

# オーガニックストレージシステム

## Organic Storage System

### あらまし

ユビキタス社会は、ITが社会の隅々にまで浸透し、データがあらゆる場所で発生する社会でもある。これらのデータを蓄積し管理することができれば、ビジネスの観点からも、あるいは社会のセキュリティという点からも大きな意義を持つだろう。しかし、ペタバイト級のデータが蓄積されてくると、その蓄積・管理は簡単ではない。

本稿では、まず大量データの蓄積・管理の課題について触れ、その課題を解決するためにはデータに自律性を与えることが重要であることについて述べる。また、実際にデータに自律性を与え、ペタバイト級のデータを蓄積・管理するためのアーキテクチャとして研究開発を進めているオーガニックストレージシステムについてその概要を説明する。

### Abstract

In the upcoming ubiquitous-computing world, we will see massive data flows generated by trillions of ubiquitous IT components surrounding our life. At any instant, this data will form a snapshot of human activity and will be very valuable from the point of view of business and safety in society. However, the size of such a snapshot could be in the multi-petabyte region, and the management of so much data is not easy. In this paper, we describe the types of problems that will be encountered when dealing with multi-petabyte amounts of data and how we can solve them with autonomic data management technology. Then, we introduce a novel storage architecture called the organic storage system for realizing autonomic data management.



武理一郎（たけ りいちろう）  
ITコア研究所 所属  
現在、データベースシステム、ストレージシステムの研究開発に従事。

## まえがき

ユビキタス社会は、ITが社会の隅々にまで浸透し、データがあらゆる場所で発生する社会でもある。発生する大量のデータを蓄積し管理することができれば、ビジネスの観点からも、あるいは社会のセキュリティという点からも大きな意義を持つだろう。

しかし、現実にはこれらの大量データを蓄積することは容易ではない。日本国内だけを考えても数千万台の携帯電話やPDAがデータの生成源として機能する。これらの機器が毎日1Mバイト（M=メガは10の6乗）のデータを生成すると仮定すると、その総量は毎日数十Tバイト（T=テラは10の12乗）というものになる。1箇月で1Pバイト（P=ペタは10の15乗）以上、1年では10Pバイト以上となる。

しかも、この大量のデータはどんどん蓄積されていく性質のもので、1年経ったから消せるというものではない。ユビキタス社会が生成するデータは、ビジネスシーンで生成されるにしても、個人生活の中で生成されるにしても極めてパーソナルな性質のものであり、そこには「人生の記録」「個人・法人の自分史」といったセンチメンタルバリューが付きまとう。そのようなデータを消すことに、人は強い抵抗を覚えるだろう。つまり、ユビキタス社会は、大量に生成されるデータが蓄積され、それを消すことができずに背負い続けることになる社会でもある。

本稿では、このようなユビキタス社会で必要とされるストレージシステムとはどのようなものなのか考えてみたい。

## 大量データ保管の課題

大量データ保管の難しさは、データがその入れ物の捕われの身となることをいかに避けるかにある。

データが少量のうちは、データを例えばあるハードディスク装置から別のハードディスク装置へ物理的に移動させることは難しい作業ではない。また、データのフォーマットを例えばMPEG-2からMPEG-4へと論理的に変換することも容易だろう。

ところが、データの量が大量になってくるとこのような作業は容易ではなくなるばかりか、実施不能に近い難しさを伴うことになる。例えば、1Pバイトのデータがあるストレージシステム上に存在する場合、このデータを別の1Pバイトのストレージ

システムに移動することを想像してみよう。一般的なデータ転送性能である10Mバイト/秒という速度で移動すると仮定すると、この移動には3年以上かかることになる。

ITは5年も経てば旧世代のものとなってしまう、システムの入替えが必要となる。その際には旧システムに存在したデータを新システムへ移動しなければならない。サイズが膨らみ流動性を失ったデータは、古いハードウェアに閉じ込められ、古いデータフォーマットのままで時を過ごすうちに、保守部品がなくなり、読み出してくれるプログラムすらなくなってしまい、その入れ物とともに古びていく。

そのようなことにならないために著者らが研究を進めているのがオーガニックストレージシステムである。

## オーガニックストレージシステム

データが特定の入れ物に捕われることなく、入れ物から入れ物へと自由に渡り歩きながら、常に新鮮さを保つためには、データに自律性を与えることが重要である。

現在のITシステムでは、親分はデータを処理するプロセッサである。データはプロセッサに処理してもらって受身の存在であり、プロセッサ側から読出してもらわないと何もできない。

これに対し著者らはデータが自律性を持ち、自分の意思で移動することを可能にしようとしている。この世界では親分はデータ自身で、プロセッサはそのデータにサービスする立場として働くことになる。

データにこのような自律性を与えることができれば、データはコンピュータが古くなってしまいう前に新しい場所を自分で見付けて引っ越すことができるようになり、また、自分を利用するプログラムがなくなってしまう前に、新しいプログラム用に変換してもらおうなどしていつまでも古びずにいることが可能となるだろう。

実際にデータに自律性を与えるには、以下の三つの点が重要となる。

## (1) データの移動性能が高いこと

いくらデータが自律的に移動できるといっても、移動性能が低くては、大量のデータを実用的な時間内で移動させることができない。この性能を高めることが重要である。

## (2) 仮想化が実現されていること

ここでの仮想化とは、データの格納場所を利用者から隠蔽することである。データが自分の都合で勝手に移動してもそのデータを利用するサーバがデータを見失わないように仮想化を実現しておくことが必要である。

## (3) 自己記述性が実現されていること

自己記述性とは、そのデータがどんなデータなのかをデータ自体が説明できることを指している。例えば、あるデータが遺伝子研究成果の場合、どの遺伝子に関する情報か、解析するにはどのプログラムが必要かなどのメタデータ（対象データが何処に格納されているかを示すデータ）がきちんと付いていれば、そのデータはどこに行っても迷子になることなく将来にわたって利用できる。

このような考えのもとに、開発を進めているのがオーガニックストレージシステムである。オーガニックストレージシステムはモジュラ構成を持つストレージシステムである。各モジュールはパソコンのようなものであり、多数のモジュールが全体として一つのストレージとして機能する。多数のモジュールが一つのシステムとして動作する様子を、多数の細胞が一つの生体を構成する様子になぞらえて、「有機体のような」という意味でオーガニックという言葉を使っている。

オーガニックストレージシステムと従来のストレージシステムのアーキテクチャの比較を図-1に示

す。従来のアーキテクチャでは多数の記録メディアを比較的少数の制御装置で束ねて管理している。制御装置は外部のサーバなどからの読み書き要求に応じてサービスを行う受身の存在である。そのため、例えば、RAID装置のデータをMTライブラリ装置に移動しようとするとき、ストレージシステムの外部にサーバを用意し、このサーバを通して移動作業を行うことになり、移動性能はこのサーバに制限されることになる。

これに対してオーガニックストレージシステムでは、モジュール群が互いに自律的に（外部からの指示によらずに）通信を行う機能を備えている。このため、データ移動をモジュール群が並列に行うことができるため（図中の横向きの矢印群）、モジュールの数を増やすことで移動性能を高めることができる。

高いデータ移動度を実現するためには、容量あたりの通信・処理性能を高くするしかない。このためには、従来よりも少数のメディアに対して一つの制御装置を割り当て、全体としては従来よりも多数の制御装置を並べることになる。オーガニックストレージシステムのモジュラ構造は、粒度（各モジュールが備える記憶容量）を小さくすることにより、容量あたりの通信・処理性能を十分に大きくすることを可能としている。

オーガニックストレージシステムでは、上位の制御装置を置いて全体を管理する方式ではなく、多数

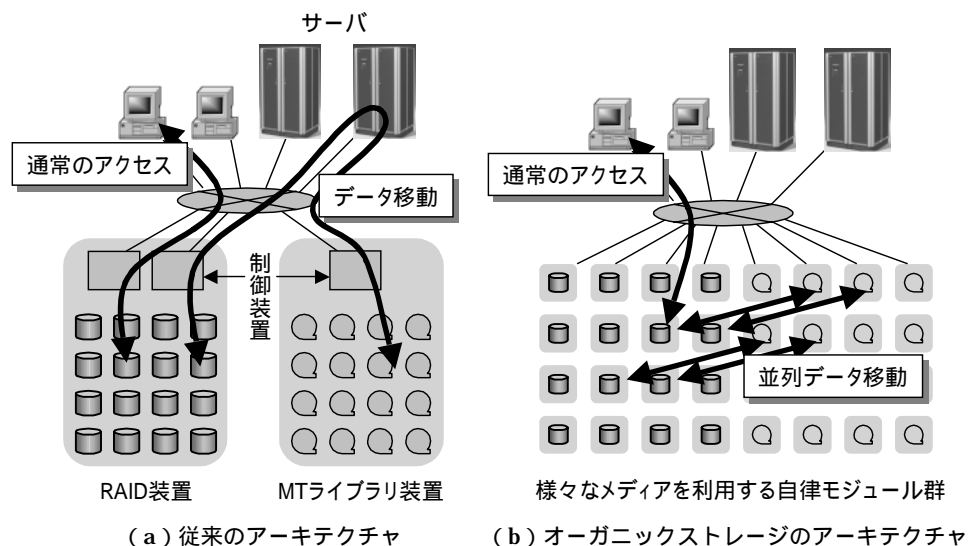


図-1 オーガニックストレージシステムと従来ストレージシステムの比較

Fig.1-Diagrammatic comparison between organic storage system and conventional storage system.

のモジュールのそれぞれが自律的に全体の一部として機能する方式を採っている。上位の制御装置を置く方式は、実現は容易であるがスケラビリティに制限が発生する可能性が高いからである。

利用者があるディスクボリュームを利用しようとすると、利用者のディスクドライバはオーガニックストレージシステムのメタデータサーバに問い合わせさせてメタデータを受け取り、それが指し示す自律モジュールに読み書きコマンドを投げる。コマンドを受け取ったモジュールはコマンドに応じた読み書きのサービスを行う。

メタデータサーバに存在するメタデータはキャッシュであり、間違っていてよい。メタデータの実体は、各モジュールにおいて対応するデータと組にして保全される。利用者に渡るメタデータが間違っていればモジュールからは利用者にエラーが返り、再度、新しいメタデータがメタデータサーバに要求される。このような頑健なプロトコルにより、メタデータサーバがシステムの信頼性・スケラビリティの妨げになることを防ぎ、柔軟な仮想化を実現することを可能としている。

## 関連分野の研究活動

本研究と関連のある分野の研究活動について述べる。

コンピュータアーキテクチャの観点からは、オーガニックストレージシステムは、ディスクなどのデータ格納装置の近傍に処理装置を置き、データ処理をサーバ側ではなくデータ格納装置側である程度行わせるという取組みの一つと言えるだろう。

このような取組みとしては、データベースマシンや高機能ディスクの研究開発に関する長い歴史がある。比較的新しい研究としては、高機能ディスク装置やデータベースマシン<sup>(1)-(5)</sup>があり、高機能ディスク装置の一種であるオブジェクトディスク<sup>(4)</sup>については標準化作業も行われている<sup>(6)</sup>

これらの研究はデータの近傍の処理能力を主にデータベース演算やコンテンツレベルの高度なアクセス機能の実現に使おうとするものである。これに対して著者らは、その処理能力をまずはデータ再配置などの運用管理のために使うことをねらっている。ただし、データベース演算などの高度な処理も重要であると考えており、今後はそのような処理をオー

ガニックストレージシステムに取り込んでいくことも検討していきたい。

## む す び

現在のオーガニックストレージシステムは、本稿で挙げた高いデータ移動性能、仮想化、自己記述性という三つの目標の内、最初の二つを実現している。高いデータ移動性能と仮想化の組合せは、データアクセスを保証したままでデータ再配置やバックアップ・リストアを、高速・自律的に実行することを可能とし、ストレージシステムの運用管理の様々な局面でブレイクスルーをもたらすものと考えている。目標の3番目に挙げた自己記述性についてはまだ実現できていない。

今後はセキュリティ機能（暗号化、改ざん不能性の付与）などの上位機能の開発や、データグリッドやデータのライフサイクルマネジメントといった文脈の中での自己記述性に関する検討を予定している。

## 参考文献

- (1) M. Kitsuregawa et al.: Functional Disk System for Relational Database. Proc. of 3rd International Conference on Data Engineering, 1987, p.88-95.
- (2) David J. DeWitt et al.: The Gamma Database Machine Project. *IEEE Transactions on Knowledge and Data Engineering*, Vol.2, No.1, p.44-62 (1990).
- (3) M. Kitsuregawa et al.: The Super Database Computer (SDC): System Architecture, Algorithm and Preliminary Evaluation. Proc. of 25th Hawaii International Conference on System Sciences (HICSS-25), 1992, p.308-319.
- (4) Garth A. Gibson et al.: Network Attached Storage Architecture. *Communications of the ACM*, Vol.43, No.11, p.37-45 (2000).
- (5) H. Yokota: Autonomous Disks for Advanced Database Applications. Proc. of International Symposium on Database Applications in Non-Traditional Environments (DANTE'99), 1999, p.441-448.
- (6) SNIA OSD Technical Work Group.  
[http://www.snia.org/tech\\_activities/workgroups/osd/](http://www.snia.org/tech_activities/workgroups/osd/)