

Quantum®

ホワイトペーパー

# 「FOREVER」コンテンツ・アーカイブ

ベストプラクティスと検討課題

## 目次

クアンタムについて.....	3
はじめに .....	3
「Forever」コンテンツ・アーカイブの要件 .....	4
ワークフローの制作段階に関する検討課題 .....	4
数十年にわたりデータを保護するためのベストプラクティス .....	5
ファイルとオブジェクト・ストレージ形式の比較.....	7
アーカイブ・ストレージのストレージ技術 .....	7
データの分類とタグ付け、エラスティック検索.....	8
オンプレミスとパブリック・クラウドの検討課題.....	9
結論.....	11

## クアンタムについて

弊社は20年以上にわたり、世界最大のデジタル・コンテンツ・アーカイブの構築とサービス提供を行っています。世界中の大手スタジオ、放送局、スポーツ・リーグ、企業向けに、コンテンツのキャプチャ、処理、分析、保存を目的とした最先端テクノロジーを構築してきました。

詳細については、[www.quantum.com/jp/](http://www.quantum.com/jp/)をご覧ください。

## はじめに

メディア企業にとって最も価値のある資産は、デジタル・コンテンツ、つまりファイルです。これらのコンテンツから価値とインサイトを引き出すことは、今日のビジネスの推進力であり、明日の市場と経済の推進力でもあります。ファイルおよびオブジェクトとして保存されるこれらのコンテンツは、飛躍的に増加しており、2025年までに、この「非構造化」データは地球上のすべてのデータの80%を占めると推定されています。

コンテンツ・プロデューサーとマネージャーは、これらのデータを貴重な資産と同じように扱い始めることが不可欠です。

- これらのコンテンツは何十年にもわたってアーカイブして保護する必要がありますが、クラウドが正解とは限りません。
- データを失うわけにはいきません。価値が高いため、データを保護し、災害から守るために複数の場所に保存する必要があります。

高い成長率により、これらのデータ・セットはすぐにペタバイト、エクサバイト規模に増加する可能性があります。数十年にわたってこの規模でデータを保存および保護する場合は、一層の複雑さが加わります。

しかし、データの保存と保護はほんの始まりに過ぎません。インサイトの獲得、コンテンツ制作の迅速化、新しいチャンネルを使用したコンテンツの配信、新しい方法によるファンとのつながりのために、これらのデータを使用するユーザーとアプリケーションがデータにアクセスできるようにする必要があります。したがって、コンテンツは管理され、検索しやすく、分析しやすいものでなければなりません。

「スターウォーズ」や「蒸気船ミッキー」の原版をどのように保存しますか？ペープ・ルースの本塁打のフィルム映像をどのくらいの期間保存しますか？患者の記録、MRI画像、CATスキャンをどのくらいの期間保存しますか？裁判で使用された監視映像は？新薬の設計と製造に使用されたゲノム・シーケンサー画像は？

弊社のお客様の多くはこれを「Forever (永遠の) コンテンツ・アーカイブ」問題と呼んでおり、ほとんどのスタジオ、放送局、ポストプロダクション・ハウス、広告代理店は、今後数年間で計画を策定する必要があります。

このホワイトペーパーでは、以下の項目について検討していきます。

- 「Forever」コンテンツ・アーカイブの要件
- これらデータのライフサイクルのハイパフォーマンス段階に関する検討課題
- 数十年にわたりデータを保護するためのベストプラクティス
- ファイルとオブジェクト形式の比較
- アーカイブ・ストレージのストレージ技術
- データの分類とタグ付け
- エラスティック検索
- オンプレミスとクラウドの検討課題
- 今後の展望

## 「Forever」コンテンツ・アーカイブの要件

まず、「Forever」コンテンツ・アーカイブに必要な基本的要件と属性から始めましょう。

属性	説明
ペタバイトおよびエクサバイト規模への拡張が可能	高解像度カメラは、4K、8K、16Kの解像度で高フレーム・レートのコンテンツを生成します。未加工のコンテンツがすべて使用されるわけではありませんが、その多くは保存され、完成品は数十年にわたって保存されます。このようなデータ・レートでは多くのコンテンツ・アーカイブがPB規模で急増する可能性があり、弊社の一部のお客様の場合は数百から数千PBにもものぼります。
最速の「ホット」ストレージから「コールド」ストレージにまで及ぶストレージ・サービス	すべてのコンテンツ・アーカイブは、迅速に取り込んで処理する必要のあるデータから始まります。この段階では速度が重要で、メディア・ワークフローを加速するためにNVMeフラッシュ・ドライブなどの技術が使用されています。一方、大規模クラウド・プロバイダーは、長期のコールド・ストレージ用にテープ技術の導入を進めています。もちろん、ハード・ドライブは今後もその役割を担っていきます。したがって、コンテンツ・アーカイブでは、ストレージ技術を組み合わせて使用する必要があります。
データの保護が必要：失うわけにはいきません	保存したデータは保護しなければなりません。この規模でデータを保護するための検討課題とベストプラクティスについて説明します。
データは数年または数十年にわたり簡単に検索およびアクセスできることが必要	かつて誰かが言いました「アーカイブ内のデータは価値を生むまで価値はない」。数年前あるいは数十年前に作成されて保存されたデータを取り出す必要が生じることがありますが、それがいつ発生するか常に予測できるとは限りません。

これらの基本的要件を念頭に置いて、「Forever」コンテンツ・アーカイブ戦略を設計する際の検討課題とベストプラクティスのいくつかを見てみましょう。

## ワークフローの制作段階に関する検討課題

ビデオ・コンテンツが作成されると、迅速にアップロードして取り込む必要があります。また、ファイルが大きいため、非常に高速のストレージ・システムが必要です。その後、スタッフとアプリケーションがしばらくの間データを処理しますが、通常、アクティブに処理されている間、データはファイル形式です。

現在、これらの「高速」ストレージ・プールはNVMeフラッシュ・ストレージ上に構築されており、ファイル、ブロック、またはオブジェクト・インターフェイスを使用してそのデータにアクセスするオプションがあります。重要な要件は、超高速のストリーミング性能に加え、メディア制作とポストプロダクションのワークステーション・エコシステム、編集およびVFXのアプリケーション、およびさまざまな資産管理プラットフォームとの統合です。

従来のスケールアウトNAS設計とTCP/IPなどのネットワーク・プロトコルは、作成されるデータの量と、データの取り込みと共有に必要な速度の両方に直面して機能しなくなり始めています。NVMeフラッシュ上で実行されるソフトウェア定義のファイル・システム・クラスターが好まれるようになってきており、レンダリングとトランスコーディングのためにアプリケーションや他のサービスと統合する特別なクライアントもあります。

## 数十年にわたりデータを保護するためのベストプラクティス

データベースまたは仮想環境のバックアップは、プランとしてはかなりシンプルです。バックアップ・アプリケーションを使用してネットワーク経由でサーバーをバックアップするか、またはスナップショットを使用して継続的にデータを保護します。このデータは、コンプライアンス目的で7年間保存され、その後期限切れになります。

何十年も保存する必要がある超大規模コンテンツ・アーカイブを保護する場合、プランは大幅に異なります。

課題	検討課題
「大きすぎてバックアップできない」データ	バッチ・バックアップ方式を使用してこのデータをバックアップしようとする、ネットワークに過度の負担がかかり、多くの場合、バックアップ・ウィンドウの大きさが十分ではありません。
法外な費用がかかるデータ複製	多くのユーザーは、このデータをスケールアウトNASクラスター間で複製することによって保護しようとしています。製品コストとサービス・コストの両面で、これは非常に高コストになります。
データをフラッシュまたはディスクに数十年間保存する場合のコスト急増	フラッシュとディスクは高性能でニアライン・ストレージに最適ですが、どちらのメディアも数十年にわたってデータを保存するには高コストで信頼性は低くなります。
災害からの保護	災害から保護するには、データを2か所または3か所にコピーすることが強く推奨されます。
ランサムウェアおよびサイバー攻撃からの保護	ランサムウェア、マルウェア、その他のサイバー攻撃の脅威が高まっています。100年アーカイブのデータをランサムウェアの脅威から保護する必要があります。

### 推奨される保護方法としてのイレイジャー・エンコーディング

イレイジャー・エンコーディングまたはイレイジャー・コーディングは、超大規模メディア・アーカイブを保護するためのベストな方法として登場しました。イレイジャー・エンコーディングを使用すると、オブジェクトはチャンクに分割されて複数のノードに分散され、場合によっては複数のサイトに分散されます。イレイジャー・コーディングのアルゴリズムは、保護と安全およびストレージの効率向上を目的とした調整が可能です。

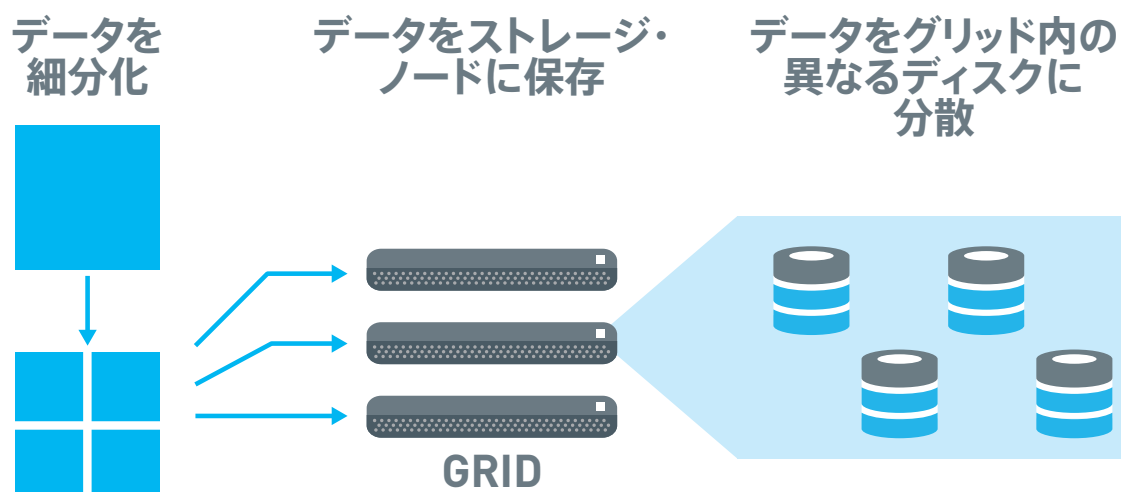


図1 - イレイジャー・コーディング図解

保存されてイレイジャー・コーディングされたデータは保護されます。バックアップは不要です。

## 地理的分散型イレイジャー・コーディングは3つのコピーでTOCが最も効率的

「Forever」コンテンツ・アーカイブにデータを保存する場合のベストプラクティスは、データのコピー×3を3か所に保存することです。これにより、ハードウェアとソフトウェアの問題、および災害から保護されます。地理的分散型イレイジャー・コーディングとは、地理的に分散した複数のサイトにイレイジャー・コードを分散できるソフトウェアのことです。

地理的分散型イレイジャー・コーディングの効率を表す例を以下に示します。

表1 - Row容量に対する割合としての使用可能容量の比較

保護方法	データ・コピー×1	データ・コピー×2	データ・コピー×3
RAIDとレプリケーションによるブロック・ストレージ	84% RAIDを使用	42% RAIDとレプリケーションを使用	28% RAIDとレプリケーションを使用
地理的分散型イレイジャー・コーディングされたオブジェクト・ストレージ*	80% イレイジャー・コーディングを使用	40% イレイジャー・コーディングを使用	61% 3-Geoイレイジャー・コーディングを使用

\* BitSpread搭載Quantum ActiveScale™ソフトウェアに基づく

## ランサムウェア対策とオフライン・ストレージ・コピーの必要性

ランサムウェア攻撃は一般化・巧妙化しています。マルウェアはサーバー上で休眠状態になっていることがあるため、破損したデータを複製またはバックアップしても効果はありません。データは破損したままです。

このため多くの企業は、ランサムウェアから保護するための実証済みかつ効果的な方法であるデジタル・テープを採用しています。テープは、デジタル・テープに保存されているデータがオフラインである、つまりネットワークからエアギャップされているという点で独特です。テープに保存されているデータは完全にセキュアです。

ライブラリに保存されているテープは「オフライン」  
つまりネットワークから「隔離」されています

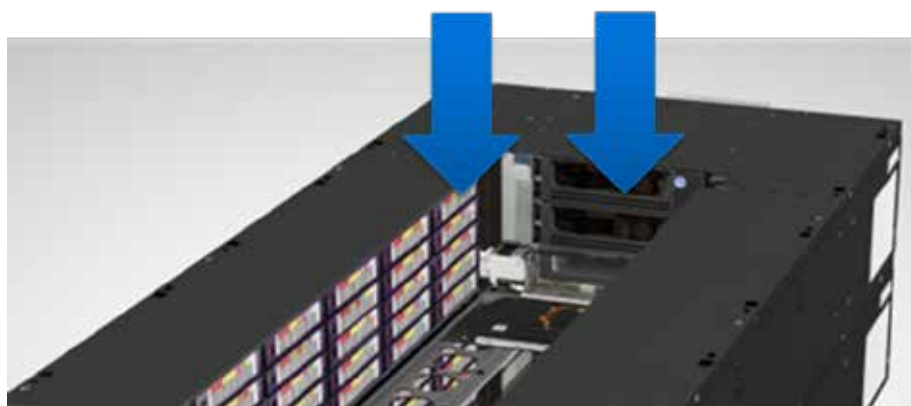


図2 - デジタル・テープ・アーカイブの「オフライン」または「エアギャップ」属性図解

## ファイルとオブジェクト・ストレージ形式の比較

コンテンツ・アーカイブに関する現在の検討課題の1つは、データ、ファイル、またはオブジェクトをどの形式に保存するかです。境界線は曖昧になっており、多くのベンダーが、ファイルまたはオブジェクト・インターフェイスをサポートする超高速ストレージと、ファイルまたはオブジェクト・インターフェイスのいずれかを備えた「超安全」なストレージの両方を提供しています。



### ファイル・システム

C:\folder\music.mp3

- データはファイルとして保存
- 階層構造
- ファイル・インターフェイスを介したアドレス指定



### オブジェクト・ストレージ

GET /object/Kbg18n7qepo  
PUT /object/Kbg18n7qepo

- データはオブジェクトとして保存
- フラットな名前空間
- S3インターフェイスを介したアドレス指定

図3 - ファイルとオブジェクト・ストレージの相違点図解

ただし、オブジェクト形式によるデータの長期保存には、いくつかの明確な利点があります。

- **大規模拡張が可能**: オブジェクト・ストアは、数十億オブジェクト、エクサバイト規模への拡張が容易です。
- **大規模環境での検索とインデックス作成が容易**: フラットな名前空間と、メタデータがオブジェクトの一部として含まれていることにより、オブジェクト・ストアとオブジェクト形式は大規模環境でも検索が容易です。
- **「クラウド・フレンドリー」**で、ユーザーはアーカイブに保存されているデータに対してクラウド・サービスを実行することができます。

## アーカイブ・ストレージのストレージ技術

下記の表は、現在アーカイブ・ストレージに利用できる最も一般的な技術を示しています。

技術	長所	短所
フラッシュ	<ul style="list-style-type: none"><li>• 低電力</li><li>• いくつかの「コールド」オプションあり</li></ul>	<ul style="list-style-type: none"><li>• 高価</li></ul>
ハード・ドライブ	<ul style="list-style-type: none"><li>• フラッシュよりも低コスト</li><li>• 保護のためのイレイジャー・コーディング</li></ul>	<ul style="list-style-type: none"><li>• 技術ロードマップが限界間近</li><li>• 電力が必要</li></ul>
テープ	<ul style="list-style-type: none"><li>• 最も低コスト</li><li>• 電力はほぼ不要</li><li>• データ寿命は30年以上</li><li>• ディスクよりも信頼性が高い</li><li>• 最も「環境に優しい」選択肢</li><li>• エアギャップされたストレージ・コピー</li></ul>	<ul style="list-style-type: none"><li>• 管理が難しい</li><li>• データへのアクセスと取得が難しい</li></ul>
光学式	<ul style="list-style-type: none"><li>• 低コスト</li></ul>	<ul style="list-style-type: none"><li>• 導入実績が少ない</li></ul>
クラウド(技術的オプションは上記と同じ)	<ul style="list-style-type: none"><li>• 「書き込んで忘れる」が容易</li></ul>	<ul style="list-style-type: none"><li>• データ・アクセス料金</li><li>• データ・セキュリティ</li></ul>
合成DNA	<ul style="list-style-type: none"><li>• 低コスト、低電力、超高密度</li></ul>	<ul style="list-style-type: none"><li>• 技術的にも商業的にもまだ実現可能ではない</li></ul>

上記に示すように、テープは「Forever」コンテンツ・アーカイブでの使用を検討すべき重要な技術です。合成DNAストレージが商業的および技術的に実現可能になるまで、テープには他のストレージ技術に比べて多くの明確な利点があります。

テープの利点:

- 最も低コスト
- 電力はほぼ不要
- データ寿命は30年以上
- ディスクよりも信頼性が高い
- 最も「環境に優しい」選択肢
- ランサムウェアから保護するためのエアギャップされたストレージ・コピー

さらに、テープ技術(磁気メディア)には、面密度の改善が続くという実現可能な技術ロードマップがあります。つまり、ハード・ドライブ技術とは異なり、より少ないスペースとより少ないコストでより多くのデジタル・データを保存できるようになるということです。

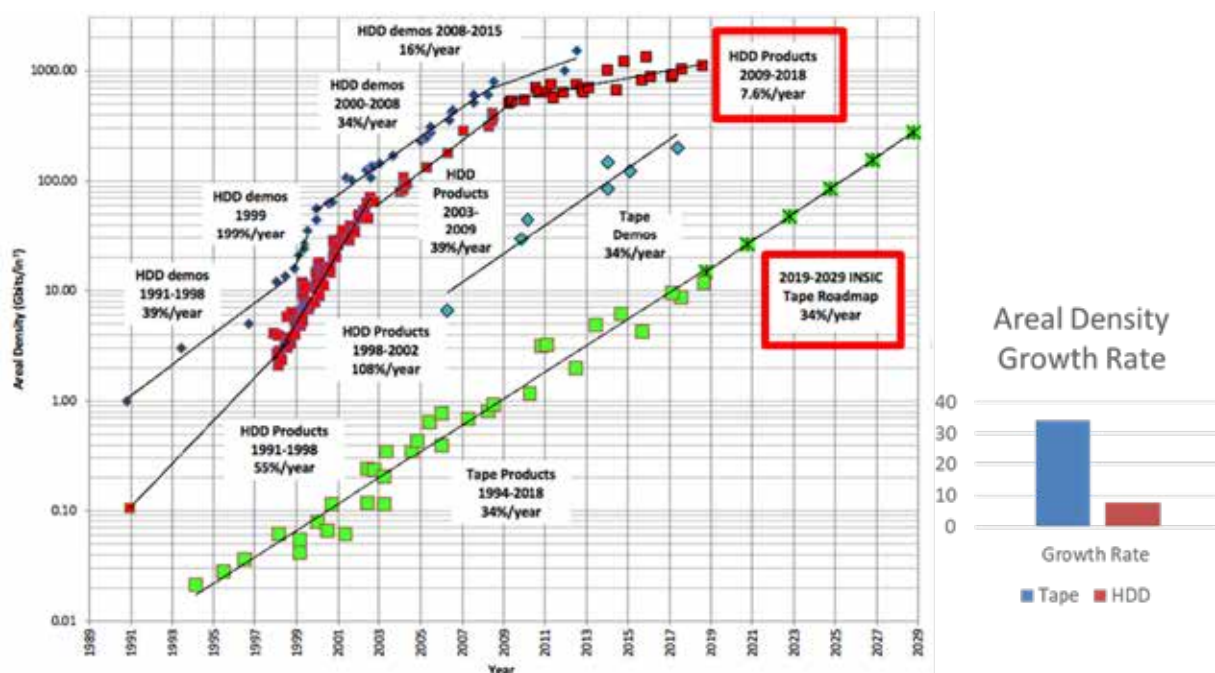


図4 - デジタル・テープとハード・ドライブの面密度ロードマップの比較

## データの分類とタグ付け、エラスティック検索

ここまで、さまざまなストレージ技術とコンテンツ・アーカイブの構築方法に関する主な検討課題について説明してきました。しかし、最大の課題はアーカイブのアクセス方法と検索方法であり、これを実現するための最初のステップは、**取り込み時におけるゼロタッチのリアルタイム・データ分類**です。

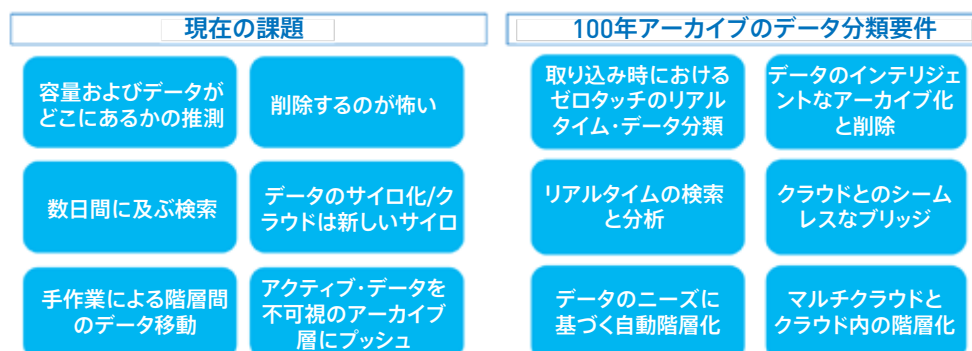


図5 - 一般的な課題とデータ分類の要件



下図に示すように、適切に設計されたデータ分類エンジンには、強力なリアルタイム分析、アーカイブされたコンテンツへの透過的なアクセス、およびスケーラブルなエラスティック検索エンジンが含まれている必要があります。

これによりユーザーは、オンプレミスとクラウドのインフラストラクチャ間で必要な時に必要な場所にデータを配置できるインテリジェントなアーカイブを構築でき、最終的にはデータをインサイトと成果にすばやく変えることのできるアーカイブを構築できます。

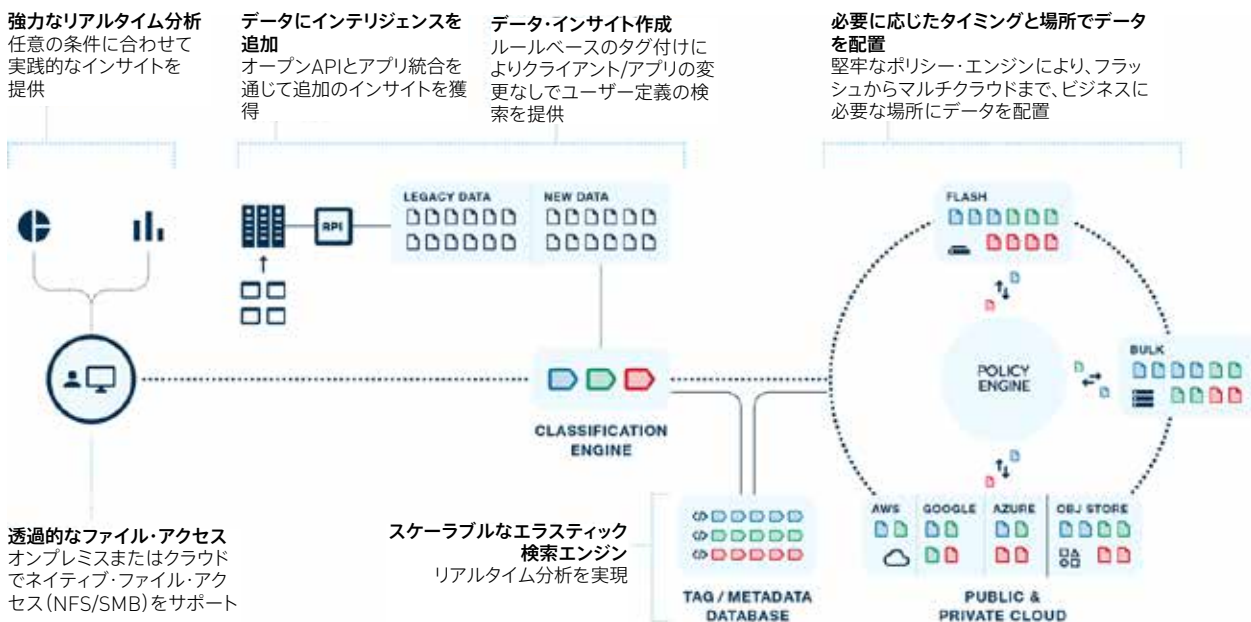


図6 - リアルタイム・データ分類エンジンの図解

## オンプレミスとパブリック・クラウドの検討課題

多くの組織が、バックアップ、アーカイブ、およびコールド・ストレージにパブリック・クラウドを検討したり使用したりしています。また、パブリック・クラウド・プロバイダーは、使いやすさの新しい基準を設定するなど、多くのことを非常にうまく行っており、最終的にはユーザーを基盤技術から分離させました。

将来はハイブリッドクラウドとマルチクラウドの両方になり、以下に概説する検討課題が重要になると考えられます。

検討課題	説明
データのアクセスおよび使用料金	ほとんどのクラウド・プロバイダーは、大量のデータを保存するための超低コストのストレージ・サービスを提供していますが、これらのデータを取得またはアクセスすると極めて短期間で高コストになってしまいます。 また、アーカイブされたデータは「価値を生むまで価値はない」ため、このようなアクセスと取得の料金を予測することは非常に難しく、予算を立てることも困難です。
データ・サイロのリスク	クラウド・プロバイダーがデータ・サイロのもう一つの形態であること気づき始めた企業が増えています。プロバイダーの切り替えは困難でコストがかかり、データ管理者はデータをクラウドに保存するときに一定レベルの制御を放棄してしまいます。
複数のクラウド・プロバイダーのサービスを利用したい	パブリック・クラウド・プロバイダーのマーケットプレイスには、常に新しいサービスが追加されています。このように競争の激しい状況は今後数年間急速に発展し続け、組織はさまざまなクラウド・プロバイダーから最高のサービスを活用したいと考えるはずで、これにより、特定のクラウド・プロバイダーにサイロ化またはロックされないように検討する必要があります。
データのセキュリティと制御	パブリック・クラウド・プロバイダーでさえ、ランサムウェア攻撃に耐性があるとは限らず、データ保護のベストプラクティスに従っている場合とそうでない場合があります。

## オンプレミスとマルチクラウド間のシームレスなブリッジ

パブリック・クラウド・プロバイダーは、このホワイトペーパーで概説したものと同一ストレージの基本技術を使用していますが、ユーザーを基盤技術から分離する優れたユーザーエクスペリエンスを開発しました。さらに、コンテンツ・アーカイブ管理者が活用したいと思う優れたサービスに満ちた大規模なマーケットプレイスを開発しました。

これが、大企業、コンテンツ・プロデューサー、およびコンテンツ・アーカイブを管理するすべての人にとって最適なアーキテクチャは、各自が構築したアーキテクチャであると弊社が考える理由です。経済性に優れ、クラウド・プロバイダー間でデータを移動する自由と柔軟性も得られます。

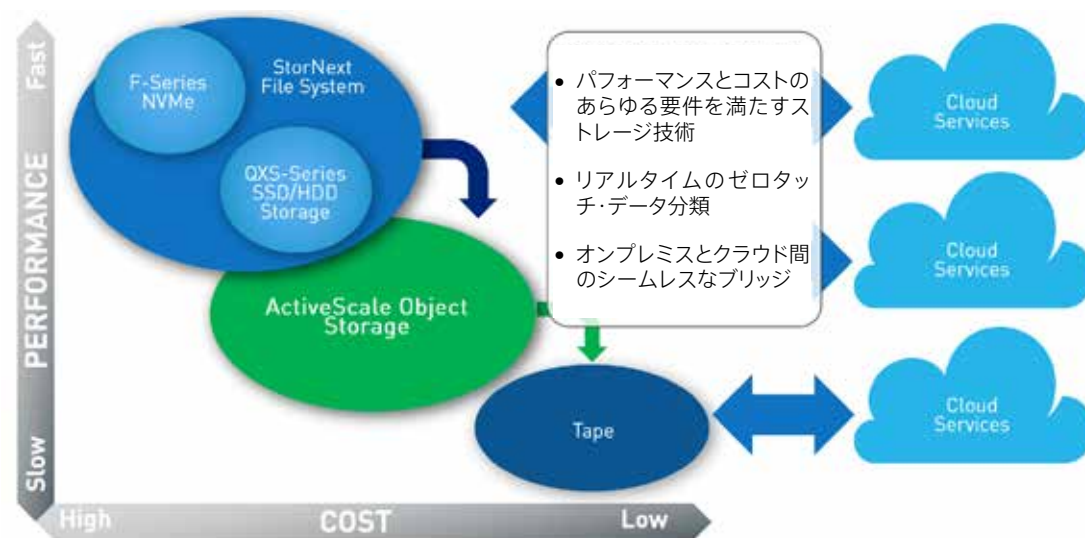
適切なデータ分類エンジンを使用し、オンプレミスとさまざまなクラウド・プロバイダー間でシームレスにデータを移動できれば、ユーザーは、最重要デジタル資産の制御、セキュリティ、データ保護を維持しながら、最高の経済性で最高のサービスを利用することができます。

## 結論

「Forever」コンテンツ・アーカイブの問題は、多くのCIO、データ・サイエンティスト、コンテンツ・クリエイター、セキュリティ・インテグレーターが今後5～10年で取り組む必要のある主要な問題であると考えられます。

弊社では、20年にわたり、お客様がこの問題を解決するのを支援してきました。お気に入りの映画やテレビ番組は、クアンタムにデジタル・アーカイブされています。過去100年分のお気に入りのスポーツ・シーンもクアンタムにデジタル・アーカイブされているかもしれません。世界中の政府機関が、国防のために、地球と気候変動の影響の研究を支援するために、そして宇宙を探索するために、クアンタムにデジタル・アーカイブを構築しています。

弊社は、世界最速のStorNext® NVMeストレージ・システムから、ActiveScaleオブジェクト・ストレージ・ソフトウェア、Quantumテープの最も低コストかつ最もセキュアな「コールド」ストレージまで、この問題を解決するために設計された技術のポートフォリオを構築しています。これらの技術とリアルタイムのデータ分類およびタグ付けならびにこれらすべての階層にデータを配置する機能を融合させ、ハイブリッドクラウドとマルチクラウドの「Forever」コンテンツ・アーカイブを構築することができます。



詳細は、[www.quantum.com/jp/](http://www.quantum.com/jp/)をご覧ください。

# Quantum®

## クアンタムについて

クアンタムの技術とサービスは、お客様のデジタル・コンテンツのキャプチャ、作成、共有を支援し、低コストで長期間保存・保護します。高性能の取り込みからリアルタイムのコラボレーションと分析、低コストのアーカイブまで、データ・ライフサイクルのあらゆる段階に適応するソリューションを搭載したクアンタムのプラットフォームは、高解像度ビデオ、画像、産業用IoTに、最速のパフォーマンスを提供します。毎日、世界中の大手エンターテインメント企業、プロスポーツ・チーム、研究機関、政府機関、大企業、クラウド・プロバイダーが、クアンタムで世界を豊かに、安全に、スマートにしています。その方法については [www.quantum.com/jp](http://www.quantum.com/jp) をご覧ください。

[www.quantum.com/jp](http://www.quantum.com/jp) • 03-6890-3038