解決するために、富士通研究所では仮想デスクトップ環境での高速表示技術「RVEC (レベック): Remote Virtual Environment Computing」を開発している。<sup>(3), (4)</sup> RVECには、以下のような特徴がある。

- ・デスクトップ画面の領域ごとに更新頻度を分析し、それぞれの領域ごとに最適な符号化方式で圧縮することによって、CPU負荷を低減し、かつトラフィックを削減する。
- ・TCP (Transmission Control Protocol) ベースとUDP (User Datagram Protocol) ベースの転送方式を切り替えることによって、パケットロスやネットワーク遅延が大きい環境でも良好な操作性を維持する。
- ・サーバ・クライアント間で利用可能なネットワーク帯域を推定し、帯域に応じて送信フレームレートを制御することによって、操作レスポンスを向

上させる。

これらの特徴によって、クラウドに蓄積されているデータをクライアントに転送することなく、データの処理や視覚化が可能となる。その結果、大量データの転送に要する時間の削減、およびデータ流出による情報漏えいの抑制が可能となる。

## 3. Dataffinic Computingの適用事例

Dataffinic Computingの適用が有効に機能する大量データとしては、映像データ、ログデータ、センサーデータ、ゲノムデータなどが考えられる。本章では一例として、大量の映像データが扱われる映像監視システムへのDataffinic Computingの適用事例を紹介する (図-3)。

国・自治体などの道路管理者は、事故防止や安全確認を目的として、道路に多数の監視カメラを設置している。これによって、映像確認や事故などの事象をリアルタイムに検知するとともに、撮影した監視映像をデータセンター (クラウド) に集約・蓄積している。蓄積された大量の映像データの活用方法として、過去に遡って子供や老人など行方不明者を検索したり、交通事故調査や犯罪捜査において詳細な状況を確認したりする映像データ分析がある。

映像データ分析では、ストレージシステムに蓄積された大量の映像データをサーバに転送して、処理を行う必要がある。映像蓄積時より高速な映像データの転送が必要となるため、ストレージシステムとサーバとの間のトラフィック量が増加する。そのため、映像データの転送が映像データ分析のボトルネックになり、分析処理の速度が低下する可能性がある。また、クラウドの映像データをクライアントに転送して確認や解析を行うと、映像データの転送によるユーザー操作性の低下とともに、映像データが流出し、プライバシーや情報漏えいの問題が発生する可能性が高くなる。

更に、クラウドに蓄積された映像データに対して行われるこれらの分析・確認・解析処理は、道路管理者の複数拠点から同時に行われる場合がある。これらの映像データ処理は大容量のメモリを必要とするため、複数のデータ処理が同時に行われた場合にサーバのメモリ容量が不足し、処理のレスポンスが



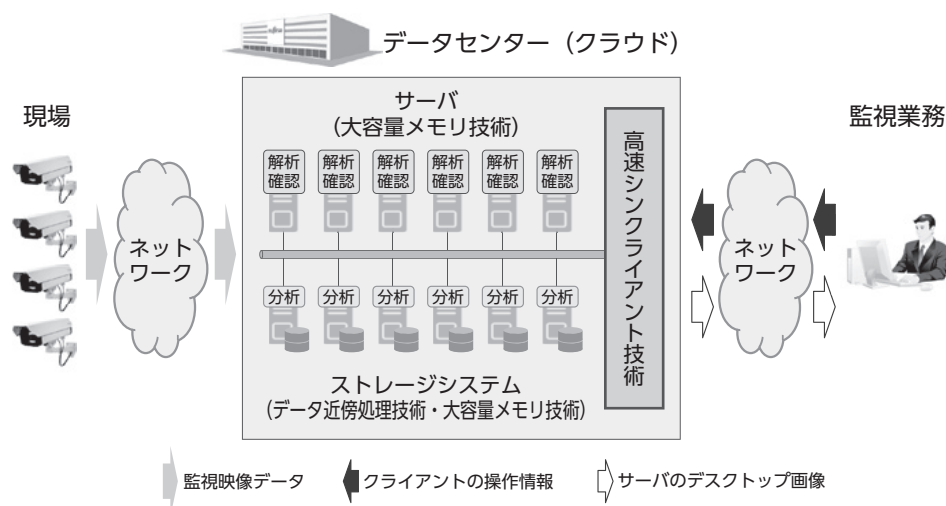


図-3 映像監視システム

低下する。そして最悪の場合には、システムがダウンする可能性がある。

このような問題を抱えた映像監視システムに対して、Dataffinic Computingを適用した場合の効果について、以下に説明する。

映像データ分析にデータ近傍処理技術を適用すると、ストレージシステムの映像データをサーバに転送することなくストレージシステム上で映像データ分析処理が可能となる。そのため、映像データの転送がボトルネックとならず、映像データ分析の処理速度が向上する。この効果を定量的に検証するために、50 GBの映像データの中から人や車などを検出する処理に要する時間を測定した。その結果、従来比10倍のデータ分析速度を実現できることを確認した。データ近傍処理によって、従来よりも多くの映像データを高速に分析できるため、交通事故の抽出や行方不明者の検出を迅速に行うことが可能となる。

また、映像の確認や解析に対して高速シンクライアント技術を適用すると、クラウド上にある映像再生・解析ソフトをクライアントから操作して、映像を確認・解析できる。その結果、監視業務を外部に委託した場合でも、クラウドに蓄積されている映像データをクライアントに転送することなく、映像の確認や解析を行えるため、映像データの流出を抑制できる。また、映像データの分析・確認・解析処理が同時に行われる場合に、大容量メモリ技術を映像

監視システムに適用すると、従来と比較して、10倍の容量を持つメモリを搭載できるため、サーバのメモリが不足することなく、十分な操作性と安定動作を確保できる。

このように、映像監視システムに対して、Dataffinic Computing, すなわちデータ近傍処理技術、高速シンクライアント技術、および大容量メモリ技術を適用することによって、システムの処理性能、セキュリティ、安定性を向上させ、信頼なビジネス環境を実現できる。

#### 4. むすび

本稿では、大量データの利活用を支えるデータ中心アーキテクチャーに関して、富士通研究所で取り組んでいる分散ストレージを用いたDataffinic Computingを紹介した。

今後は、大量データを更に効率よく蓄積・管理する技術の開発を進めるとともに、お客様と連携して実証実験を行うことで、お客様のビジネスを真に支えることのできる技術の素早い提供を進めていく。

本稿に掲載されている会社名・製品名は、各社所有の商標もしくは登録商標を含みます。

#### 参考文献

- (1) 富士通研究所：分散ストレージ上で大量データを高速処理する基盤技術を開発。

<https://pr.fujitsu.com/jp/news/2018/09/20.html>

- (2) 富士通研究所：物理メモリを追加せずインメモリ処理性能を向上させるメモリ拡張技術の実証実験に成功。

<https://pr.fujitsu.com/jp/news/2018/09/19-1.html>

- (3) 松田正宏ほか：モバイル環境向け高速シンククライアント技術Mobile RVEC（モバイルレベック）. FUJITSU, Vol.63, No.5, p.537-542 (2012).

<http://img.jp.fujitsu.com/downloads/jp/jmag/vol63-5/paper07.pdf>

- (4) 山田幸二ほか：GPU内蔵ハードコーデックによる高画質・低帯域を実現する画面転送方式. 第23回日本バーチャルリアリティ学会大会, 2018年9月.

<http://conference.vrsj.org/ac2018/program2018/pdf/11C-2.pdf>

## 著者紹介



**田村 雅寿** (たむら まさひさ)

(株) 富士通研究所  
プラットフォーム革新プロジェクト  
データシステム技術の研究開発に従事。



**吉田 英司** (よしだ えいじ)

(株) 富士通研究所  
ICTシステム研究所  
データシステム技術の研究開発に従事。



**山田 幸二** (やまだ こうじ)

(株) 富士通研究所  
ICTシステム研究所  
動画像符号化技術・動画像符号化応用  
技術の研究開発に従事。