

コンピュータ用テープによる デジタルデータの長期保存

一般社団法人 電子情報技術産業協会
情報・産業社会システム部会
技術企画・標準委員会
テープストレージ専門委員会

アジェンダ

アジェンダ

- デジタルデータの氾濫 - ビッグデータ時代の到来
- コンピュータ用テープメディアの活用
- デジタルデータの長期保存という課題
- 長期保存＝マイグレーション＋自動化
- 放送映像業界での活用例
- 最新IT技術のあらゆる分野への活用

デジタルデータの氾濫

- ビッグデータ時代の到来 -

デジタルデータの氾濫

- ビッグデータ時代の到来 -

Big Data とは、生成されたコンテンツではなく、消費されたコンテンツでもない。
Big Data とは、データを取り巻いているすべてのデータに関わる分析である。

科学的な実験からソーシャルネットワーク、
スマートテクノロジーまでをカバー



認識されていなかった課題を明らかにし、そのソリューションを見出す

Source: IDC's Digital Universe Study, sponsored by EMC, June 2011

デジタルデータの氾濫

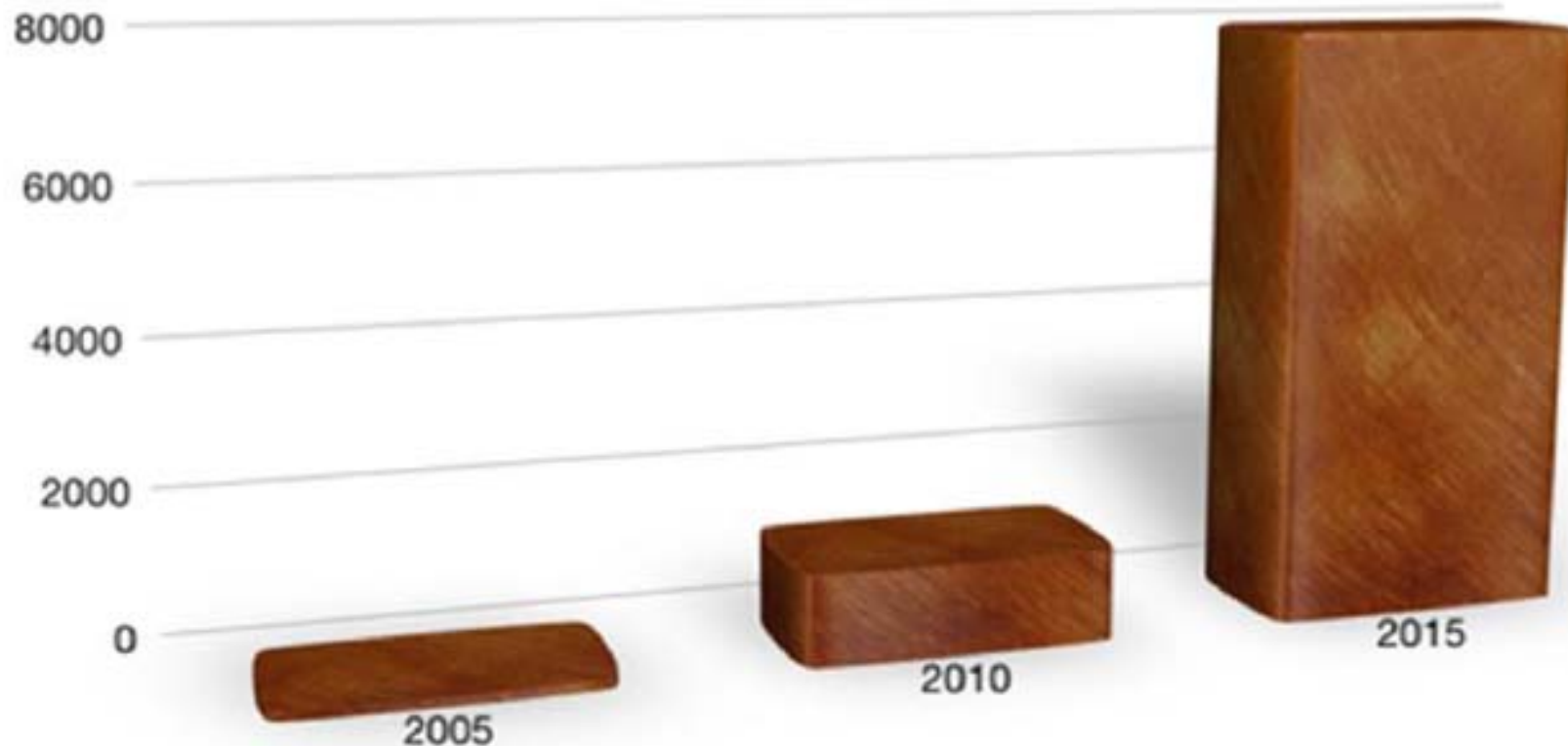
- ビッグデータ時代の到来 -

- ビッグデータとは？
 - 所謂あらゆる種類のデジタルデータ
 - SNS, YouTubeなどの出現で加速
 - 人類文明の夜明けから2003年までに生成されたデータ量と同じ量のデータが2日毎に生成されている
- データの重要度
 - 期間限定と半永久保存の二極化
 - 半永久保存には階層管理、ポリシー管理というコンセプト

デジタルデータの氾濫

- ビッグデータ時代の到来 -

10年間のデジタルユニバースの膨張：容量（エクサバイト）



Source: IDC's Digital Universe Study, sponsored by EMC, June 2011

デジタルデータの氾濫

- ビッグデータ時代の到来 -

デジタルユニバースとパブリッククラウド、2015 年



Source: IDC Digital Universe Study, sponsored by EMC, June 2011

デジタルデータの氾濫

- ビッグデータ時代の到来 -

Big Data とは、生成されたコンテンツではなく、消費されたコンテンツでもない。
Big Data とは、データを取り巻いているすべてのデータに関わる分析である。

科学的な実験からソーシャルネットワーク、
スマートテクノロジーまでをカバー



認識されていなかった課題を明らかにし、そのソリューションを見出す

Source: IDC's Digital Universe Study, sponsored by EMC, June 2011

デジタルデータの氾濫

- IT業界のビッグデータへの対応 -

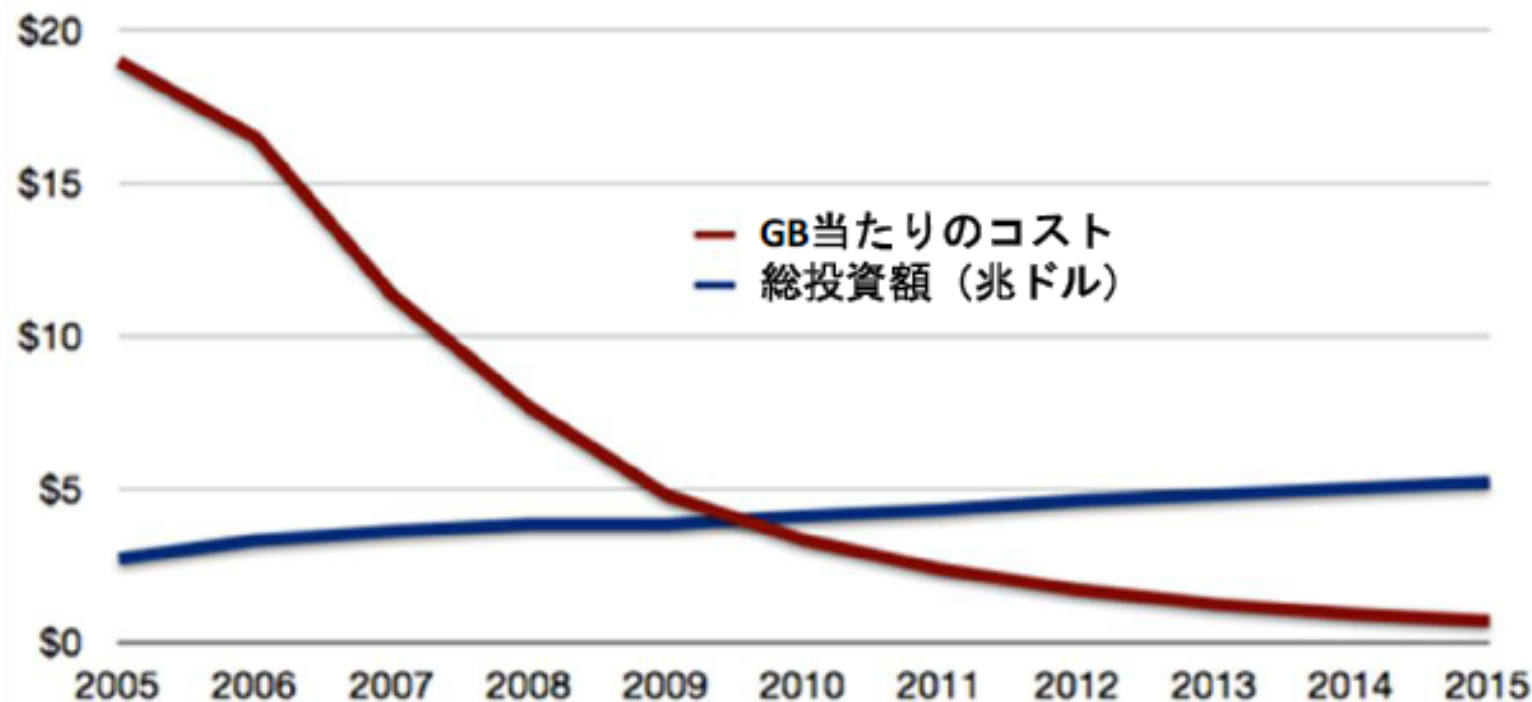
- 安全に半永久に保存できる仕組みづくり
 - 期間限定と半永久保存の二極化
 - 半永久保存には階層管理、ポリシー管理というコンセプトの導入
- 再利用できるデジタルアーカイブ
 - 検索性、アクセス性
- 半永久、実はコストとの戦い

デジタルデータの氾濫

- IT業界のビッグデータへの対応 -

- 半永久、実はコストとの戦い

デジタルユニバースの膨張に関するパラドックス（逆説）：コストの下落と投資額の増加



Source: IDC's Digital Universe Study, sponsored by EMC, June 2011

コンピュータ用テープメディアの活用

コンピュータ用テープメディアの活用

- 信頼性

- 60年の歴史
- 近代コンピュータテープは転写、定期巻き戻しの心配無用
- 加速試験でLTOメディアは19年以上性能変化なし

- 大容量、コンパクト

- LTO5カートリッジはVHSの半分サイズでHDビデオ50時間分、4MBの写真375,000枚を記録可能

- 高速な転送速度

- デジタル化、ファイル化する作業の時間を短縮
- LTO5の場合、2時間HDビデオを約6分で、4MBの写真を1分間に2100枚コピー可能

* 保管環境の空調費用等は含まず。

コンピュータ用テープメディアの活用 つづき

- 低コスト

- 容量あたりのコスト、消費電力は最少クラス
- 保管時電気代不要*

- 将来性

- 8世代までのLTOロードマップ
- テープのファイルシステム化によりいっそう使いやすく身近に

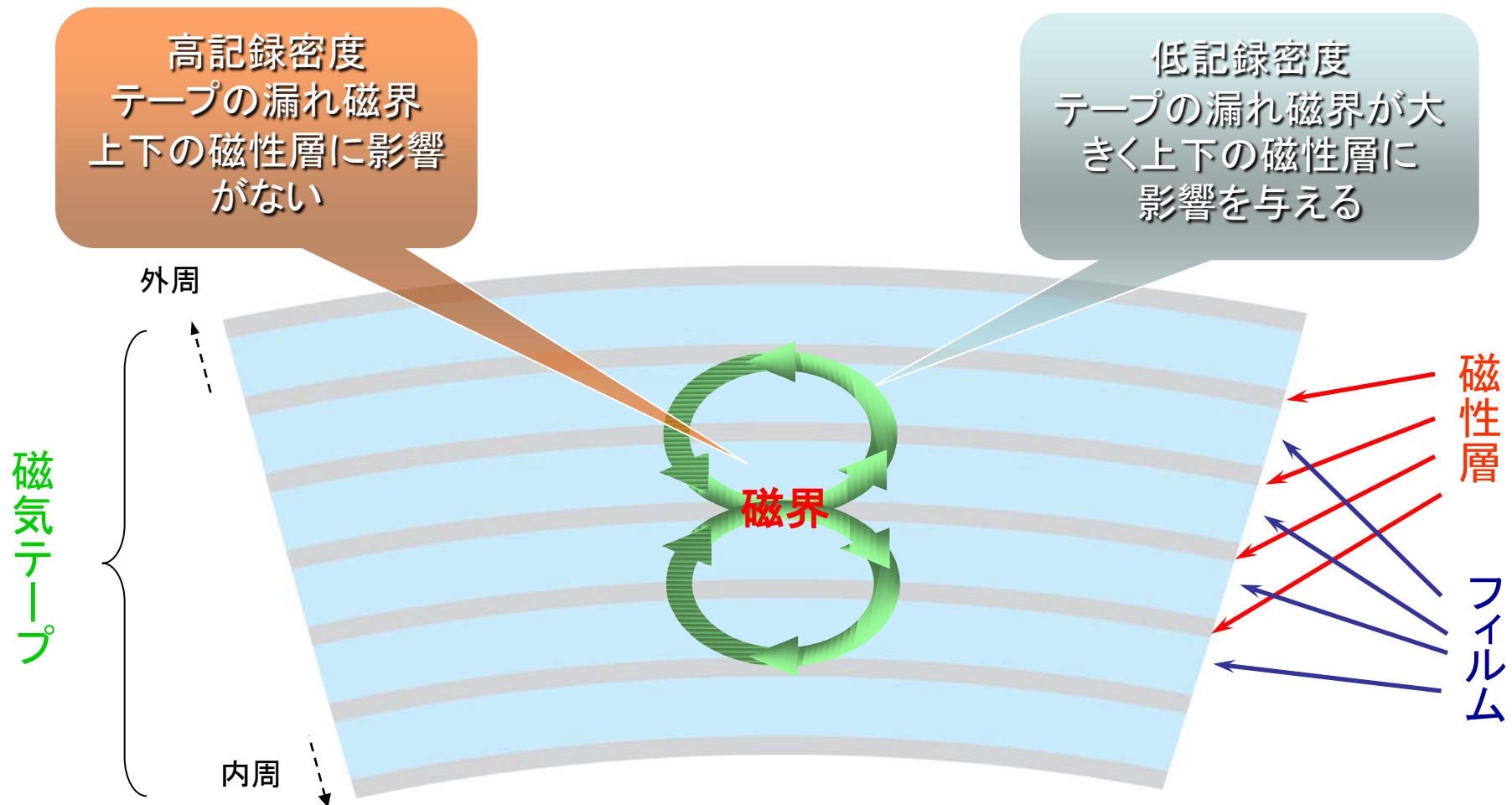
* 保管環境の空調費用等は含まず。

信頼性 - 長期保存性

- テープは記録した媒体をドライブから取り外して保管可能
 - リムーバブルの媒体
 - オフラインでオペミスから保護
 - オフサイトで災害時にも安心
 - もちろん電気代も不要
- 同系列のテクノロジーであれば、最新のドライブに入れ替えても、前世代の媒体を読み出すことが可能
- テープ媒体は、ベースフィルム・磁性体・バインダーなどの素材の改良や生産技術の改良により高品質となり、長期保存性が向上。定期的なテープの巻き直しは不要

現在主に用いられているテープ（メタル及び蒸着）では、適切な使用及び保存条件であれば、30年以上の寿命が予測されている。ただし、システムベンダーからは、システム利用形態（製品の変遷など）を考えて、10年程度ごとのマイグレーションが推奨されていることが多い。

信頼性 – 磁気転写について



現在のテープでは、磁気転写の発生はない！

信頼性 – テープの巻き直しについて

1. 低テープテンション化

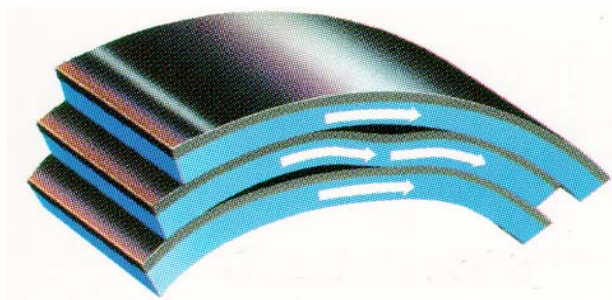
- オープンリールテープ : 1
- 初期のカートリッジテープ : 約 2/3
- 最近のカートリッジテープ : 約 1/3～1/4

オープンリールテープ
のテンション値を1とし
たときの、相対比表示

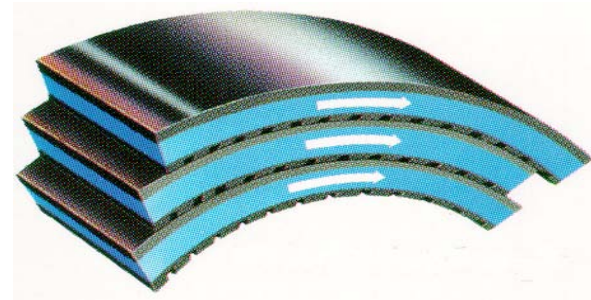
2. 磁性層の化学的安定性の向上

3. バックコート採用

- テープ表面圧縮効果を低減する
- 走行変動要因を効果的に除去し、テープを安定走行させる



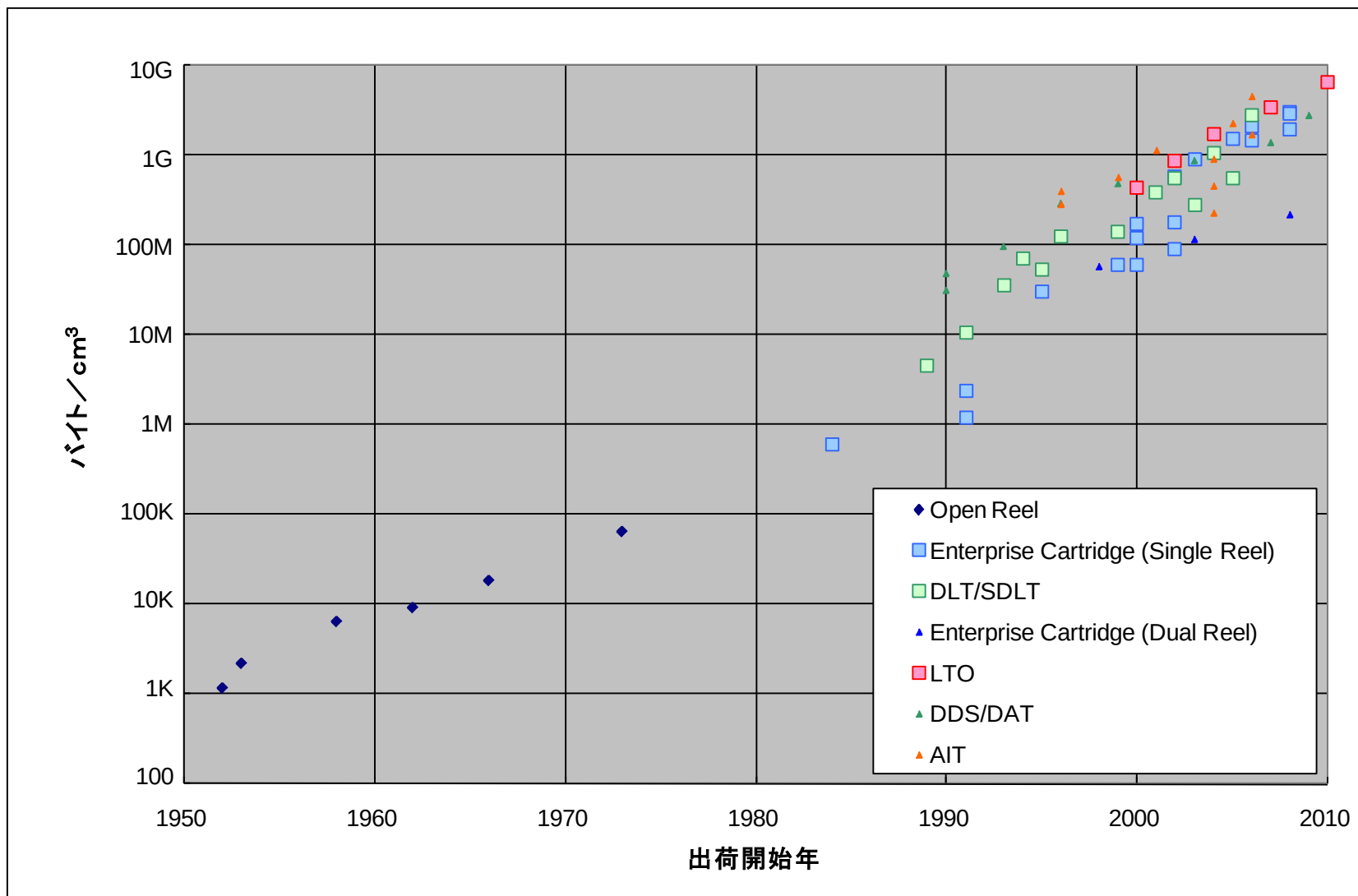
バックコーティングなし



バックコーティングあり

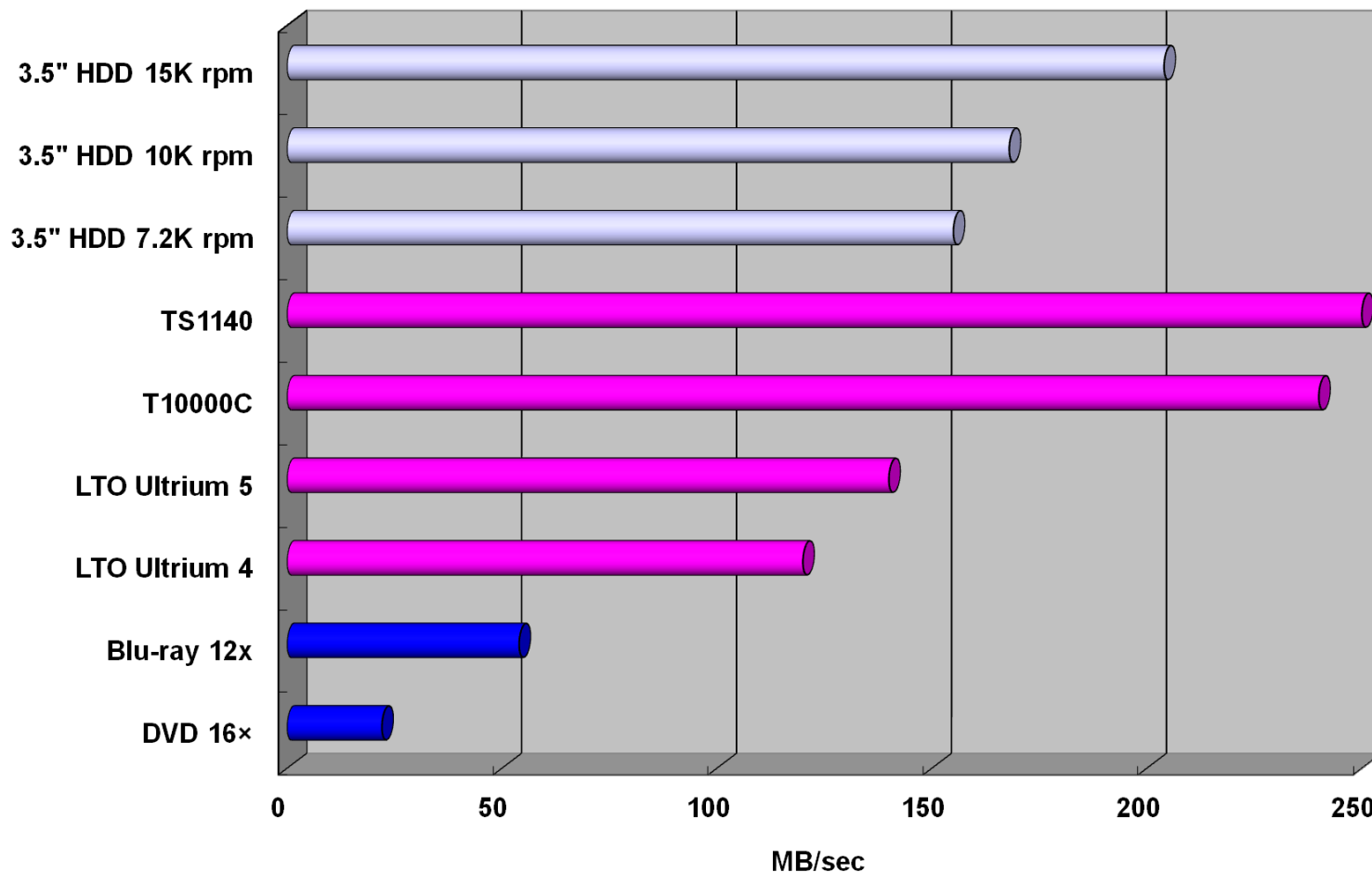
現在のカートリッジテープでは、テープの巻き直しは不要!!

大容量・コンパクト



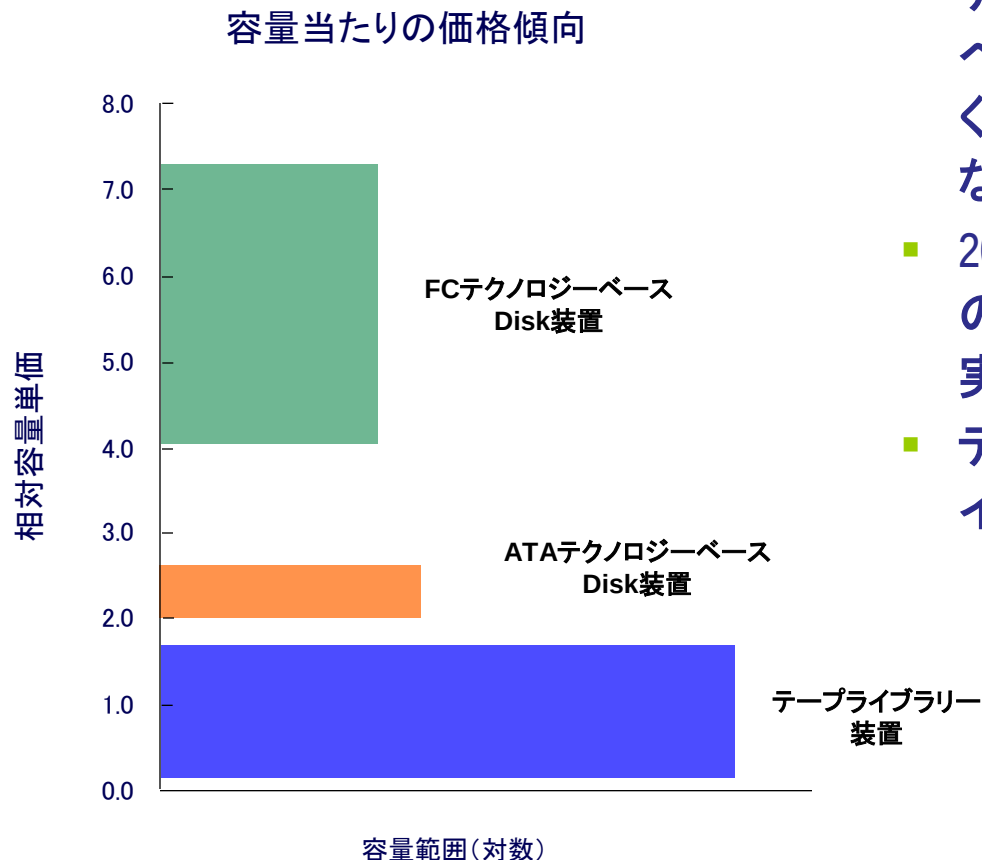
高速な転送速度

max sustained data rate



テープドライブは、非圧縮時の値
HDD/光Discは、最外周の最高値

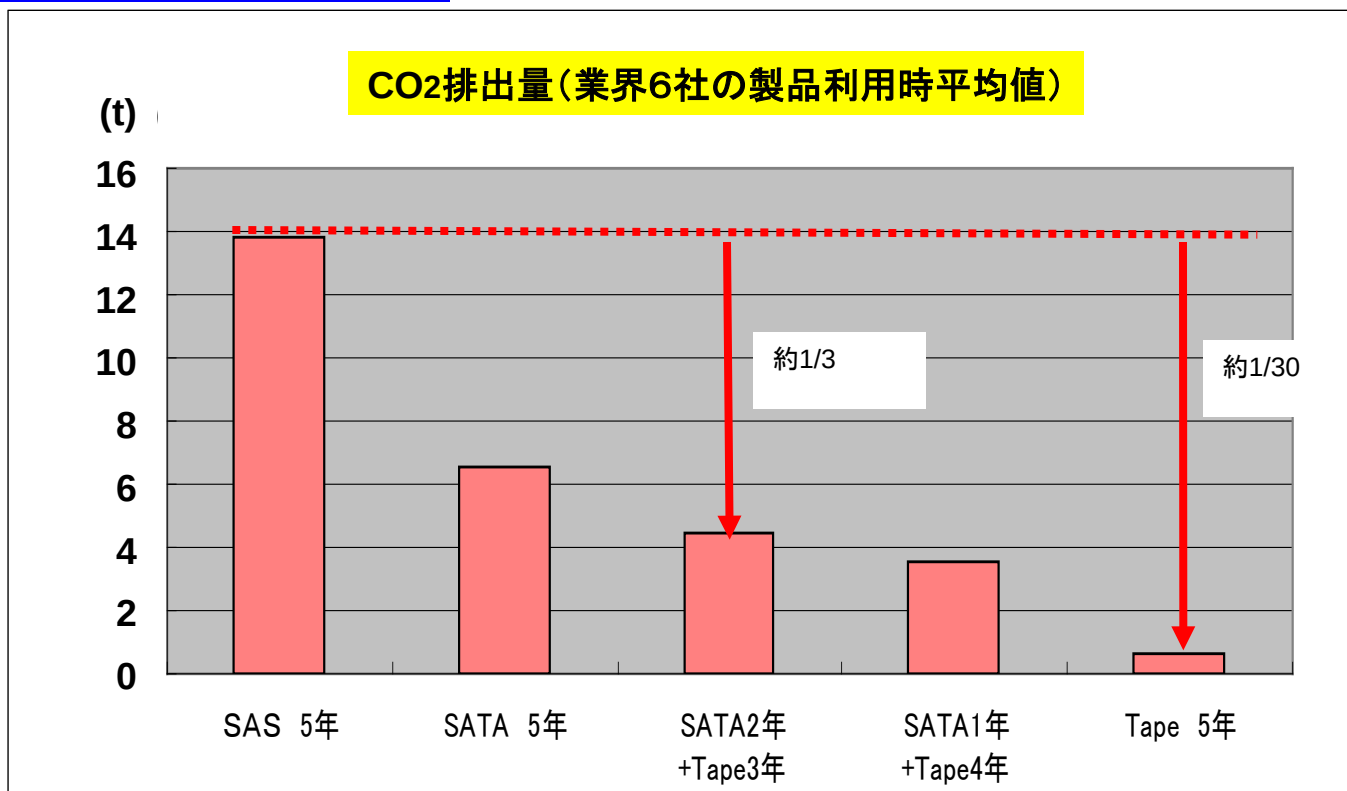
低コスト - 容量あたりの低価格性



- テープライブラリー装置ではHDDに比べ媒体が安価なため、収納巻数が多くなるほど、容量当たりの価格は安くなる
- 2007年現在では、19インチラック程度の設置面積で、400~560TBの容量を実現可能
- テープドライブを複数台搭載可能なライブラリー装置のみ
 - テープドライブを複数台搭載すると装置全体の価格は上昇する
 - 比較的安価で低容量帯のオートローダは含んでいない

低コスト - ストレージの消費電力比較

テープストレージを利用することによって、CO₂年間排出量を1/3～1/30程度に削減

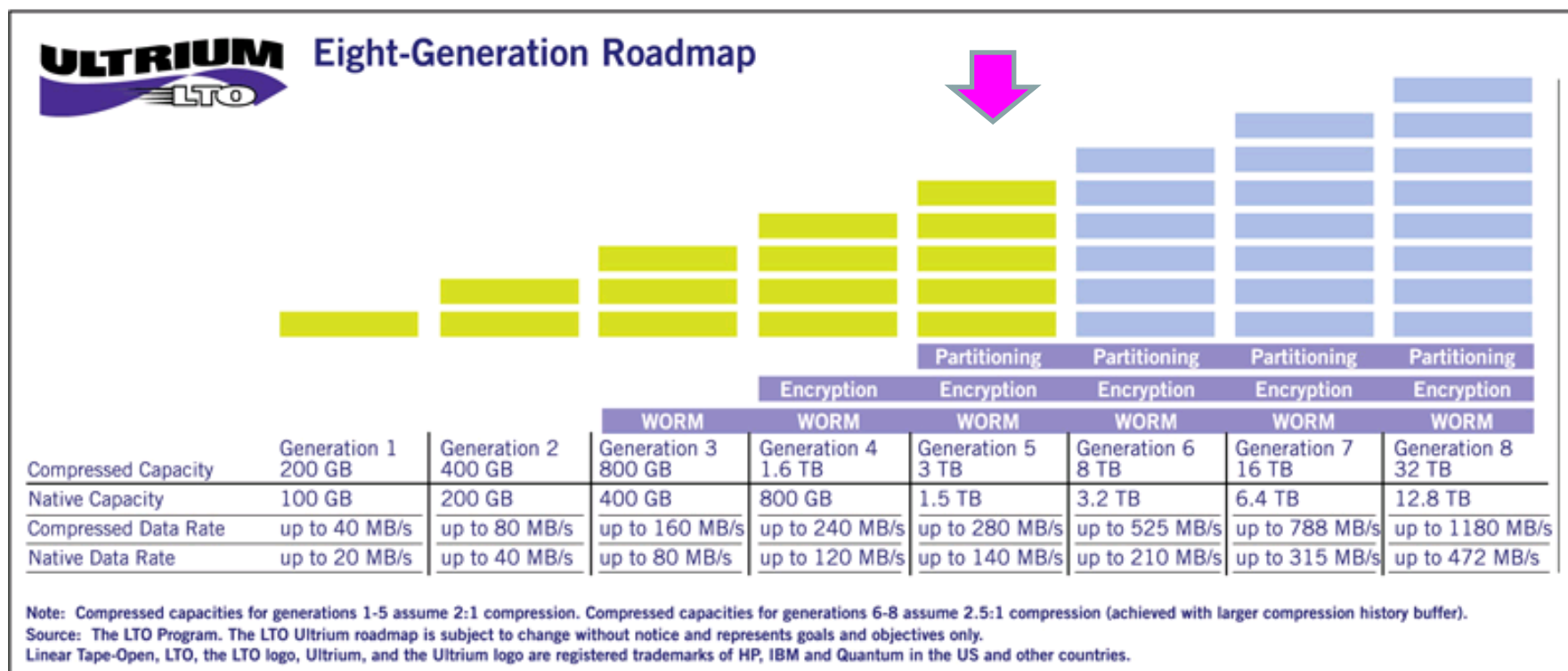


注1. 計算に使用したデータはカタログ値です。

注2. 各装置への保存期間は推奨するものではなく実際の運用では、データ毎に時間の経過に対応して最適な記憶装置とそれぞれの保存期間を決める必要があります。

将来性 - 進化を続けるテープメディア

LTO テープ技術のロードマップ



LTOコンソーシアムHPより <http://www.lto.org/technology/roadmap.html>

将来性 - テープのファイルシステム化

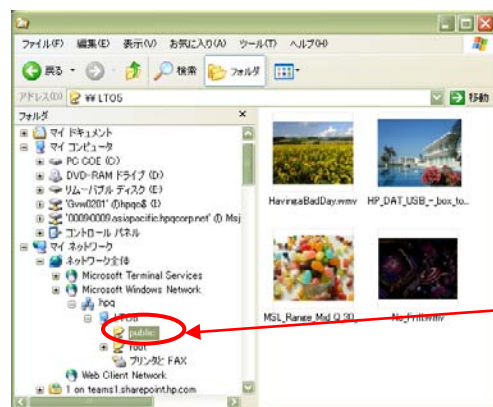
- テープをUSBメモリー、CDと同様に使える
- パーティション機能とLTFSで実現
 - LTO 5からサポート

ファイルシステム情報
(インデックス)

ファイルのデータ情報
(データ)

ファイルインデックス
ファイル名と位置情報

ファイルポインター



テープカートリッジ

テープ(LTO)導入メリット その他にも

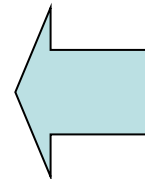
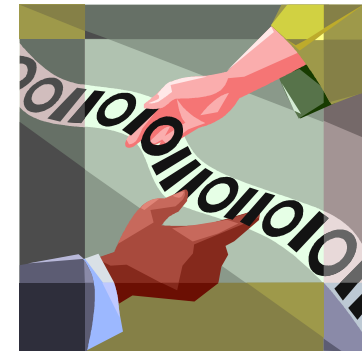
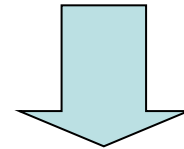
- 拡張性
 - テープライブラリにカートリッジを追加するだけで容易に容量拡張可能
- 改ざん防止
 - WORM (LTO3以降サポート)
- データ保護
 - オフライン保管可能 (オペミス、ソフトウェアバグ、ウイルス・ハッカーの影響を受けない)
 - 外部保管可能 (災害対策)
 - 暗号化 (ハードウェア暗号化による低コスト、パフォーマンスの保証) (LTO4以降サポート)

デジタルデータの長期保存という課題

データの保存とは？

目的 = 情報の再利用

- 情報のデジタルデータ化
- デジタルデータの解釈
- デジタルデータの保存



情報のデジタルデータ化

写真の場合

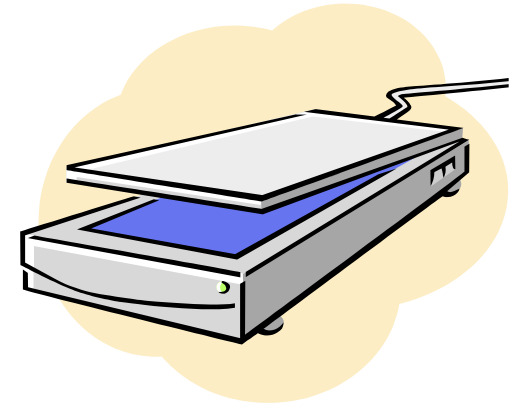
- スキャナーでデジタルデータにする

デジタルデータのフォーマットは？

- JPEG TIFF RAW ...
- 将来にわたって解釈可能な必要がある

一緒に保存するデータは？

- 写真のタイトル
- 撮影者・場所・時間...
- カメラ・レンズ・フィルム...
- 印画紙
- 写真にできたシミ？
- 色褪せ具合??
- 傷ついた箇所???



コンテンツ と メタデータ

⇒ 情報の種類・保存の目的によって様々

デジタルデータの解釈

デジタルデータを解釈して、情報を取り出す。

課題：デジタルデータ化した方法（フォーマット・メタデータなど）がわからないと、情報に戻せない。

解決策：

1. 遠い将来でも解釈方法が失われない方法でデジタルデータ化する。
= 標準化
2. より新しいデジタルデータ化の方法に変換することを繰り返す。
= 論理マイグレーション



デジタルデータの保存の課題

記録したデジタルデータを

記録した通りに間違い無く読み出す。

**課題： 100年以上の長期に亘ってデータの読み取りを
保証できるデジタルデータ記録方法が存在しない。**

－ ふたつの寿命

- 記録メディアの寿命
- 記録・再生システム(ハードウェアやソフトウェア)の寿命

デジタルデータの保存可能期間

	メディア	システム (ハードウェア)	システム (ソフトウェア)
保存期間	メディアがデータを保存していること。(媒体寿命)	メディアからデータを読み出し、システムのメモリに格納できること。	ハードウェアで読み出したデータを他のプログラムやユーザーが取り扱えること
3年	通常使用可能	データの保存で使用したハードウェアが使用可能	対応ソフトウェアで使用可能
10年	テープ・光ディスクなら読み出し可能。	新製品は無くても、動作するものが残っている可能性が高い。	上位互換システムで使える場合もある
30年	テープ・光ディスクなら読み出し可能という資料もある。	<u>通常、残っていない。</u>	<u>通常、残っていない。</u>
100年	<u>電子媒体では難しい。</u>	残っていない	残っていない

JEITA テープストレージ専門委員会 長期保存分科会による想定であり、実際のケースにおいて必ずしも当てはまるものではありません。

デジタルデータ保存の手法

解決策： 3つの手法

- ミュージアム手法

- メディアはそのまま保存。システムもそのまま保存。記録当時のまま保存・保守を続ける。

- エミュレーション手法

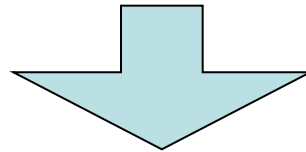
- メディアはそのまま保存。システムは更新するが古いメディアも読めるように作る。

- マイグレーション手法

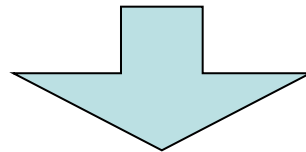
- 古いメディアから新しいメディアへ。システムも更新。

デジタルデータ保存の解決策

デジタルデータを100年以上保存するには、
現在の技術では困難である。

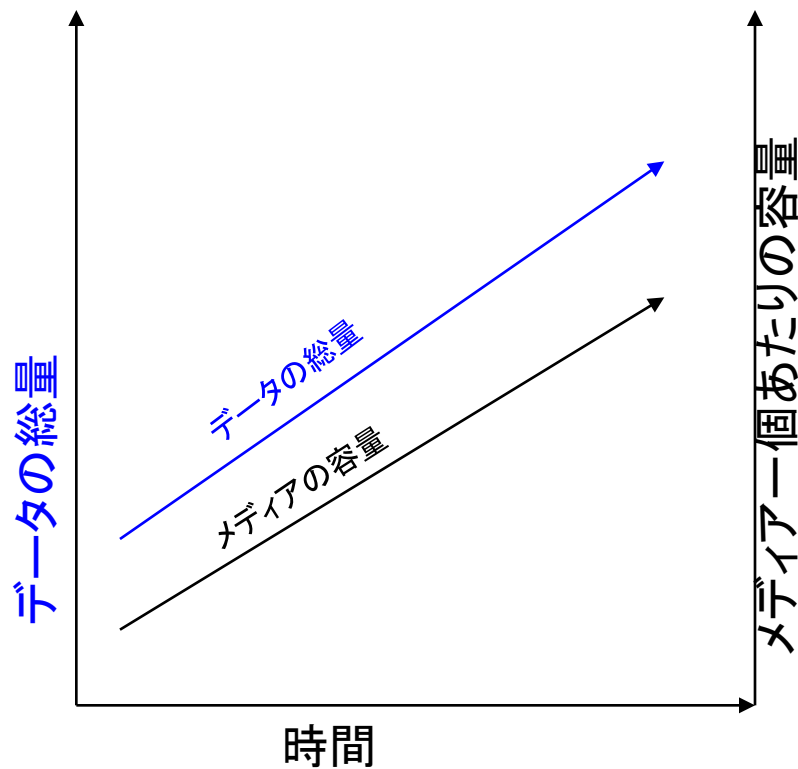


継続的にマイグレーションを繰り返し、
データを読み出し可能な状態に保つ。

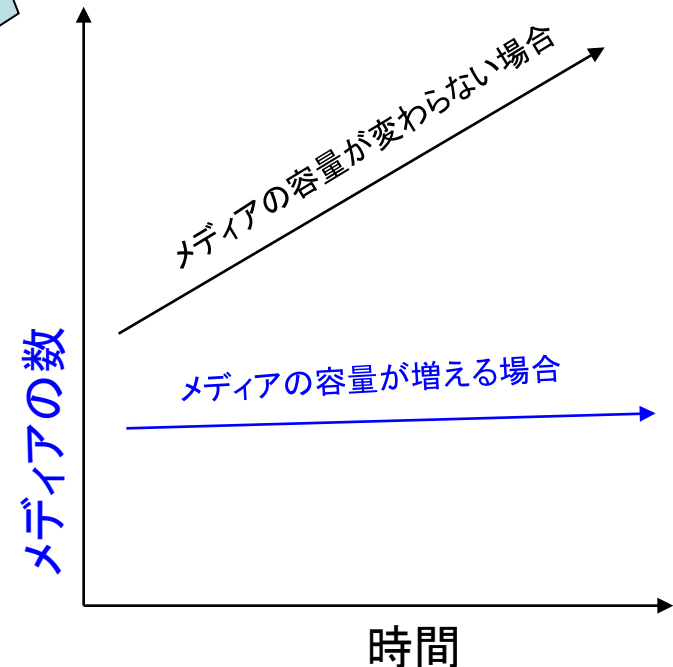
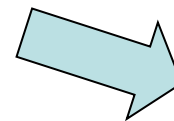


**(事実上)100年以上の長期保存が
可能な記憶装置を実現できる。**

長期データ保存における物理メディア数



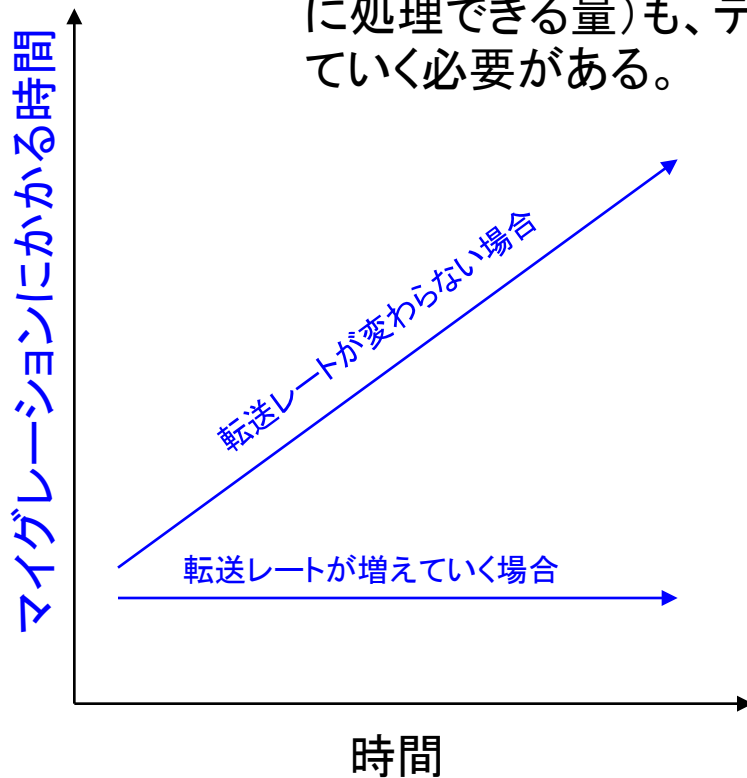
保存データを最新のメディアに移動することによって、物理メディアの総数を増やさないことができる。



保存すべきデータ量が増大するのと同様に、物理メディアあたりのデータ容量も増大していくと考えられる。

データマイグレーションの処理帯域

一定の間隔で継続的にデータマイグレーションを行うには、データマイグレーションの処理帯域（一定時間内に処理できる量）も、データの増加に合わせて、増大していく必要がある。



考え方:

一定の間隔でデータマイグレーションを行うために、データマイグレーションの処理帯域は総データ量の増分と比例して増大しなければならない。

解決案:

データマイグレーションの並列化

- データパスの数を簡単に増やせる構成
- リムーバブルストレージによる distributed storage

長期保存＝マイグレーション＋自動化

データマイグレーションに求められる要件

マイグレーションによる長期保存システムの要件

- 手に入れやすい装置を使用する

一般に市販されているものなど、代替品が手に入る装置。またメディア・装置にかかるコストも重要

- 信頼できる装置を使用する

多くのユーザーに広く使われている装置などはある程度の信頼性を備えていると考えられる。装置が故障すれば保管データが失われる可能性が高まる。

- マイグレーション先のシステムが、保存環境を改善する

現在の保存環境を改善する方向にデータを移行させなければ、保存データはその損傷の危険によりさらされることとなる。

- マイグレーション対象データがマイグレーション先のシステムでも読出せる

マイグレーションしたデータを読み出す環境が常に存在することを保証する必要がある。

マイグレーション時のデータ処理要件

- **マイグレーション対象デジタルデータを、そのままの形でマイグレーションする**
 - マイグレーションの過程でデータの変換を行った場合、その変換方法が失われた場合、元のデジタルデータに再変換することができなくなる。このような事態を避けるためには、記録したビット列に対してどんな種類の変換も加えることなく、そのままの形でマイグレーションを繰り返すことが重要である。
(論理マイグレーションは、情報のデジタル化の方法を変更することであり、元のデジタルデータに再変換する必要はないので、この要件の対象外である。)
- **マイグレーションの際には、対象デジタルデータと関連する情報を一緒にマイグレーションする**
 - デジタルデータの解釈に必要な情報のことをメタデータと呼ぶ。デジタルデータのマイグレーションにおいては、メタデータをデジタルデータ本体と一緒にマイグレーションすることが必要である。

マイグレーション作業の要件

- **マイグレーションが正常に終了したことを確認する**

デジタルデータのマイグレーションが正常に終了していなければ、そのデジタルデータは完全とは言えず、結果的にデジタルデータを失ったことと同じことになってしまう。よってマイグレーションの正常終了を確認することは重要である。

- **マイグレーションの実行とその結果を記録する**

デジタルデータのマイグレーションを実行した際に、その結果などを記録することで、記録したデジタルデータの完全性や原本性を確認することができる。

- **マイグレーションにおける戦略**

ただ漫然とマイグレーションを繰り返すだけではデジタルデータの長期保存は実現できない。ここで述べているそれぞれの必要要件を満たす記録方式を用いたシステムを構築することは、マイグレーションにおける正しい戦略のひとつであると考えられる。

またマイグレーションの周期についても、それぞれのシステム構成に応じて適切に設定する必要がある。

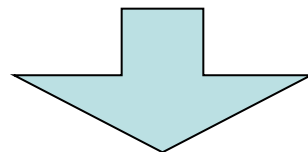
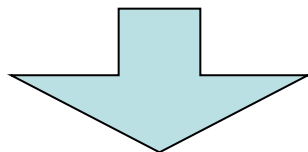
デジタルデータの長期保存用 標準フォーマットと 自動マイグレーションシステム

デジタルデータの長期保存用標準フォーマット

- 必要要件を満たすフォーマットを策定
 - マイグレーション用のメタデータを定義
 - 関連するファイルをまとめてマイグレーションすることを意識
 - テープ用ファイルシステムを使用することにより、ファイルレベルで他メディアとの互換性を保つ
 - データの利用者が、マイグレーションが行われたことを知らなくても、目的のデータを読出すことができる仕組み

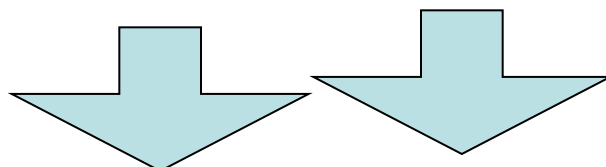
標準フォーマットの利用

デジタルデータの長期保存用フォーマットの標準化



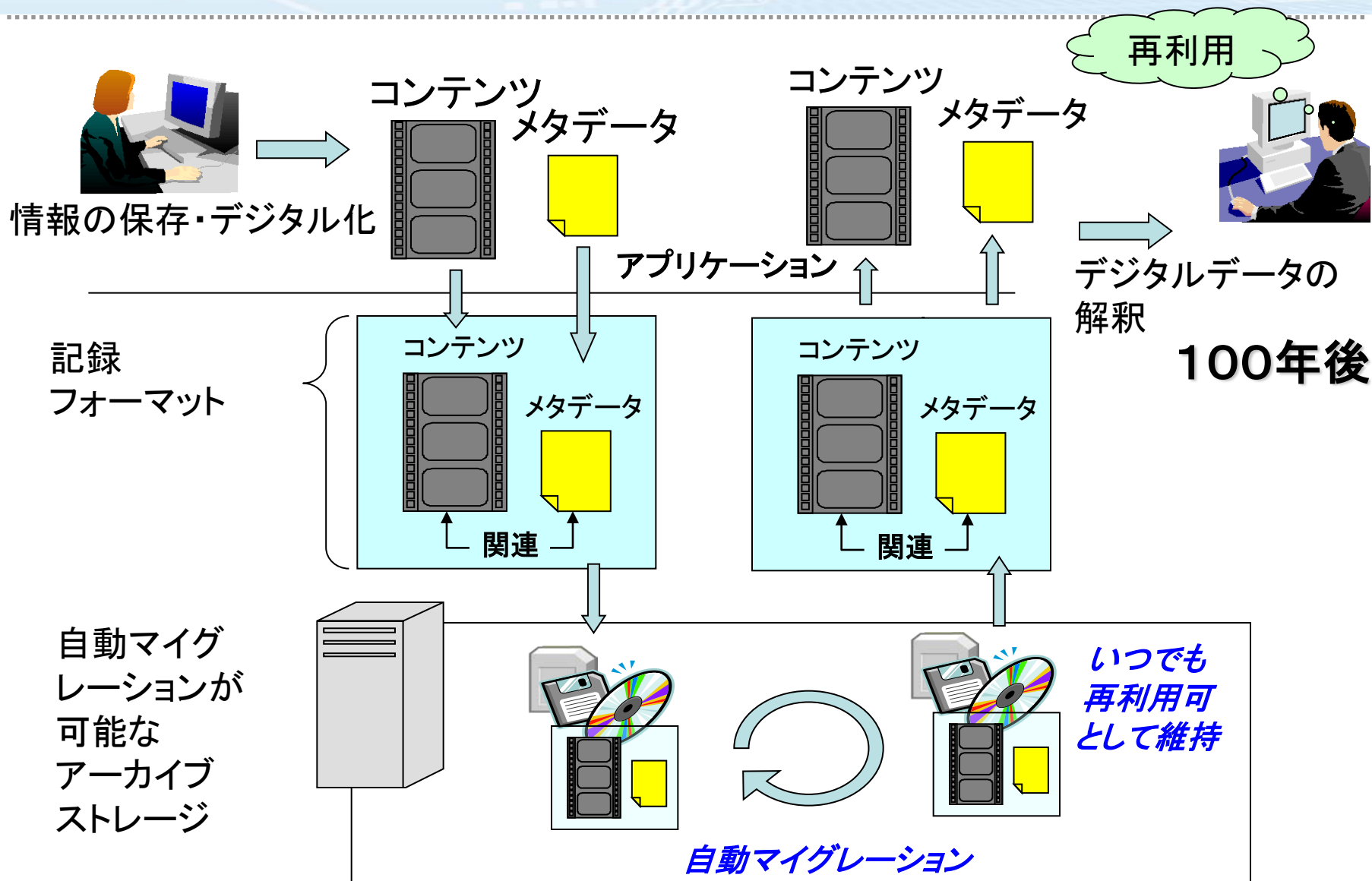
異なるシステム・メディア間での継続的なマイグレーションが可能になる。

自動的にマイグレーションを繰り返すことが可能になる。



特定のシステムにしばられない
ユーザーシステムから独立した
マイグレーションシステムへ

自動マイグレーションシステム



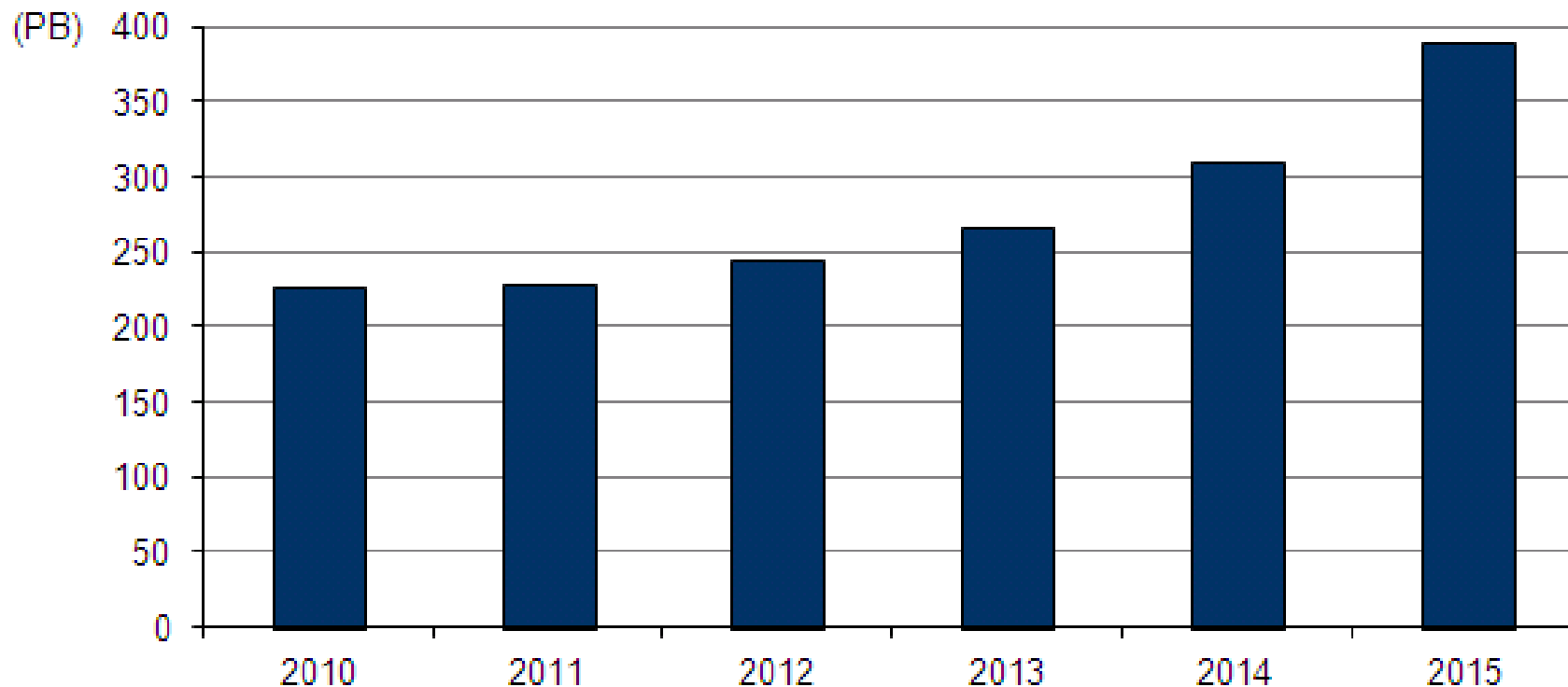
放送映像業界での活用事例

放送業界でのトレンドと課題

- 業界のトレンド.... 工数削減、効率化
 - ファイルシステム化
 - データのコンバージョン不要
 - 管理の容易化
 - 作業の効率化
- 課題
 - 過去のビデオ資産の運用が非効率
 - ファイルシステム化への移行
 - 現行ビデオテープシステム保守終了
 - 過去のビデオ資産のマイグレーションが必要
 - メディアの選択が課題
 - コスト、転送速度、記録容量、継続性、将来性

国内放送コンテンツ向けストレージインフラ市場 容量予測

- ・ 2010年 国内市場規模は、投資金額583億3,500万円、容量規模225.9PB
- ・ 2010年～2015年の容量の年間平均成長率11.5%、記録媒体の構成比は大きく変化
- ・ ファイルベースを柱にアーカイブやアセット管理でビジネス機会が拡大



出典: Source: IDC Japan, 10/2011



課題解決のためには.... LTO 5のメリット

放送業界でのLTO導入メリット

- 安い

- 従来のテープ及びディスクシステムから安価なLTOへ
 - 2日分(約50時間)以上のHDデータが1本に
- 保管時電気代不要*

* 保管環境の空調費用等は含まず。

- 速い

- LTO 5の転送レートは140MB/s(1120Mbps)
 - リアルタイムHD (32-64Mbps)の約20倍

- 便利

- データの大容量化に伴う高容量メディアの要求
 - LTO 5では1カートリッジあたり1.5TB
- フォーマット等の制約なし
 - HD,SD,グラフィック、等混在記録可能
- テープのファイルシステム化 (LTO 5ではパーティションを採用)
 - HDD,USBメモリと同様に使用できるためデータ交換媒体として最適

放送業界でのLTO導入メリット

• 安い

- 従来のテープ及びディスクシステムから安価なLTOへ
 - 2日分(約50時間)以上のHDデータが1本に
- 保管時電気代不要*

* 保管環境の空調費用等は含まず。

• 速い

- LTO 5の転送レートは140MB/s(1120Mbps)
 - リアルタイムHD (32-64Mbps)の約20倍

• 便利

- データの大容量化に伴う高容量メディアの要求
 - LTO 5では1カートリッジあたり1.5TB
- フォーマット等の制約なし
 - HD,SD,グラフィック、等混在記録可能
- テープのファイルシステム化 (LTO 5ではパーティションを採用)
 - HDD,USBメモリと同様に使用できるためデータ交換媒体として最適

放送業界でのLTO導入メリット

- 安い

- 従来のテープ及びディスクシステムから安価なLTOへ
 - 2日分(約50時間)以上のHDデータが1本に
- 保管時電気代不要*

* 保管環境の空調費用等は含まず。

- 速い

- LTO 5の転送レートは140MB/s(1120Mbps)
 - リアルタイムHD (32-64Mbps)の約20倍

- 便利

- データの大容量化に伴う高容量メディアの要求
 - LTO 5では1カートリッジあたり1.5TB
- フォーマット等の制約なし
 - HD,SD,グラフィック、等混在記録可能
- テープのファイルシステム化 (LTO 5ではパーティションを採用)
 - HDD,USBメモリと同様に使用できるためデータ交換媒体として最適

放送業界でのLTO導入メリット

• 安い

- 従来のテープ及びディスクシステムから安価なLTOへ
 - 2日分(約50時間)以上のHDデータが1本に
- 保管時電気代不要*

* 保管環境の空調費用等は含まず。

• 速い

- LTO 5の転送レートは140MB/s(1120Mbps)
 - リアルタイムHD (32-64Mbps)の約20倍

• 便利

- データの大容量化に伴う高容量メディアの要求
 - LTO 5では1カートリッジあたり1.5TB
- フォーマット等の制約なし
 - HD,SD,グラフィック、等混在記録可能
- テープのファイルシステム化 (LTO 5ではパーティションを採用)
 - HDD,USBメモリと同様に使用できるためデータ交換媒体として最適

最新IT技術のあらゆる分野への活用

最新IT技術のあらゆる分野への活用

- 医療

- 電子カルテ
- グラフィックデータの高解像化

- 監視カメラ

- カメラ台数の増加(ロンドンの例)
- 高解像度化、カラー化
- Webカメラ等による低価格化、導入の容易化

- 研究機関

- 測定デバイス、システムの進化
- 莫大な捨てられないデータ

まとめ

長期アーカイブは

- あらゆる情報が
- 適切なタイミングで
- 簡単に出し入れできて
- 安全にかつ低コストに保存されていること

コンピュータ用テープは

- ファイル化されたデータなら何でも記録
- 高速:1120Mbps [140MB/s] (LTO5)
- LTFSによりUSBディスクのように使える
- オフライン、オフサイト保管でオペミス、ソフトウェアバグ、ウイルス・ハッカーの影響を受けない、災害対策
- 暗号機能で情報漏洩のリスク低減

ありがとうございました