

# 急増する非構造化データに直面する企業 IT における スケールアウト型オブジェクトストレージの使いどころ

近年、企業において非構造化データが急増しています。これらの非構造化データの多くは、研究開発における実験データや製造設計関連データをはじめ、継続的な投資の結果として蓄えた貴重な企業資産です。しかし、それは IT 部門が管理できる範囲や量を越え、ユーザ部門が独自に管理調達するストレージ装置に保管され、企業内でサイロ状になりはじめています。また、制約の多い社内ファイルシステムに満足できない社員は、個人契約のクラウド (BYOC: Bring Your Own Cloud) を業務で利用し、企業の管理が及ばないところで、社内外とデータ共有しています。増え続けるデータのバックアップやアーカイブ、DR (災害復旧) や BCP (事業継続計画) 対策も企業にとって頭の痛い問題です。

こういった課題の解決には、オンプレミスでスタンドアロン利用を前提とした従来型のストレージ装置ではなく、「次世代ストレージ」とも言われる、スケールアウト型オブジェクトストレージが適しています。本ホワイトペーパーでは、このスケールアウト型オブジェクトストレージと、その代表的な製品である「CLOUDIAN HyperStore」の特長と仕組み、その企業 IT における使いどころについて紹介します。

## 1 非構造化データ管理の課題

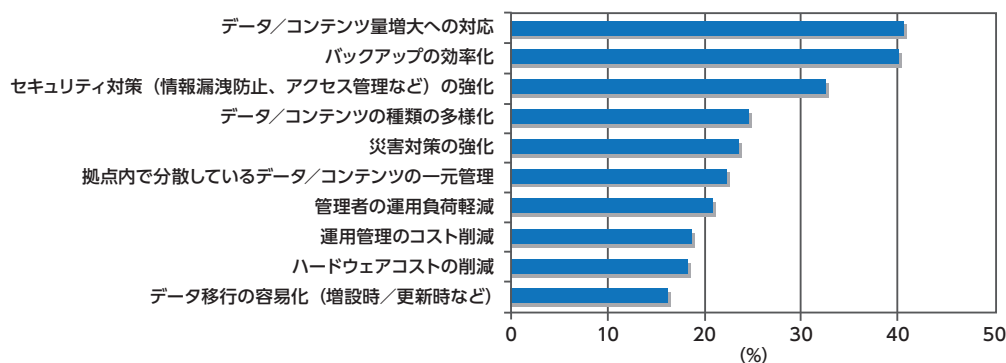
“企業 IT は非構造化データの増加とともに、その膨大な量が引き起こす課題への対処が求められる”

IDC が 2014 年 10 月に国内企業 300 社に対して実施した調査では、現在管理している非構造化データ向けストレージ容量を 100TB (テラバイト) 以上とした回答者は 13.0% であり、国内企業で大容量の非構造化データ向けストレージを管理する企業はまだそれほど多くない状況と言える。しかしながら、今後 2 年間の非構造化データ向けストレージ容量の増加を 10% 以上と見込む回答者が 34.7% となり、51% 以上の大

幅な増加を見込む回答者も 10% を越えている。

(同調査における非構造化データ管理の課題の) 上位 10 項目を見ると、非構造化データの増加に伴うデータ / コンテンツの増加への対応や効率的な管理、また、データ保護への対応が国内企業における非構造化データ管理の課題となっていることが分かる。

非構造化データ管理の課題



n = 300

Notes:

『2014 年 国内ファイル/オブジェクトストレージ市場予測と競合分析: 台頭する Software-Defined Storage の影響を探る (IDC# J15531001、2014 年 12 月発行)』を基に作成

上位 10 項目を抜粋

複数回答

Source: IDC Japan, February 2015

出典: IDC White Paper「クラウドファン、エンタープライズへの普及戦略を強化: オブジェクトストレージのソリューション展開を拡大」Sponsored by CLOUDIAN (2015 年 2 月)

この IDC の調査結果が示すように、非構造化データの急増とともに、多くの企業が新たなデータ管理の課題に直面しています。クラウドファンがお客様からご相談を受ける実務レベルの課題は、以下のように深刻です。

- 非構造化データは、ユーザ部門が独自にストレージ装置を調達しサイロ状に保管され IT 部門の管理が及ばない
- 各部門でバックアップ・アーカイブデータをテープ等で長期保管はしているものの、実際には活用しない「死蔵データ」となっ

ている

- ファイル等の保管に各ユーザ部門に割り当てられたストレージ容量は充分ではなく、また外部からのアクセスも制限されることから、複数の個人契約クラウドが業務利用されている
- 企業レベルで不特定多数の利用者を対象としたパブリッククラウドストレージサービスを利用しているが、全てのデータを外部に預ける訳にはいかない
- DR や BCP のために複数拠点にデータをバックアップする必要性はあるが、手間やコストを考えると本格的に展開できない

### 非構造化データは企業内でサイロ状に増加中

- 定型・構造データを扱う業務システムと異なり、ユーザ部門のデータストレージには、非構造化データが急増しているのでは？

情報システム部門		ユーザ部門					
業務系	情報系	研究開発	設計製造	運用	マーケティング	営業	管理
DB データ	ファイル	実験データ	設計図面	モニタ	画像	顧客情報	電子書類
	メール	製造品質	運用監視	動画	顧客情報	顧客情報	電子書類
	アクセスログ	ログ	ログ	写真	SNS ログ		

「事業部門のストレージ調達は、事業部門に任せています」(複数の企業 IT 担当者)

「事業部門からストレージ容量を制限すると言われていました」(製造業 IT 担当者)

「情報システム部門では既存システムの運用 / 保守に多くの時間が割かれている」(Japan IT Market 2015 TOP10 Predictions J15990181, December, 2014)

© Copyright 2010-2015 Cloudian KK. All rights reserved.

### 長期保管データは、実際には使わない、使えない


- テープ等の低コストの記録媒体にデータを長期保存はするものの、実際には「死蔵データ」になっているのでは？

「過去の記録を所轄官庁に提出する必要がありました。倉庫に行って、テープからデータを読み出すのはたいへんでした」(外資系金融 IT 担当者)

「商品売上の季節変動をビッグデータ分析しますが、過去10年間のデータをテープから読み出すだけで、かなりの時間を使いました」(コンサルティング担当者)

「テープが劣化するので3年毎に書き換えています」(製造業 IT 担当者)


「過去データはテープに記録しています。使うことではなく、保管することが目的です」(メディア IT 担当者)



© Copyright 2010-2015 Cloudian KK. All rights reserved.

### 個人契約クラウドを業務利用しデータを外部保存(BYOC)

- CIOの管理が及ばない個人契約クラウド (Bring Your Own Cloud) を業務に活用している社員が増えているのでは？

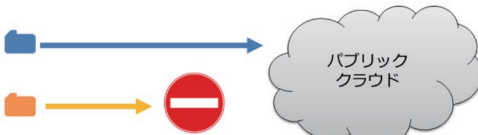


「社内調査をしたら、300人以上が個人契約クラウドを使い開発業務をしていました。実際にはもっといると思います」(米国大手ソフトウェア開発企業 CIO)

© Copyright 2010-2015 Cloudian KK. All rights reserved.

### パブリッククラウドに全てのデータは預けられない

- 柔軟性が高く、経済的なパブリッククラウドの活用を始めたものの、「外部には預けられない」データがあり、自社内システムと重複管理をしているのでは？




「全てのデータをクラウドには預けません」(多数 IT 担当者)

「クラウドの利用は禁止されています」(官公庁関係者)

© Copyright 2010-2015 Cloudian KK. All rights reserved.

### 複数拠点におけるデータ保護と管理の手間や費用が重荷

- 災害に備えた地域冗長の必要性は認識しているものの、複数拠点にデータ保護するコストが重荷となり、プロジェクトの優先度が下がっているのでは？



「日本だけではなく、海外拠点も含めて3重にバックアップするプロジェクトの最終承認に時間がかかっています」(機器製造 IT 担当者)

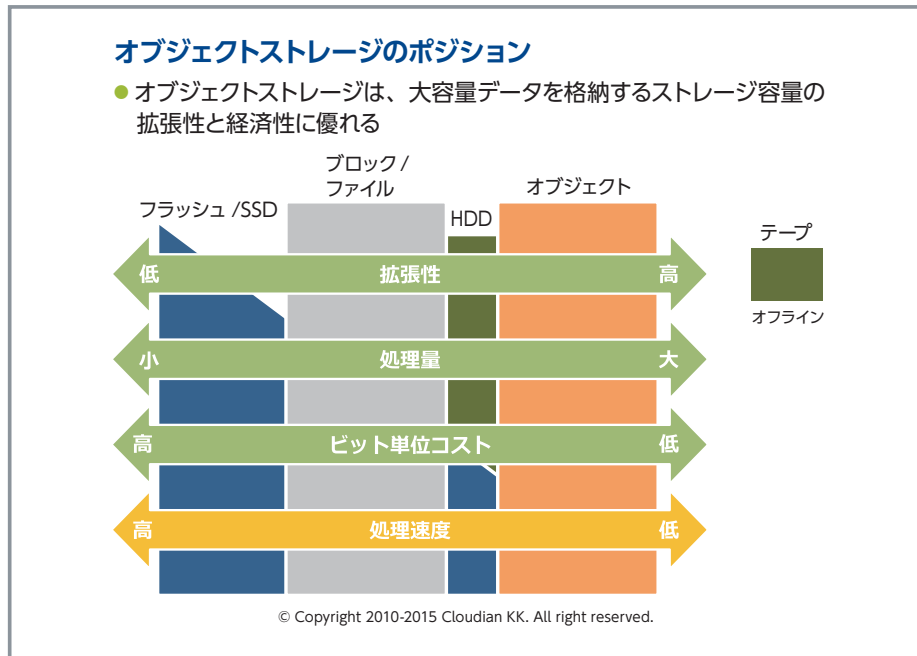
© Copyright 2010-2015 Cloudian KK. All rights reserved.

## 2 スケールアウト型オブジェクトストレージの概要

“スケールアウト型オブジェクトストレージは、大量データの保存に最適”

このような企業 IT が直面する非構造化データ急増に伴う課題の解決には、ブロックストレージやファイルストレージに比べ、「拡張性

が高く」、「処理量が大きく」、「経済的」なスケールアウト型オブジェクトストレージが適しています。なかでも汎用サーバを活用する SDS (Software Defined Storage:ソフトウェア定義のストレージ) が注目を集めています。



この SDS によるスケールアウト型オブジェクトストレージの特長と仕組みを説明します。

### (1) オブジェクトストレージは大量データの管理を柔軟に行える構造

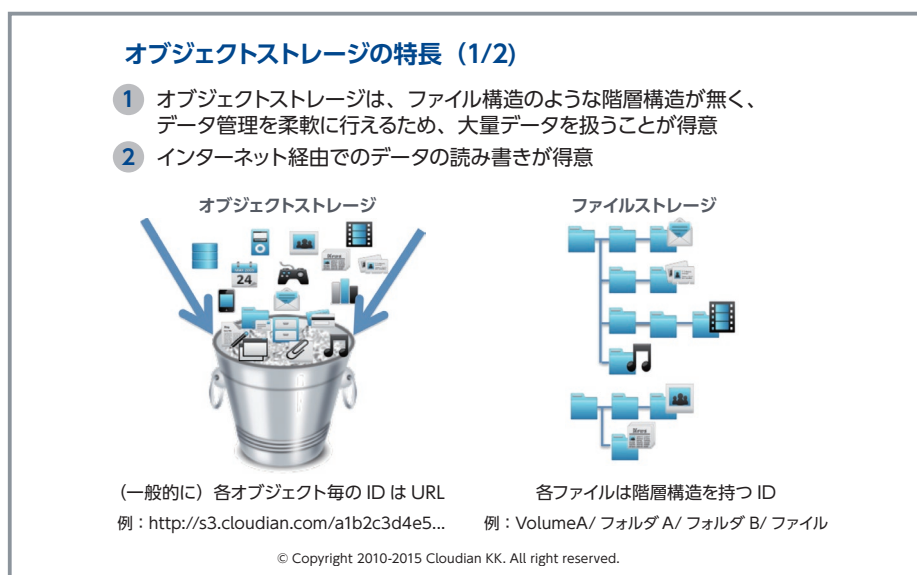
オブジェクトストレージは、ファイルストレージのような、ディレクトリやフォルダといった階層構造がなく、大きなバケツにデータを放り込むかのようにデータを保存します。

ファイル構造のように、(1) 階層構造を維持管理する仕組みが不要で、(2) ディレクトリやフォルダのサイズに制限がなく、(3) 目的の

ファイルに到達するために階層をたどる必要がなく、(4) ファイルの物理的な格納場所からデータ移動するとファイルの識別子 (ID) が変わってしまうといった制約がないため、大量データを柔軟に管理できます。

### (2) オブジェクトストレージに保存するデータの識別子 (ID) には URL が付与される

オブジェクトストレージに格納されるデータには、インターネットで利用される URL が識別子 (ID) として付与されるものが多く、インターネット経由でデータをやり取りすることに適しています。



### (3) オブジェクトストレージのデータには属性情報をカスタマイズして付与できる

オブジェクトストレージでは、各データに属性情報（メタデータ）をカスタマイズして付与できます。データと属性情報をセットにして扱えることから、オブジェクトストレージではデータやファイルのことを「オブジェクト」と呼びます。

のちほど紹介する CLOUDIAN HyperStore では、この属性情報を使いパブリッククラウドサービスの Amazon Web Service (AWS) が提供する Amazon S3 や Glacier に一定期間経過したファイルを転送し保存しておく自動階層化という機能を実現しています。

#### オブジェクトストレージの特長 (2/2)

- 3 各オブジェクトに属性情報（メタデータ）をカスタマイズして付与でき、データ管理が柔軟・効率的

一般的なファイル属性（ファイルサイズ、作成日、ファイル作成者等）だけではなく、データ管理に伴う属性をカスタマイズし付与可能

(例) CLOUDIAN HyperStore の Amazon S3 への自動階層化 (AWS AutoTiering)

- CLOUDIAN HyperStore に格納されたデータを一定期間後、パケット単位で Amazon S3 に自動的に転送し、プライベートクラウドとパブリッククラウド間で階層化（ハイブリッドクラウド）



© Copyright 2010-2015 Clouidian KK. All right reserved.

### (4) スケールアウトの仕組みがあれば、システム拡張に伴う運用負担を軽減できる

大量データに対応するオブジェクトストレージは、一般的にスケールアウトという仕組みが備わっています。スケールアウトとは、新しい装置を追加してゆくだけで、そのシステム全体の容量を拡張し

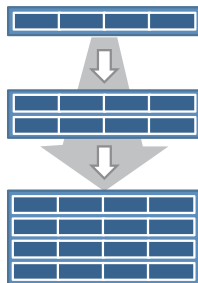
てゆくことです。一方のスケールアップとは、データが増え、装置のデータ容量が一杯になったり、性能が劣化してきた場合などに、設備を買換え、より大容量や高性能のものに置換してゆくことです。この場合、置換の時間に加え、その都度、データを移行するという運用作業も発生します。

#### スケールアウトでシステム拡張に伴う運用負担を軽減

- スケールアウトの仕組みにより、ストレージ容量拡張に伴う作業コストやデータ移行に伴うリスクが飛躍的に低減

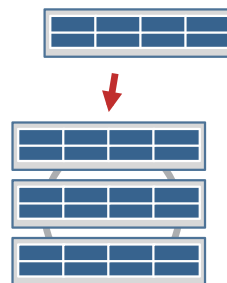
##### スケールアップ

- 設備買換・置換することで全体容量を拡張
- 都度、データ移行作業



##### スケールアウト

- ノードを追加してゆくことで全体容量を拡張
- システムがデータ移行

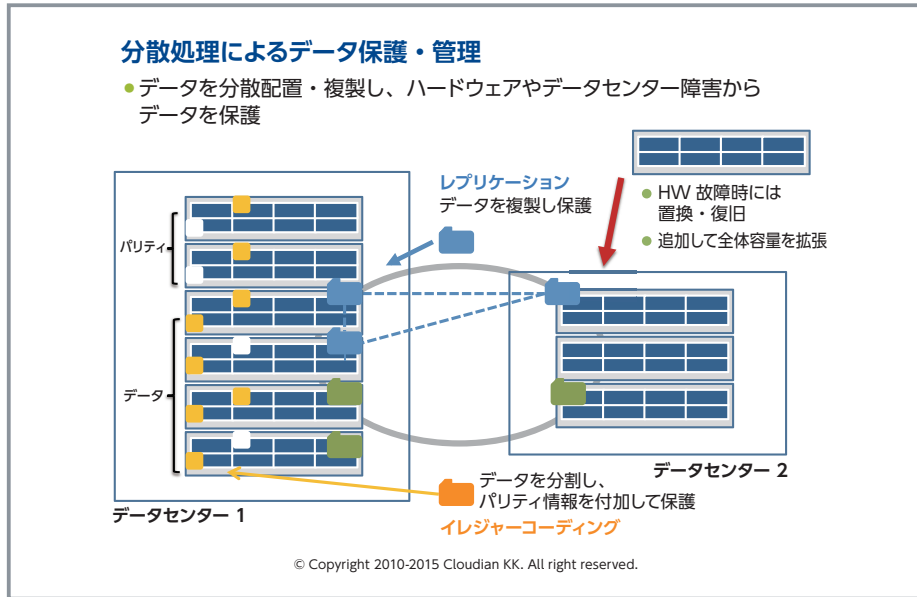


© Copyright 2010-2015 Clouidian KK. All right reserved.

## (5) スケールアウト型オブジェクトストレージは分散処理で柔軟にデータ保護

装置を追加してゆくスケールアウト型の場合、仮にひとつの装置が故障した場合でも、そこに保管されたデータが喪失したり、サービスが停止しないよう、「分散処理」が行われます。分散処理とは、分散する複数の装置を、あたかも「ひとつのシステム」のように統合制御することです。

この分散処理により行うレプリケーションという仕組みでは、データの複製を作り複数の装置に保存し、仮に1つのハードウェアに障害があったとしても、データが喪失しないように保護します。また、イレジャーコーディングという仕組みでは、ひとつのデータを複数に分割したうえで「消失訂正符号」というパリティ情報を追加し、分割したデータの一部が欠けたとしても復元できるようにすることで、データを保護します。

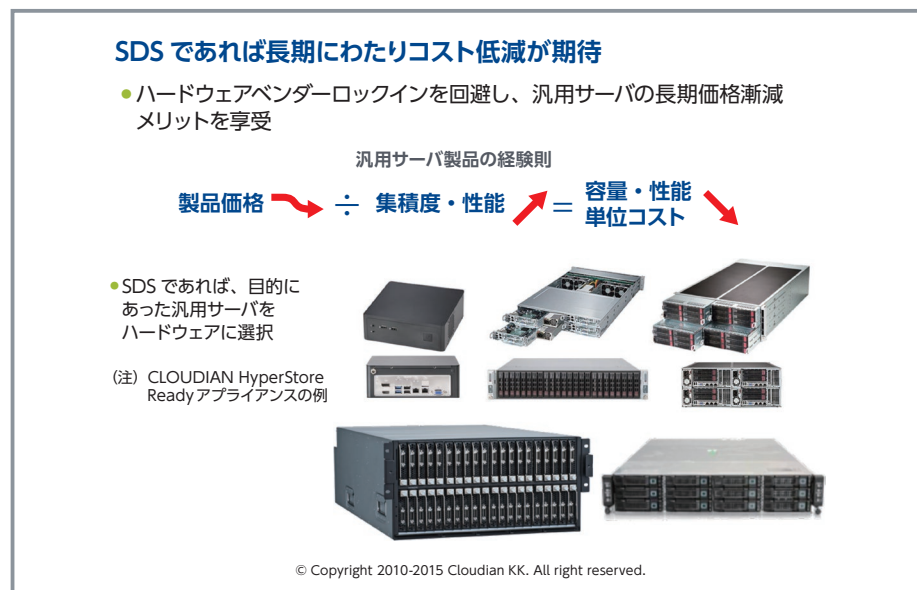


## (6) ソフトウェア定義のストレージ (SDS) であれば長期にわたりコスト低減

これらの仕組みをソフトウェア製品として提供するストレージは、最近では SDS (ソフトウェア定義のストレージ) と呼ばれ始めました。SDS では、従来のようなストレージ専用の物理的な装置ではなく、ソフトウェアが、データの保存領域に使う汎用的なサーバの内蔵

ディスクを、仮想的に「ひとつのストレージ空間」に統合制御します。

汎用サーバを使うメリットは、汎用サーバの容量や性能あたりの単位コストが下がり続けていることから、ストレージのコストを長期にわたり低減させることが期待できます。また、さまざまな種類の汎用サーバがあるため、利用目的にあったストレージに利用することもできます。



## CLLOUDIAN HyperStore の概要

### “数多くの独自機能を備えるスケールアウト型オブジェクトストレージの SDS、CLLOUDIAN HyperStore”

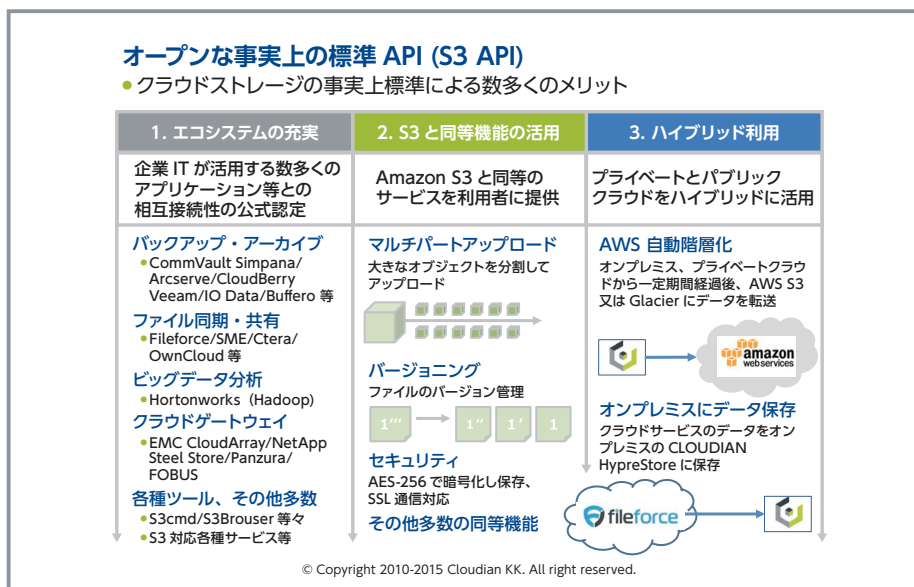
ここまで SDS のスケールアウト型オブジェクトストレージの特長と仕組みを説明してきましたが、その代表的な製品である CLLOUDIAN HyperStore は、これらの特長を高度に備えています。加えて、数多くの独自機能を備えており、そのいくつかを紹介します。

### (1) CLLOUDIAN HyperStore はオープンな事実上の標準 Amazon S3 API に準拠

CLLOUDIAN HyperStore の「もっとも分かりやすい」特長は、クラウドストレージの事実上の標準ともいえる Amazon S3 の API に高度に準拠する API をアプリケーションへのインターフェースとして提供している点です。これにより、Amazon S3 対応の数多くのアプリケーション等を「そのまま」活用できます。

また、CLLOUDIAN HyperStore を使うことで、Amazon S3 と同等のサービスや機能を提供できます。つまり、サービスプロバイダーだけではなく、企業自身が構築するプライベートなクラウドでも、社員に Amazon S3 と同等のサービスが提供できます。

さらに、CLLOUDIAN HyperStore で企業構内のストレージやプライベートクラウドを構築すれば、パブリッククラウドとプライベートクラウドをハイブリッドに活用することができます。CLLOUDIAN HyperStore の標準機能である自動階層化を使い、格納したデータを一定期間後、AWS Glacier に自動的に転送しておけば、地域冗長も兼ねた長期間のアーカイブ先となります。また、クラウド型のファイル共有サービスである Fileforce と合わせて活用すれば、データを不特定多数が利用するパブリッククラウドに預けることなしに、企業構内に設置した CLLOUDIAN HyperStore に保存しておくこともできます。



### (2) CLLOUDIAN HyperStore は複数データセンター間で柔軟にデータ保護

CLLOUDIAN HyperStore は、複数のデータセンター間で同期・非同期を選択しながら、データの複製を行い保護できます。

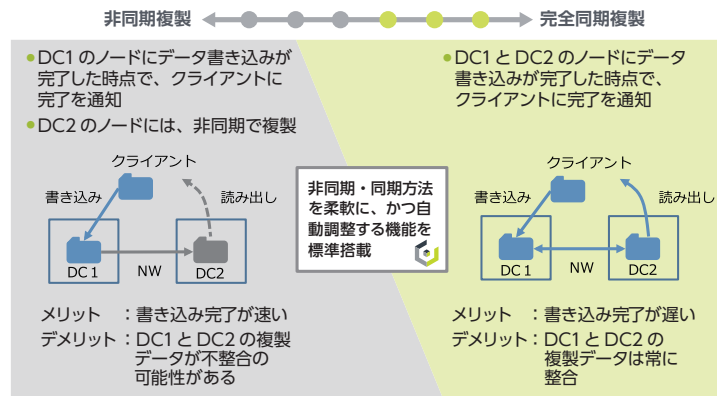
非同期で複製を行う場合、データセンター間で複製が完了するまで、複数のデータセンターのそれぞれのデータの整合性は異なります。一方、同期で複製を行えば、複数のデータセンターにあるデー

タは常に整合性を保つことができますが、ネットワーク経由で複製を同期させるため、書き込み処理完了は「遅く」なります。このようにデータの整合性と処理の速さはトレードオフの関係です。

CLLOUDIAN HyperStore であれば、利用目的に応じて同期と非同期複製を選ぶことができます。さらに同期か非同期かの 2 者択一ではなく、その水準を自動的に調整することもできます。

### 複数データセンター間の同期・非同期複製を選択

- 複数データセンター間の複製の同期・非同期を目的に応じて選択し、柔軟にデータの整合性を調整する機能が標準搭載



© Copyright 2010-2015 Cloudean KK. All right reserved.

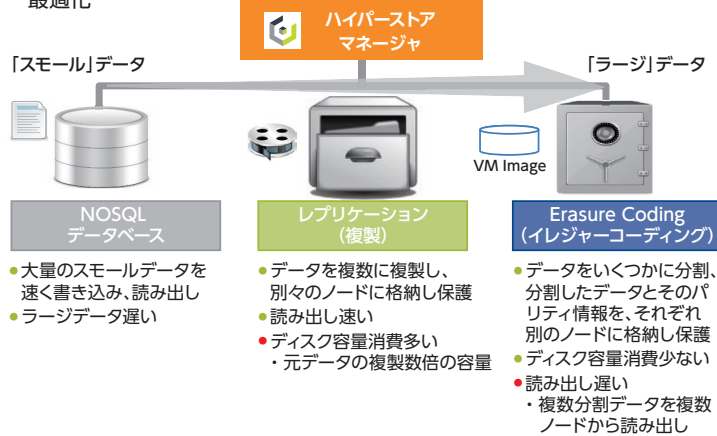
### (3) CLOUDIAN HyperStore はデータの種類と目的に応じて最適化

CLOUDIAN HyperStoreには、特許出願中のハイパーストアマネージャという機能が実装されています。これは、データのサイズが小さい場合には NOSQL データベースで迅速に処理し、データサイズが大きい場合には、レプリケーションかイレジャーコーディング

ングかを選択し、データを保存できます。この場合、レプリケーションを選べば、データを速く読み出すことができますが、ディスク容量を多く必要とします。一方、イレジャーコーディングを選べば、データの読み出しが遅くなりますが、ディスク容量を効率的に利用できます。

### データの種類と目的に応じた最適化

- ハイパーストア・マネージャ (特許出願中) により、小から大までデータサイズに適した処理性能や、データの利用目的に応じたデータ保護方法を選択し最適化

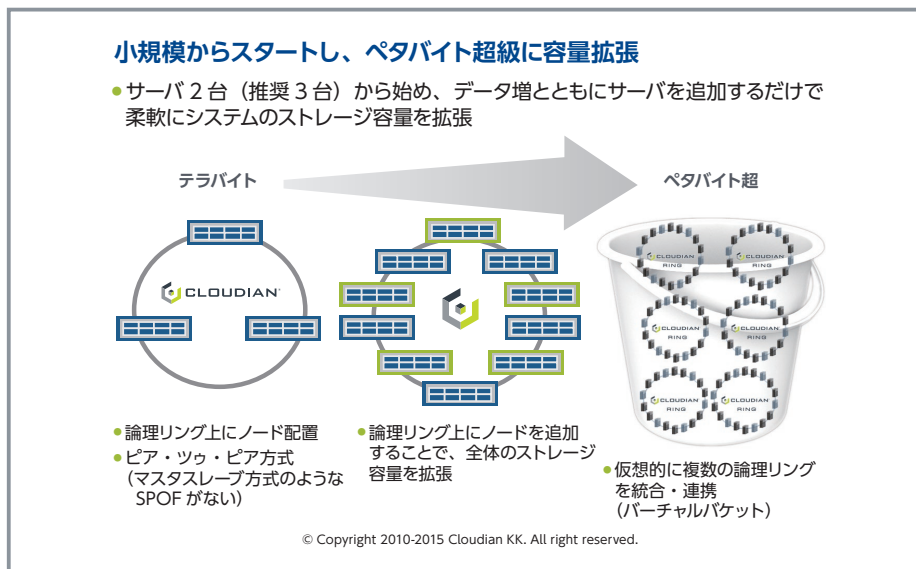


© Copyright 2010-2015 Cloudean KK. All right reserved.

#### (4) CLOUDIAN HyperStore は小規模からスタートし、ペタバイト超に拡張

CLOUDIAN HyperStore は最小 2 台（推奨 3 台）からスタート

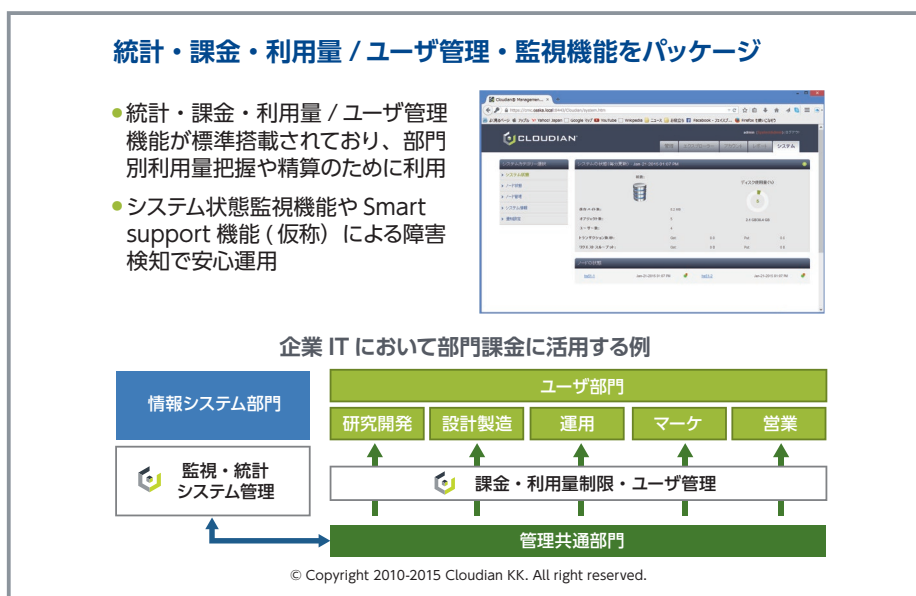
し、汎用サーバを追加するだけで、ストレージシステム全体の容量をペタバイト超に柔軟に拡張することができます。



#### (5) CLOUDIAN HyperStore はターンキーで利用開始

CLOUDIAN HyperStore には統計・課金・利用量・ユーザ管理・

監視機能も標準パッケージに含まれており、企業 IT で活用する際、管理共通部門が部門別課金等を行うことが可能です。





## 4 CLOUDIAN HyperStore の企業 IT における使いどころ

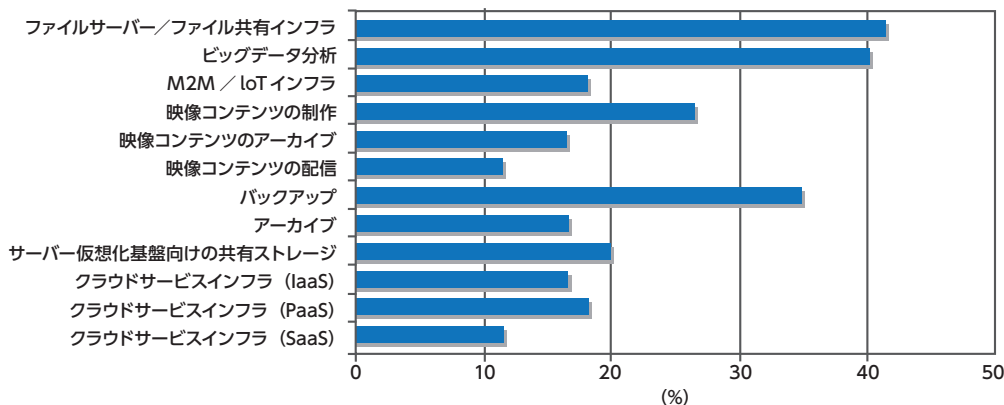
ここからは、企業 IT が抱える課題を解決するための CLOUDIAN

HyperStore の使いどころについて、説明していきます。

オブジェクトストレージに対するワークロード別利用意向は、回答率の高い順に「ファイルサーバー / ファイル共有インフラ」(41.7%)、「ビッグデータ分析」(40%)、「バックアップ」(35.0%)、「映像コンテンツの制作」(26.7%)、「サーバー仮想化基盤向け共有ストレージ」(20%)となった。ファイルデータの管理や

バックアップといった従来のワークロードに加え、ビッグデータ分析といった新たなワークロードでの利用意向も高い結果となった。このほか映像コンテンツ関連やサーバー仮想化環境向けインフラなど、多様なワークロードでの利用意向を持っていることが分かる。

スケールアウトオブジェクトストレージのワークロード別利用意向



n = 60

Notes:

「2014 年 国内ファイル / オブジェクトストレージ市場予測と競合分析：台頭する Software-Defined Storage の影響を探る (IDC# J15531001、2014 年 12 月発行)」を基に作成

回答対象は、スケールアウトオブジェクトストレージの「導入を検討中 / 導入予定がある」「導入済み」の回答者

10%以上の回答を得た項目を掲載

複数回答

Source: IDC Japan, February 2015

出典: IDC White Paper「クラウドファン、エンタープライズへの普及戦略を強化: オブジェクトストレージのソリューション展開を拡大」Sponsored by CLOUDIAN(2015 年 2 月)

### (1) CLOUDIAN HyperStore を容量制限を気にしない共通データストレージに活用

業務系や情報系のストレージは IT 部門で調達管理しているものの、研究開発部門等ライン部門のストレージの調達や管理はユーザー部門に任せているケースを冒頭で紹介しました。これらの部門におけるデータの大部分が、「非構造化データ」です。ひとつの例になりますが、ある部門独自で管理するモニターカメラの動画は 5 分間で 10MB でした。1 台のカメラ画像であれば 1 年間で 1TB、複数拠点に 100 台配置すれば、100TB です。このように、部門

毎のニーズから生まれた非構造化データを部門別でストレージ装置に保管しはじめるとサイロ状になります。

CLOUDIAN HyperStore であれば、スケールアウトの仕組みを備え、多数利用者が共有できるマルチテナントという機能もあります。そのため、まずは非構造化データが多い部門への導入から始め、他の部門が利用し始める都度、そのストレージ容量を柔軟に拡張してゆけます。その結果、やがては IT 部門が統合管理のうえで、全部門のファイルやデータを格納できる、容量制限を気にしない共通データストレージ基盤にすることができます。

## 容量制限を気にしない共通データストレージ

- パブリッククラウドストレージサービスと同等の、容量制限を気にしないデータストレージ環境を企業自身で構築・管理・運用



スケールアウト  
オブジェクトストレージ

(例) CLOUDIAN HyperStore に  
モニター動画を保存した例  
5分間毎に約10MB(カメラ1台)

© Copyright 2010-2015 Cloudian KK. All right reserved.

## (2) CLOUDIAN HyperStore をビッグデータ分析のための大量データ保管に活用

CLOUDIAN HyperStore で構築した共通ストレージ基盤にデータを保存しておけば、ビッグデータ分析のために大量に蓄えたデータの保管場所として活用できます。

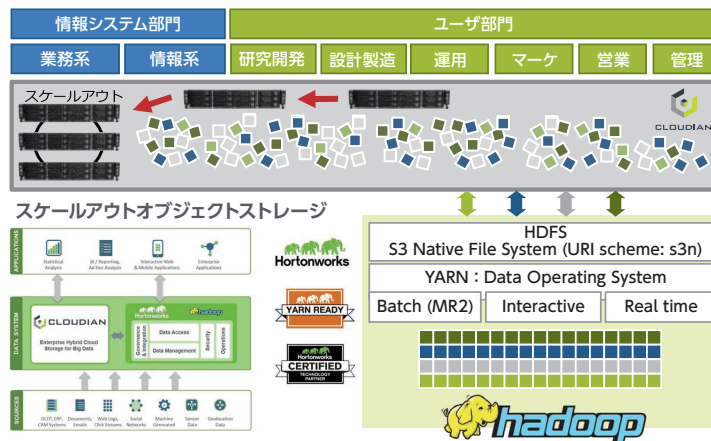
ビッグデータ分析のツールとしては Hadoop が多く活用されていますが、Hadoop 専用の HDFS (Hadoop File System) に貯めたデータを他の目的に利用する際には、別のストレージに移

し替えることが必要となり、余分なストレージを消費することになります。

Hadoop は S3 API のインターフェースも備えていることから、CLOUDIAN HyperStore に Hadoop 分析に利用する多種多様なデータを、とりあえず貯めておき、Hadoop 分析が必要な時にだけ、読み出して活用することができます。特に CLOUDIAN HyperStore は、Hortonworks から Hadoop との相互接続性の公式認定も受けています。

## ビッグデータ分析ストレージ

- Hadoop のファイルストレージ基盤として一元活用し、データ格納場所の重複を回避



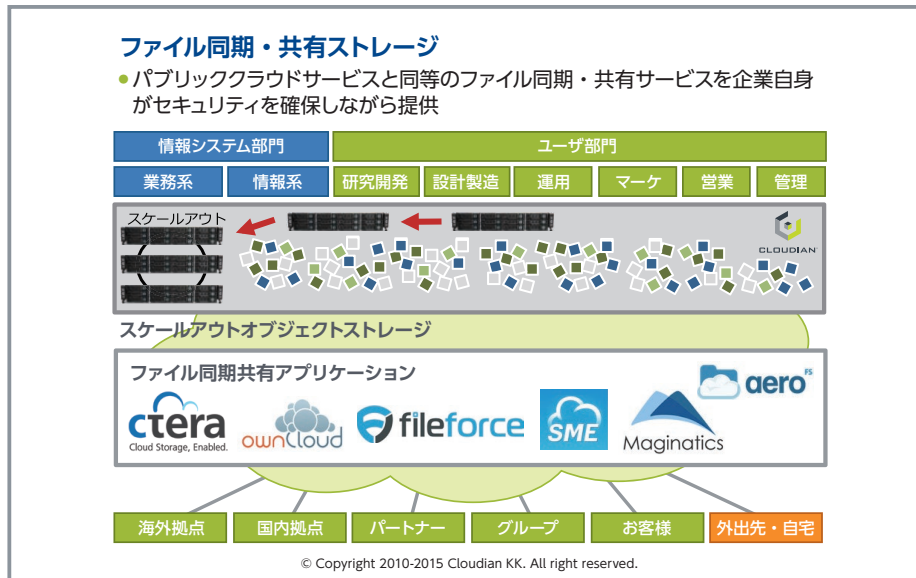
© Copyright 2010-2015 Cloudian KK. All right reserved.

### (3) CLOUDIAN HyperStore をファイル同期・共有ファイルのデータ保管に活用

CLOUDIAN HyperStore は、各種ファイル同期・共有アプリケーションとともに活用できます。

最近では、BYOC (Bring Your Own Cloud) とも言われるように、社員が個人契約しているクラウドストレージサービスにファイル等

を保存し、社内だけではなく、社外とファイルを共有しながら共同作業をするというケースも多くあります。ただし、コンプライアンスの観点から、外部にデータを預けることに制約がある企業が多いことも事実です。CLOUDIAN HyperStore とファイル同期・共有アプリケーションを合わせて使えば、いわゆる企業内 Dropbox といったファイル同期・共有サービスを、IT 部門のセキュリティ管理のもとで社員に提供することができます。

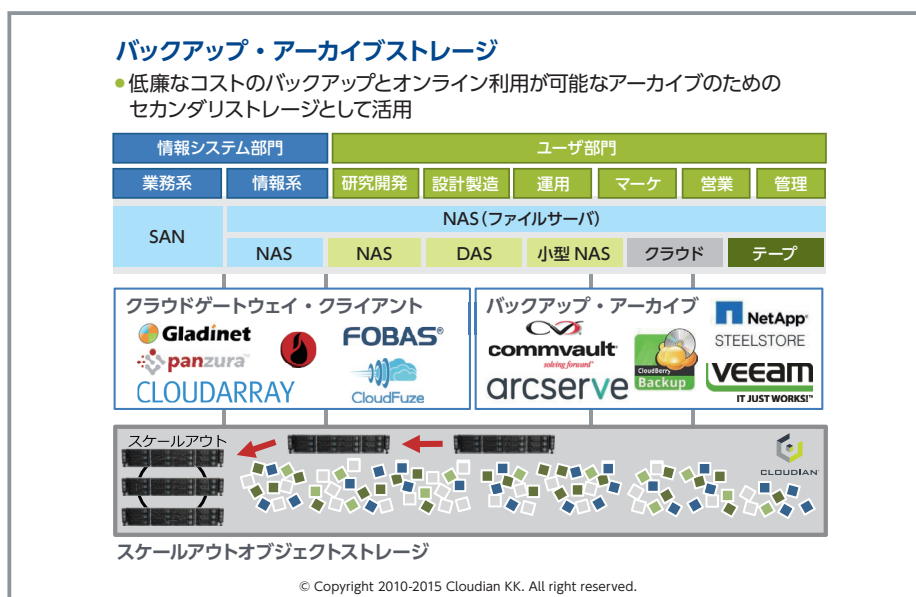


### (4) CLOUDIAN HyperStore を SAN・NAS の 2 次ストレージやバックアップに活用

サイロ状となった非構造化データは、各ユーザ部門の SAN、NAS、DAS といったストレージ装置に保存されています。CLOUDIAN HyperStore で共通ストレージ基盤を作るにしても、それらのストレージ装置を全て置換するとなると、データ移行の時間やコストもかかります。

「クラウドゲートウェイ」や「バックアップ・アーカイブ」のアプリケーションには、SAN や NAS のプロトコルを S3 API に変換する

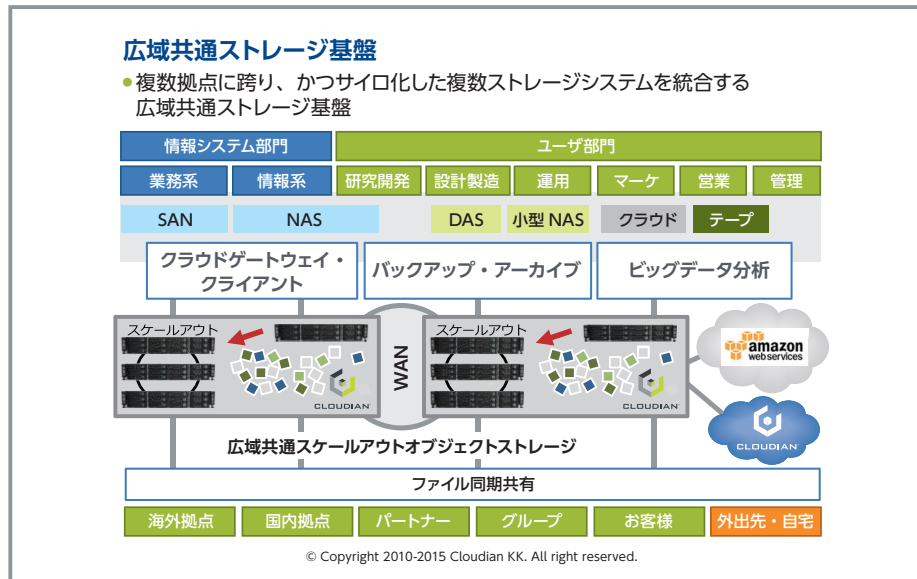
製品が多くあります。こういったアプリケーションを活用することで、現在のストレージを 1 次ストレージとして使い、データがあふれた場合やバックアップ・アーカイブのための 2 次ストレージとして、CLOUDIAN HyperStore を活用するといったように、ストレージを目的に応じて使い分けることが可能です。つまり、速い処理が必要な場合には、高性能な SAN や NAS を使い、アクセスする頻度の少ないデータは、段階的に CLOUDIAN HyperStore にデータを転送し階層化して保存するということです。これにより、高価で高機能なストレージを効率的に活用でき、その一方で、万が一のためのバックアップやアーカイブのコスト低減につながります。



## (5) CLOUDIAN HyperStore で複数データセンターにデータを分散配置し DR、BCP 対策

CLOUDIAN HyperStore を使えば、DR や BCP のために、この共通ストレージ基盤を広域の複数データセンターに分散配置することも可能です。さらに、AWS の Amazon S3 と Glacier や、

Cloudian HyperStore を採用するパブリッククラウドストレージサービスにバックアップやアーカイブすることもできます。つまり、各部門毎にサイロ化した非構造化データを、複数拠点やクラウドに分散配置して保護しながらも、企業 IT 部門が統合管理できる「広域共通ストレージ基盤」となるのです。



## 5 さいごに

2年間で倍増し続ける非構造化データに直面する企業 IT における課題は、その急増する量に柔軟に対応することです。これに対し、オンプレミスでスタンドアロン利用を前提とした従来のストレージ装置で対処し続けられれば、ユーザー部門別のサイロとならざるを得ず、企業として統合管理がますます難しくなります。

このような課題の解決には、本ホワイトペーパーで紹介した「次世代ストレージ」とも言われるスケールアウト型オブジェクトストレージが適しています。その代表的な製品である CLOUDIAN HyperStore であれば、企業 IT が必要とするあらゆるシーンで、さまざまなアプリケーションや最適なハードウェアと組み合わせ、幅広く活用できるのです。

### クラウドファンについて

日本と米国を開発拠点とするクラウドファンは、パブリッククラウド、プライベートクラウド、オンプレミス環境でハイブリッドに活用できる SDS (Software Defined Storage: ソフトウェア定義ストレージ) である [Cloudian HyperStore] をソフトウェア製品及びアプライアンス製品により提供しています。国内外大手プロバイダー、エンタープライズが採用する Cloudian HyperStore は、複数データセンター間を含み、データ複製・分散配置によるデータ保護をしながら、汎用サーバ 2 台からペタバイト超級にまで経済的に、柔軟にスケールアウトします。統計・課金・管理機能も実装済みであり短期間に利用開始できます。



クラウドファン株式会社

[www.cloudian.jp](http://www.cloudian.jp) | [info@cloudian.com](mailto:info@cloudian.com)