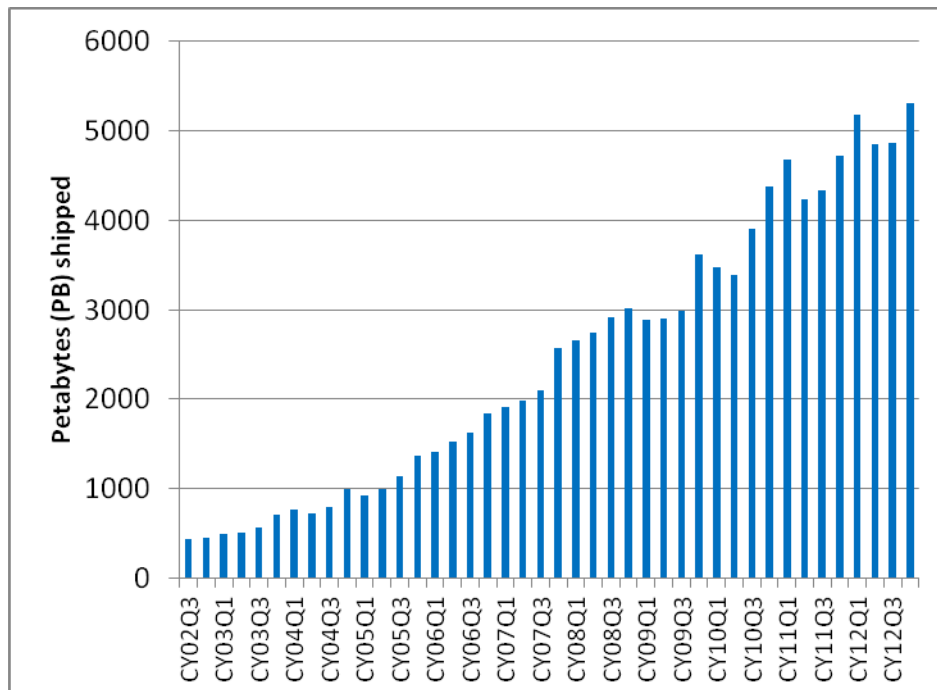

第1章 テープストレージの未来

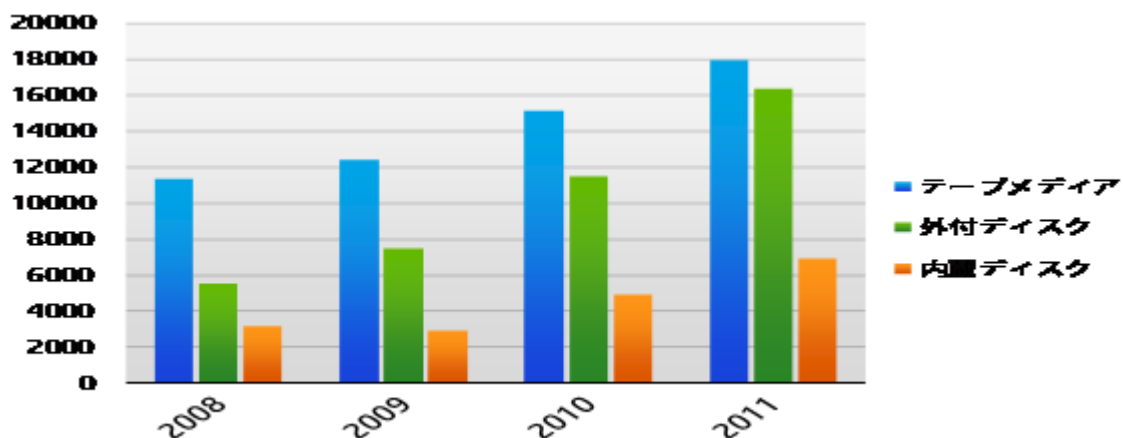
—米国 IT 業界データからテープストレージの将来を検証— 2013/7

ビッグデータ、クラウドの時代、今後多くのデータが集まるデータセンターにおいて、従来以上にテープストレージは、欠かすことのできないものになりつつあるようだ。特に、コスト面、長期アーカイブ、災害対策のためのオフサイト保管といった観点から、テープストレージが果たす役割は依然として高く、今尚多くの米国企業がテープの開発、採用に投資を続けている。この章では、それを米国の記事、レポートから検証していきたい。

検証 その1：テープメディアの2012年出荷容量は20エクサバイト！

年々増加するデータ量に対する低コストストレージの要求の高まりは、テープメディアの出荷容量に現れている。2012年Q4のSanta Clara Consulting Groupのデータによるとテープメディア総出荷量は年々増加しており、2012年には20エクサバイト(非圧縮、1エクサバイト=1,000ペタバイト=1,000,000テラバイト)で、外付けディスク容量を超えている。今後もデータの爆発的増加に伴い、さらに増加し続けると見られている。





検証 その2：クラウドデータのおよそ50%はテープに保存される

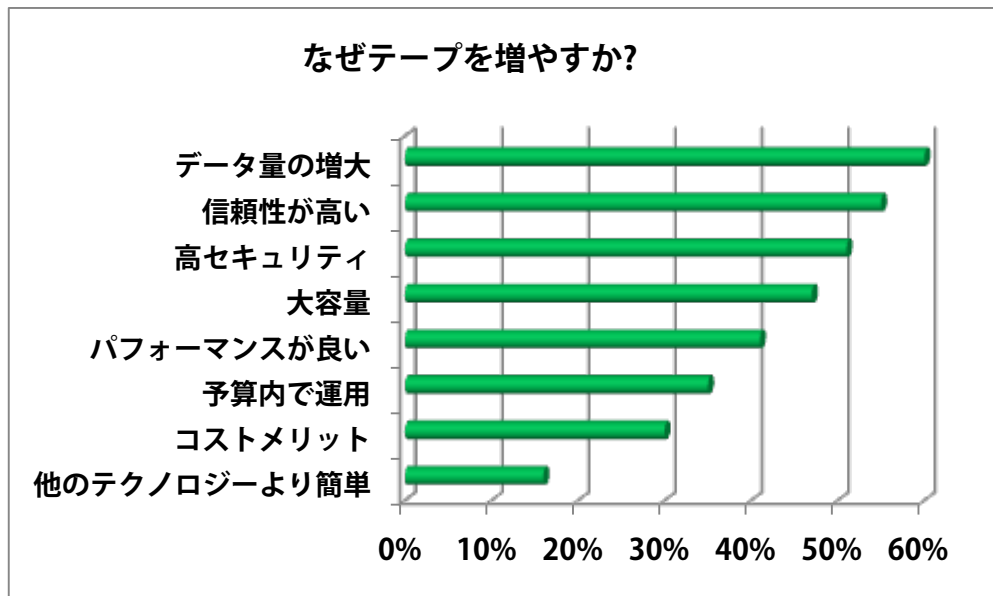
今後も継続してデータの集まるクラウド。Enterprise Storage Group のレポート[ESG Research Report_Digital Archive Market Forecast_June 2010] によると、2013年にはクラウドデータの50%がテープに保存されると予測されている。

クラウド業者のデータセンターでは SLA (Service Level Agreement) に沿って、tiered storage architecture(階層ストレージ構成)、または Hierarchical Storage Management (HSM：階層型ストレージ管理)などを使い、データをより効率よく、安全に管理する必要がある。効率よく、つまり低コストで安全にデータを保管できるデバイスが必要とされているのである。

検証 その3：ビッグデータ時代、データ量の増大、信頼性、セキュリティが重要視されてきている

2012年にUSとカナダの企業を対象に実施された Storage IT decision makers の調査結果によると、77%のユーザーがテープをバックアップ、アーカイブに使用しており、今後3～5年でテープの使用を継続、または拡大すると答えたユーザーは87%に達した。その理由として、**データ量の増大**に対応できること、**信頼性の高さ**、**セキュリティの高さ**などが挙げられている。

ビッグデータ時代の今日、年率50%以上で増え続けるデータをどこに保存するのかは、喫緊の課題である。もちろん予算も限られるため低コスト、低消費電力であることも重要であるが、更には管理コストの面からも信頼性が高いこと、低コストで高いセキュリティが実現できること等が求められている。また拡張性が高く、容量の増設が簡単であることも、予想外のデータの増加に対応できるという点で、テープストレージが使いやすいと考えられているようだ。またオフラインで保管できるテープは、内部、外部からの意図的なデータへの攻撃、オペミス、災害等によるリスクに対して非常にセキュリティ度が高いと言える。

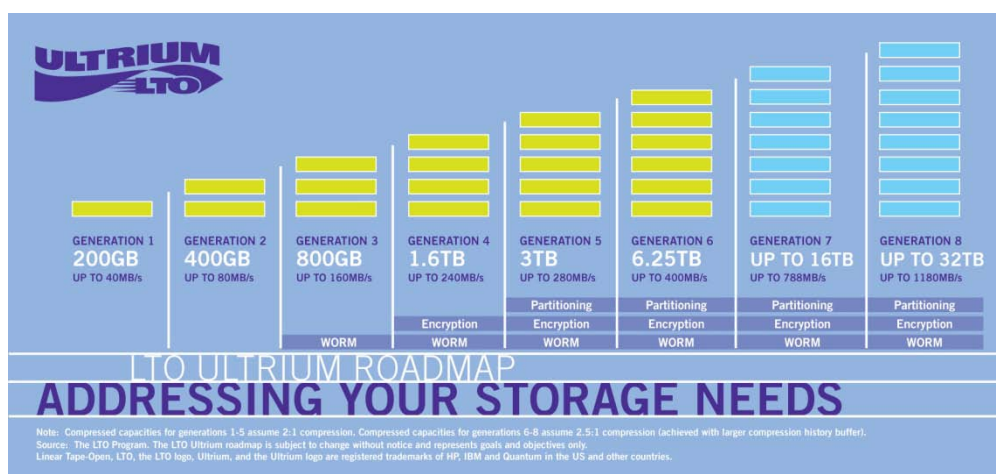


検証 その4：進化し続けるテープテクノロジー

世界で最初のテープドライブがこの世に現れてから既に 60 年を超えているが、テープの技術、パフォーマンスはそれ以来継続的に進化し続けている。最新の LTO Ultrium (以下:LTO) 6 テープドライブでは 2.5:1 圧縮で 6.25TB の容量を、一時間当たり約 1.4TB のスピードで記録できる。

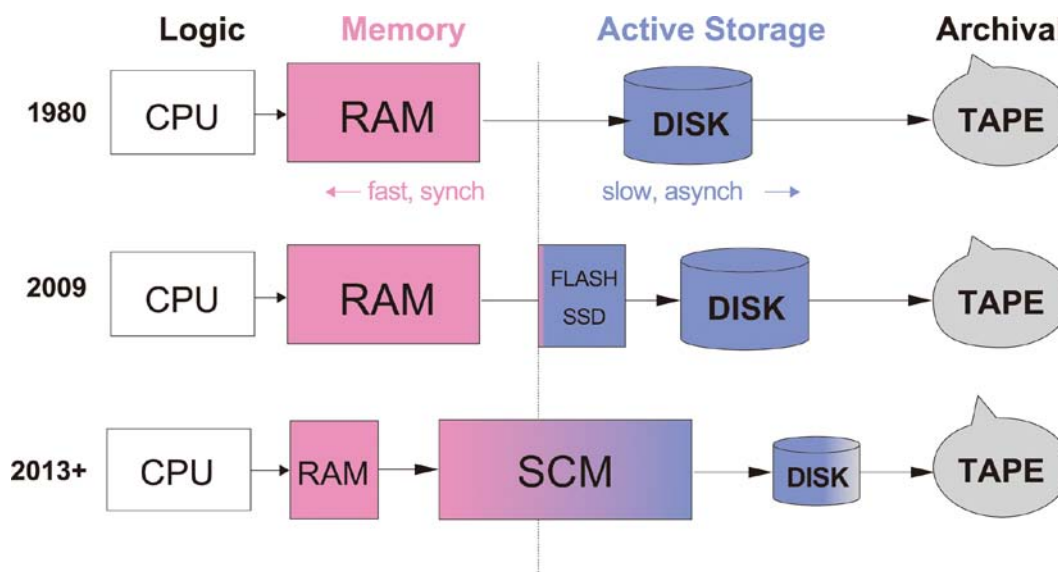
実際これは最初のテープドライブと比較して約 84 万倍もの性能アップ(転送レート)にもなっている。

多くのテープストレージベンダー、テープメディアベンダーは現在も最新のテープテクノロジーの開発にしのぎを削っており、これは裏を返せば、今後もテープストレージには非常に多くの期待が寄せられているという証でもある。



米 SNIA(Storage Networking Industry Association)の予測では、今までディスクが担っていたアクティブ

ストレージの範囲は今後 FLASH SSD に一部切り替わり、2013 年以降には FLASH SSD は SCM(Storage Class Memory)に置き換わっていくとしている。さらに SCM はその守備範囲を拡大し、メモリーの範囲では RAM の領域を、アクティブストレージの分野ではディスクの領域に拡大していくとの見方をしている。一方テープはアーカイブの分野で継続的に主流であり続けるとの見方をしている。これはアーカイブに求められる要件を満たすテクノロジーがテープ以外に見つからないためといえるだろう。



SNIA, Advanced Storage Industry Standards and Relationships for Enterprise and Cloud Computing. Data Storage Expo 2009

米国データセンター運用要件からテープストレージの将来を検証

現代のデータセンターでは、テープストレージはやはり欠かすことのできないものになりつつあるようだ。以下ではデータセンターのデータ保護に必要な要件を具体的に解説し、テープストレージに期待されている役割について説明していこう。

Information Week の調査によると、328 ユーザーのうち 3/4 が 1TB 以上のデータを管理していることがわかる。これはかなりのデータ量だろう。さて次に問題になるのは、これらのデータをどこに保管するかということだ。以下その鍵となる 3 つの要素について説明していきたいと思う。

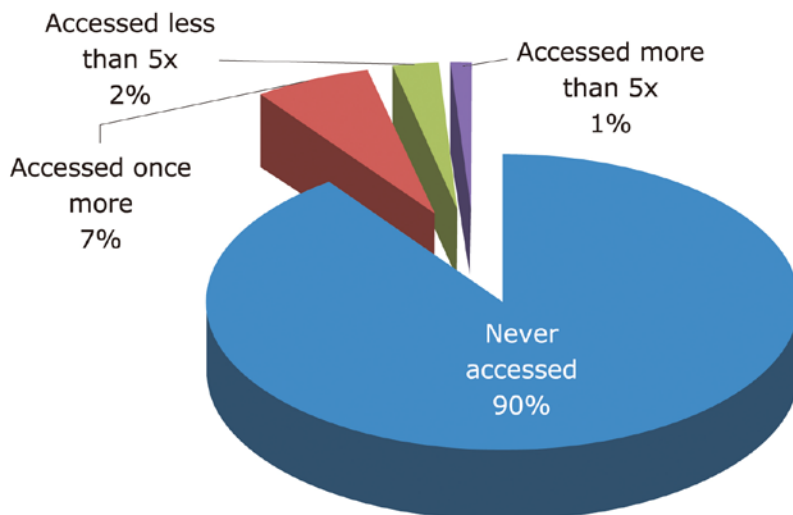
要素その 1：低コストのストレージ

今後も増え続けるデータを、単純に次から次へとディスクを増設して保存するのはあまり賢い方法とはいえない。なぜならそのような運用はコスト的に高くつくからだ。

ここで重要なのは、コスト計算はハードウェアの購入コストだけでなく、消費電力、空調、さらには

データセンターのフロア管理コスト、メンテナンスコストなども考慮する必要がある点だ。一般にアクセスされるデータの 80%以上は作成されてから 90 日以内のものであると言われている。最近 University of Santa Cruz により実施された調査の結果では、NAS に保存された 90%以上のデータが二度とアクセスされず、また 6.5%のデータがたった 1 回しかアクセスされていないといったことが明らかになった。

Research into disk access patterns on a NAS



Source: University of California, Santa Cruz

それにもかかわらず、法規制等さまざまな理由からこれらのデータも長期的に保管しなくてはならないわけだが、それらを全てディスクに保管するのではコストがかかりすぎる。その結果として出てきた概念が **tiered storage architectures**(階層ストレージ構成)という考え方である。簡単に言えば、頻繁にアクセスされるデータはアクセス時間の短いデバイス(数秒程度)に、そうでないものはアーカイブデータとしてテープのような低コストのデバイスに保存するというものだ。

容量コストは一般的に代表的なテープ媒体である LTO 6 カートリッジで Gbyte あたり 4.4 セント(約 3.6 円)[2012 年 12 月現在のレート換算]とかなり安いですが、電力コストを含む TCO(total cost of ownership)で考えた場合、その差はさらに大きくなる。なぜならシステムディスクの場合、データアクセスがあると無いとに関わらず常時ディスクを回転させる必要があるが、テープの場合はデータの保存時と読み出し時以外、テープカートリッジはライブラリーの中に保管されているだけなので、余計な電力は必要ないためだ。

データアーカイブの運用方法の調査としては Clipper Group の調査レポートが有名だが、ここでは **tiered storage architecture**を採用しているデータセンターでの 5年間のデータ保管運用方法について、ディスクのみの運用(D2D)とディスクとテープのハイブリッド(D2D2T)運用の比較をハードウェアとメディアの購入コスト、さらには電力とデータセンターのフロア面積を考慮して行っている。

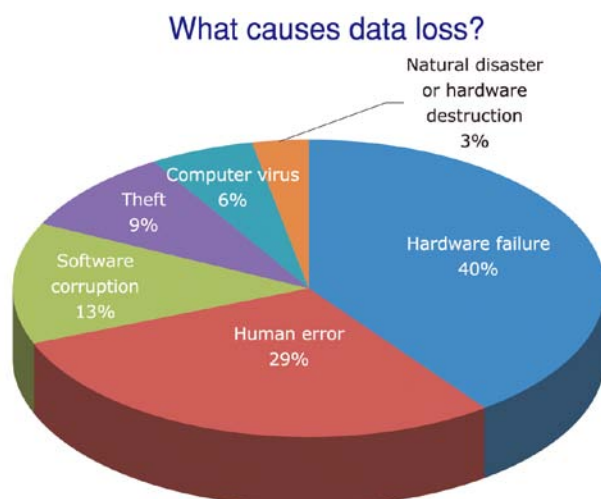
Clipper Group の試算の結果、データアーカイブにかかる購入費用は、ディスク(SATA)のみのシステム構成では、テープを組み合わせた場合に比べると 23 倍以上もコスト高になるとの結果が出た。さら

に電力コスト等も考慮すると、なんと 290 倍もの開きが発生するという結果も出ている。

要素その 2：複数階層のデータ保護

データ保護は異なる複数の壊滅的なデータ損失のリスクを、どれだけ緩和できるかということに尽きる。それらはハードウェアの問題かもしれないし、人為的なミスかもしれない。さらにはソフトウェアの問題、盗難、コンピュータウイルス、自然災害といずれもひとつのテクノロジーで補えるものではない。

ちなみに上記は Pepperdine 大学の Dr David M Smith 氏の調査結果での順番で、ハードウェアの問題が 40%に対して人為的なミスが 29%とかなり多く、以下 13%、9%、6%、3%という割合だ。オフライン、オフサイト保管できるテープは人為的なミス、盗難、コンピュータウイルス、自然災害といった要因に対して強さを発揮できると言える。以下はより安全なデータ保護を行う上で必要な 3 つの項目である。



Source: Dr David M Smith, Pepperdine University

1.複数階層でのデータ保護

理想的には 3 つのコピーを異なる場所に保管することが推奨されている。またそのうち 1 つは地震とか洪水といった災害に対する復旧用に、遠隔地にデータを保管することが必要だ。

このような要件には、可搬性に優れているテープを使用することで、低コストで容易にデータを外部保管することが可能である。

2.最低でも 1 つのオフラインコピーを作る

これはシステムの分離を実現するためのもので、コンピュータウイルス、意図的な破壊工作、人為的なミスといったリスクからデータを保護するのに有効だ。

オフラインデータ保管といえばやはりテープに敵うものはないだろう。実績があり、運用も簡単、し

かも低コストで実現できる。

3. アクセスの無いデータの保護

「情報漏洩」という言葉は最近良く耳にするが、これが起きた時は単純にデータの紛失にとどまらず、社会的信用度の失墜、訴訟等により会社の存続に関わる事態に発展する可能性すらある。

KPMG の調査結果によると、世界中で情報紛失の影響を受けた人々の人数は 2008 年の 9200 万人から 2009 年には 1 億 9 千万人にも跳ね上がっており、この流れはさらに加速すると見られている。そのような背景から漏洩したデータを読めないようにする「暗号化」はデータセキュリティ戦略の最も重要な機能となってきている。

そのような中、LTO 4 以降のテープドライブでは低コストでパフォーマンスに影響の無い暗号化機能を実現し注目を集めている。

要素その 3：長期にわたるデータ保存性

いかなるアーカイブソリューションを選択する場合でも、保管メディアは十分なメディア寿命があるものを選択することが重要だ。一般的にはテープメディアはディスクの 4~6 倍の寿命があるといわれている。JEITA での加速テストでも LTO メディアは 20 年以上の保管寿命が確認されている。

また、もうひとつ重要なことは、増加するデータ量に対して簡単にかつ低コストで容量が拡張できることだ。テープライブラリーの場合は単純にカートリッジを増やすことや、新しい世代のテープドライブに交換することで簡単に容量を拡張できる。もちろんテープメディアのコストパフォーマンスの高さは言うまでも無いだろう。

ここまで説明してきたように、データセンターにおけるテープの役割は以下のように変わりつつある。

- ・リカバリー時間に余裕がある運用でのバックアップ
- ・古いデータ(アーカイブ)のディスクベースバックアップからの移行
- ・長期保管データの保存 法令順守のための高信頼性、低コストアーカイブ
- ・システムからの分離、災害対策のためのオフサイト保管

テープ以外のテクノロジーで上記要件を満たせるものは現在のところ見つかっていない。それが今尚多くの米国企業がテープの開発、採用に投資を続けている証拠でもあるのだ。中小企業の多い日本でもストレージ統合の波は迫ってきている。データセンター先進国米国でのこのようなうねりが日本の IT 産業に到来するのは意外と早いかもしれない。

(社)電子情報技術産業協会(JEITA) テープストレージ専門委員会

日本ヒューレットパッカーード(株) 井上 陽治