

Big Dataにより新たな価値を創出する
次世代ITプラットフォームコンテンツクラウドを実現する
仮想ファイルプラットフォーム

—Hitachi Virtual File Platform—

Virtual File Platform Realizing Content Cloud

小笠原 裕 矢川 雄一 里山 元章
Ogasawara Hiroshi Yagawa Yuichi Satoyama Motoaki

映像やメールなど非構造化データを中心としたコンテンツデータの伸びが著しい。非構造化データはファイル形式で格納され、これまでファイルデータの増大に対してはNASの導入で対応することが一般的であったが、近年の急激なデータ量の伸びに伴って、管理の煩雑さが課題となってきた。

日立グループは、これを解決するためにコンテンツクラウドというコンセプトを打ち出し、従来のNASを超える仮想ファイルプラットフォームHitachi Virtual File Platformを提供している。これにより、拠点内・拠点間でのインフラ統合や、データセンターへのファイルの自動集約などが可能になり、装置リソース、人的リソースの有効活用と、管理工数の削減を実現した。また、この製品を核とするクラウドサービスの提供により、クラウド導入を容易にしている。

1. はじめに

企業が扱うデータの量は、原則として毎年増大傾向にある。しかし、近年はデータの種類によってその保管の傾向が異なってきており、ファイルなど非構造化データの伸びが著しい。

調査会社であるIDC Japan株式会社の分類によれば、データは(1)構造化データ、(2)複製データ、(3)非構造化データ、(4)コンテンツデポの4種類に分類される。このうち構造化データは、データベース、ERP (Enterprise Resource Planning)、SCM (Supply Chain Management)、CRM (Customer Relationship Management)などのデータである。また、複製データは、レプリケーション、バックアップ、アーカイブのデータである。これらのデータは、主としてデータセンター内で生成され、格納される特徴を持っている。

それに対して、非構造化データ、およびコンテンツデポは、拠点ごとにNAS (Network Attached Storage) やファイルサーバに格納されて利用されるケースが多い。非構造化

データは、各種テキスト、画像、オーディオ、ビデオ、電子メール、CAD (Computer-aided Design) ドキュメントなどが含まれ、通常はファイル形式で保管される。また、コンテンツデポは、ニュース、株式相場、金融情報、技術情報、学術的情報、価格などの比較情報、ビデオ、音楽、ゲームなど、サービスプロバイダーによって提供されるデジタルコンテンツのことである。これらのデータも非構造化データが主体であり、専らファイル形式で生成、格納される。

非構造化データは、容量ベースで2009年～2014年の間に年率で約65%の伸びがあると予測されている(図1参照)。これにより、これまで各拠点で導入していたNASやファイルサーバが管理できる適正な容量を超え、運用管理の煩雑さが事業に与える影響を無視できない状況になる日が近いと予想される。

このような中、日立グループはこれら非構造化データの効率的な管理を可能とし、システムの導入コスト、運用管

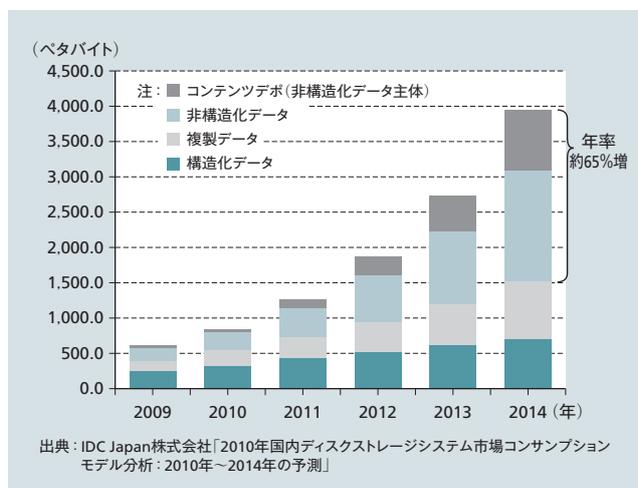


図1 | 国内データタイプ別ディスク使用容量の変化
2009年～2014年の間、年平均成長率が拡大すると予測されている。

理コストを削減するため、コンテンツクラウドというコンセプトを打ち出し、その実現に向けてファイルの仮想化技術を実装した製品Hitachi Virtual File Platform（以下、VFPと記す。）の提供を開始した。

ここでは、コンテンツクラウドの概要と、従来のNASにとどまらないクラウド対応機能を備えたVFPの特色および機能について述べる。

2. 急増するデータを管理するコンテンツクラウド

2.1 概要

コンテンツクラウドを端的に表現すると、「データの属性や種類を意識せず、安全・確実・快適に保管・活用できるシステム」である。

これまで、ファイル、メール、画像、映像といった多種類のコンテンツ（ファイル）については、それぞれにファイルサーバやNASを準備し、個別に管理したり、バックアップを作成するなどしており、その手順が煩雑になっていた。コンテンツクラウドでは、これらのデータはクラウド環境で一元的に管理・運用され、バックアップなどの作業も集約できる。このため手順が簡易化されるとともに、それぞれのサーバに分散し、むだが生じていたリソースも一元化され、効率よく利用できることになる（[図2](#)参照）。

2.2 コンテンツクラウドの段階的な実現

従来の個別運用・管理からのコンテンツクラウドへの移行にあたっては、一足飛びに採用する方法のほかに、段階的に採用していくアプローチもあり得る。つまり、まずは拠点内でサーバやNASを集約・統合し、次に運用管理を自動化し、データ管理を階層化することによってクラウドに対応していくという方法である。これにより、ユーザーは自社の環境を効率的に集約しつつ、容易にクラウド環境へ移行することができる。

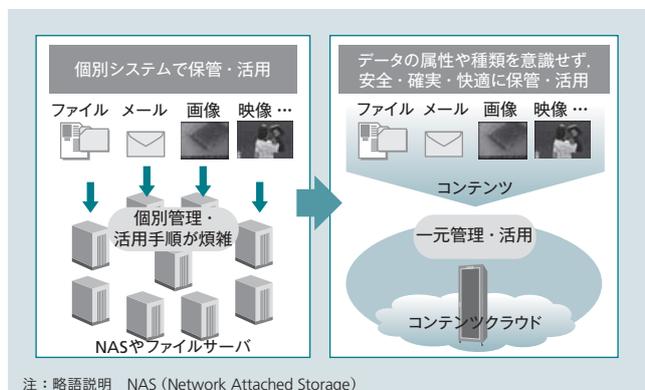


図2 | コンテンツクラウドのイメージ

従来はデータが増加するたびにNASを追加し、アイランド化を招いていたが、コンテンツクラウドではアプリケーションからのデータアクセスを集約できる。

3. コンテンツクラウドを実現する仮想ファイルプラットフォーム

日立グループは、VFPによってコンテンツクラウドの段階的な実現を可能としている。それぞれに対応する機能について以下に述べる。

3.1 集約・統合を可能とする高拡張性

多数のNASやファイルサーバを集約・統合する装置には、十分な性能と容量が求められる。VFPは、従来モデルの約2倍の高性能を実現した。また、容量については、最大15ペタバイトの容量と、従来比60倍となる最大1ペタバイトのファイルシステムサイズを実現し、NASやファイルサーバの統合に十分な許容量を備えている。

さらに、容量仮想化機能により、必ずしも設定したファイルシステムサイズに相当する容量のディスクを最初から準備する必要をなくしている。Hitachi Adaptable Modular Storage 2000シリーズのHitachi Dynamic Provisioning機能との連携で、ストレージプールを複数のファイルシステムで共有することができる。ファイルシステムに書き込みがあると、プールから自動的に容量を割り当てるため、増加するデータに対してタイムリーな容量の提供ができる。また、実際の搭載ディスクの容量が使用する容量に近づくとシステム管理者に通知されるため、ディスクの購入と増設もタイムリーに行うことができ、ストレージコストを適切に維持できる。部署の移動などで大量にファイルを削除したり、ファイルシステムを削除した場合にも、削除で生じた空き領域を解放し、他のファイルシステムに割り当てることが可能である（[図3](#)参照）。

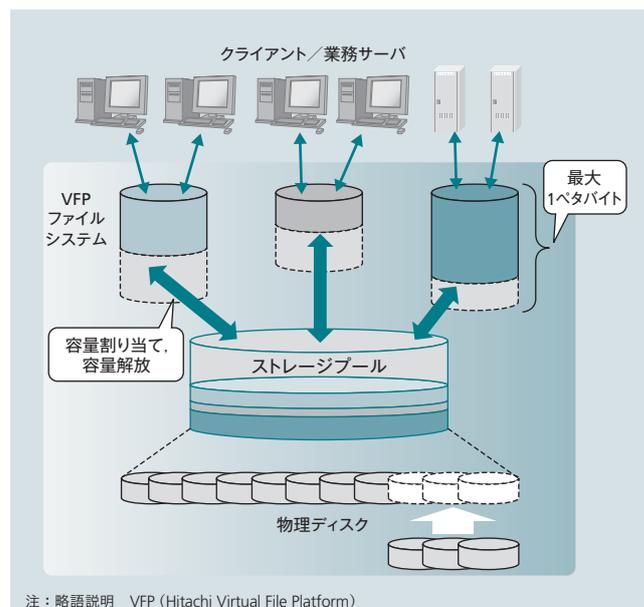


図3 | 容量仮想化機能による効率向上

あらかじめ仮想的な容量を十分大きく設定しておくことで、事前の容量設計の手間を削減できる。また、ファイルシステムが使用する物理容量は自動的に割り当てることができる。

3.2 アクセスプロトコルの異なるストレージを1台に集約

VFPは、一般的なNASのプロトコルであるNFS (Network File System) やCIFS (Common Internet File System), FTP (File Transfer Protocol) に加えて、ブロックストレージのアクセスプロトコルであるファイバチャネル, iSCSI (Internet Small Computer System Interface) にも対応している。Hitachi Adaptable Modular Storage 2000シリーズと運用管理画面を共通化し、同一の画面上でVFPとHitachi Adaptable Modular Storage 2000シリーズを管理できる。これにより、さまざまなアクセスプロトコルを持つアプリケーションからのアクセスを1台の装置に集約し、ストレージコストと運用管理コストの双方を低減できるようになっている。

3.3 外部ソフトウェアとの連携

VFPは、VMware^{※)}やHitachi Command Suiteと連携し、より効率的な運用管理を可能にしている。VMwareとの連携では、プラグインを提供し、VMware vSphere^{※)}クライアントのGUI (Graphical User Interface) から、直接VFP上にデータストアを作成し、VM (Virtual Machine) のイメージファイルを格納できるほか、バックアップやリストアも実行できる。さらにバックアップのスケジューリングも可能となっており、ストレージを意識せずにVMwareを運用できる。

3.4 システム移行を容易にするオンラインデータ移行機能

VFPでは、既存のNAS装置やファイルサーバから、オンラインでデータの移行を行う機能を提供している。ファイル仮想化機能の適用により、VFPをLAN (Local Area Network) 上に接続し、移行設定を終えた後、ユーザーは移行を開始した直後からすべてのデータにアクセスが可能である。この場合、データは移行前、移行後を問わない。このため、これまでは移行が完了してから再開していた業務を移行直後から再開することができ、システムの停止時間を大幅に削減できる。

3.5 Cloud on-Rampソリューション

VFPでは、ファイル仮想化機能により、拠点のVFPに書き込まれたデータを、自動的にセンターのHitachi Content Platform (以下、HCPと記す。)にバックアップ/アーカイブすることができる(図4参照)。

拠点ごとのVFPが、Cloud on-Ramp (クラウドへの入り口)としてセンターと統合し、あらかじめ管理者によって

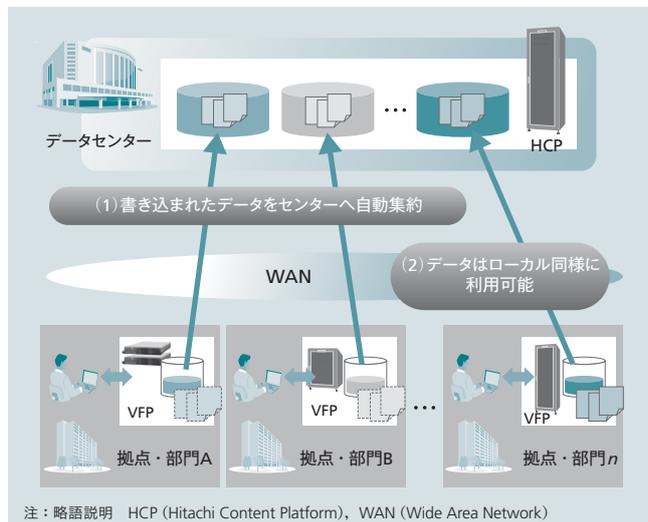


図4 | Cloud on-Rampによるデータの自動集約

自動的にセンターへバックアップ/アーカイブを行う。拠点・部門のデータをファイル仮想化することで、ユーザーは物理的な格納場所を意識する必要がない。

設定されたポリシーに従って、VFP上のデータをHCPに転送する。一定期間VFP上のデータにアクセスがなかったファイルなどは、ポリシーに基づいてスタブ化され、実体が削除されるため、VFP側のディスク容量を有効に活用することができる。常にアクセスするような重要度の高いファイルは拠点側のVFP上に配置して高速なアクセスを確保する一方、アクセスされなくなっても保存が必要なデータはセンター側のHCPに格納することで、拠点側ではディスクの容量を気にせずにシステムを使い続けることができる。なお、いったんHCP上に格納され、スタブ化されたファイルであっても、アクセスがあった場合はVFP上にリコールされるため、その後は再度拠点上での高速アクセスを行うことができる。リコールされたデータがその後アクセスされなくなった場合は、通常のデータと同様、ポリシーに従って再度スタブ化され、VFPの容量を効率よく使用することができる。

また、Cloud on-Rampソリューションでは、拠点災害などによってVFPが使用不能になった場合、VFPをリプレースした後、直ちに業務を再開することができる。これは、VFPリプレースの後、まずセンター側のHCPから管理情報を回復し、要求されたデータから順にオンデマンドでVFP側へデータを復元していくことによるものである。これにより、ユーザーはすべてのデータがHCPからリストアされるのを待つことなく、迅速に必要なデータにアクセスできる(図5参照)。

また、センター側のHCPも別センターのHCPにデータを二重化することができるため、拠点側とセンター側双方の災害に対応が可能となっている。

Cloud on-Rampソリューションでは、各拠点に対して

※) VMware, VMware vSphereは、VMware, Inc.の米国およびその他の地域における登録商標または商標である。

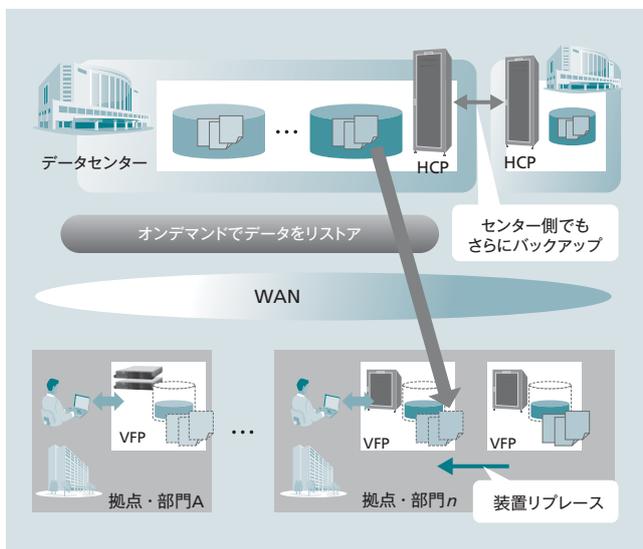


図5 | 拠点災害時の迅速なデータアクセス

Cloud on-Rampソリューションでは、拠点災害の際も装置リプレイス直後からデータへアクセス可能である。

センター側で個別の領域を割り当てることが可能なマルチテナント機能を利用できる(図6参照)。例えば、同図において、拠点Bから拠点Cのバックアップデータへのアクセスを禁ずることにより、拠点間でのセキュリティを独立させることができる。これは、社内で部署ごとのセキュリティを分離したい場合はもちろん、Cloud on-Rampによって顧客にクラウドサービスを提供しようと考えているプロバイダーにとっても便利な機能である。センター側のデータは、セキュリティを確保しつつ、一括して重複排除や圧縮を行うことにより、容量の利用効率を向上できる。さらに、拠点側でファイルに改竄(ざん)防止属性を設定した場合、センター側に格納される際にもその設定が引き継がれるため、業務ごとの安全・安心なデータ保管と、容量の使用効率向上の双方を一度に実現することができる。

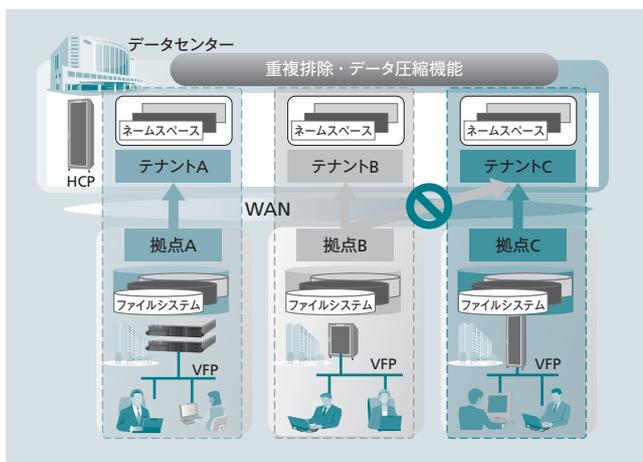


図6 | マルチテナント機能によるセキュアなデータ管理

センター側に配置するHCPのテナントに各拠点のVFPを割り当てることで、独立したセキュリティを確保する。

4. おわりに

ここでは、コンテンツクラウドの概要と、従来のNASにとどまらないクラウド対応機能を備えたVFPの特色および機能について述べた。

VFPは、爆発的に増大するコンテンツデータをファイル仮想化によってコンテンツクラウドで管理することができ、装置リソース、人的リソースの有効活用と、管理工数の削減を実現する。さらに、この製品を核とするクラウドサービスを提供しており、ユーザーのクラウド導入のハードルを下げることを可能としている。

日立グループは、引き続きコンテンツクラウドを実現する機能の追加・性能の向上を行い、顧客のデータ活用における効率の向上、コストの削減に寄与していく。

執筆者紹介



小笠原 裕

1995年日立製作所入社、情報・通信システム社 RAIDシステム事業部 ファイルストレージ開発本部 アーキテクチャ開発部 所属
現在、ファイルストレージ製品の企画・マーケティング業務に従事
情報処理学会会員



矢川 雄一

1991年日立製作所入社、情報・通信システム社 RAIDシステム事業部 ファイルストレージ開発本部 アーキテクチャ開発部 所属
現在、ファイルストレージ製品の企画・マーケティング業務に従事



里山 元章

1987年日立製作所入社、情報・通信システム社 RAIDシステム事業部 ファイルストレージ開発本部 ファイルストレージ開発設計部 所属
現在、ファイルストレージ製品の開発・設計業務に従事