

来るべきデジタル社会の危機、記憶の保管性について ー超長期保管メモリ、千年メモリの必要性和その課題ー

小林 敏夫

神奈川大学 理学部 情報科学科

〒259-1293 神奈川県平塚市土屋 2946 番地

E-mail: tkoba@purple.plala.or.jp

■ はじめに

エレクトロニクス技術の目覚ましい進歩によって、デジタル技術が個人の生活と社会に浸透しはじめて 30 年足らず、今や社会を維持、活動させるために必須な技術となり、デジタル技術への依存は抜きさしならぬ状況になっています。それに伴い、レガシーな記憶媒体、すなわち紙、フィルムなどは、特殊な用途を除いて急速に姿を消していくと思われます。そのため、利便性の影に潜む危うさ対策の必要性和、媒体、システムの脆弱性、保管コストの上昇が課題として図書館、アーカイブ関係の機関によって指摘されています。一般にはこれまであまり目が向けられてこなかったのですが、デジタル化による記憶の保管性の脆弱化が進行し、人間社会の存在の根源である記憶の保管性に危機が生じています。もう一つの問題は、デジタル技術がもたらした情報爆発があります。デジタルボーンのデータ量は、2011 年には全世界でゼタバイト (10^{21} B) 規模になると予想されています。これまで、これらの問題に対して、電子社会を支えている半導体技術側からの取り組みが殆どなされていません。半導体デバイスには、この問題を解決する大なるポテンシャルがあり、機能面のみならず記憶の保管性においても大きな役割を果たすことができます。

■ 保管の現状

記憶の保管性の危機を引き起こしている原因の 1 つは、デジタル記憶媒体の脆弱性、すなわち CD、DVD、ハードディスク、フラッシュメモリなどの寿命が 5 年ないし 20 年程度しかもたないという事実によります。また、媒体運用を支えるシステムの寿命も似たり寄ったりです。さらに、デジタルデータの保管の問題を複雑にしているのは、単にデータが残るだけでは不十分で、必要な時にデータの意味を理解できることを保障することがデジタルデータの場合難しいためです。絵画とか紙に書かれた文章などであれば、見た目で程度理解することができます。しかし、ビット列の情報は、それだけで意味を理解することは困難です。デジタルデータを保管する際の基本的な手続きについて Open Archival Information System (OAIS) というものが提唱されています。現在 ISO の標準規格となっています。このように保管データとそれを解読するための付与データ (メタデータ) とを含むデータの構成についての検討は、進んでいるようです。しかし、メタデータの長期間に渡る意味理解を保障する手段とデータの物理的な保管性については、状況は変わっていません。さらに、媒体が劣化するという問題の他に、OS、読み出し装置が変わり古い媒体が読めなくなるという問題があります。これらが短命であるため現在長期保管戦略として、マイグレーションとエミュレーションという方法が推奨されています。マイグレーションとは、デジタル資産をある世代の技術から次世代の技術へシステムごと移行させるというものです。今はマイグレーションが主流ですが、定期的に大きなコストが発生し、継続的なコスト負担が大きな問題となります。さらに、今後データ爆発が本格化した際に、何時までマイグレーションが可能なのかとの危機が示されています。

理想的には、コストを掛けずに棚に置いておくだけで保管ができる (意味理解も保障される) ことが望ましい。また、デジタルボーンのコンテンツを保管した原本がそのままアーカイブ媒体になることが望ましい。このことを実現するための第一歩として、劣化しないメモリ媒体が必要です。少なくとも 100 年以上の保管を保証できるメモリ媒体が望まれます。現在、このことを実現しているのは、良質の紙媒体を除くと、アーカイブ仕様のフィルムだけです。特に画像情報に対しては、見た目保存ができることから有利です。しかし、今後も増大し続ける種々のデジタルデータとその利便性をフィルムのみによって確保できるとは考えにくく、利便性の点で他のデジタルメディア、すなわち光ディスク、ハードディスク、半導体不揮発性メモリなどによって市場が席卷されているのが実情です。

■ アーカイブの現場での課題

● 信頼性の課題

すでに媒体の平均寿命 (MTTF) が短いことが問題であることは述べましたが、さらに現場を悩ます問題として媒体の寿命バラツキがあります。ハードディスクの故障率については、米カーネギーメロン大学と Google によって調査がされています。光ディスクの故障率については NIST から報告されています。いずれも製品ごとに大きなバラツキが有る事が指摘されています。管理するデータ量が大きくなればなるほどこのバラツキ、故障率の問題が重要になります。

● 保管コスト

保管コストを正確に見積もるのは難しい問題です。大規模なデータセンターの事例をみると、1TB の年間の保管コストは 1,500 ドルと言われています。その内訳は、媒体コストが 1/3、管理運営、システムコストなどが 2/3 であると言われています。小規模なアーカイブでは、媒体コストの割合はさらに小さくなります。媒体コストが占める割合が小さいと言う事実は、保管システムを考える上で重要な要素です。

● データ爆発

毎年生み出されるデジタルデータの量は、何時かは飽和するものと思われます。しかしながら、今のところは増加の一途を辿っています。当分この状況が続くと考えられます。

■ 期待される記憶媒体の要件

データ保管を必要とするユーザーの立場で考えると、期待される要件として以下の様なものが考えられます。

- (1) 環境、メンテナンスフリー
 - ⇒ 高い耐環境性；フォーマットとインターフェースの永続性の保障
- (2) 保管コスト全体を安くできること；媒体そのもののコストよりも、保管コストに注目
 - ⇒ マイグレーションフリー⇒ 超長期寿命を持つ媒体の実現⇒ 単純な物理インターフェースの採用
- (3) データ爆発への対応
 - ⇒ 生成されたオリジナルデータを保管した媒体が即アーカイブ媒体となること

■ 半導体技術への期待

● メモリセルデバイス

このデジタル社会における記憶の保管性の危機を救うものとして、半導体不揮発性メモリに大きな可能性があると考えられます。現在すでに市販されている主流のフローティング型半導体不揮発性メモリの記憶原理によると、他のメモリ媒体では絶対に実現できない高温（例えば 100℃以上）で、100 年以上記憶を保持できます。さらに、MONOS 型メモリを用いれば、150℃で 1000 年以上記憶を保持することも期待できます。記憶の主体となるデバイスの主要な部分は、シリコン、シリコン酸化膜、窒化シリコンと安定な物質で構成されています。既存の半導体不揮発性メモリデバイスは、記憶保持のデータも実績も豊富にあります。このことは、半導体チップで超長期保管メモリを作ることを考える上で大きなアドバンテージを与えてくれます。

● 超長期保管メモリチップ実現へ向けての課題

最終的な商品としての半導体メモリチップによる超長期保管メモリ媒体の実現には、配線と周辺回路を含むチップ総体としての信頼性をどのようにして確保するのかを考えなければなりません。耐腐食性の高い配線とパッド材料の検討、1000 年の信頼性を保証する評価手法の確立が必要です。また、妥当なビットコストをどの様に実現するのか、さらにシステム、データのフォーマット、それらの sustainability、dependability をどの様に確保していくのかなど多くの課題があります。しかしながら、これらの課題は、解決可能であると考え、解決できた際に得られる対価は非常に大きなものです。

(1) 配線材料、パッド材料

チップを構成した際、最も脆弱な要素です。配線材料としては、より安定な高融点金属、パッド材料としては、導電性酸化物などの検討が必要となります。

(2) 信頼性確保と妥当なコストの実現と意味理解の保証へ向けて

超長期保管メモリチップの信頼性を確保するために、信頼性保障をしなければならない要素を極力少なくする必要があります。そのために、メモリチップ回路の簡素化、プロセスの簡略化は必須です。このことは、妥当なチップコスト実現のためにも必須です。そのため、現在のシステムと外部メモリをつなぐインターフェースと超長期保管メモリとの間に新たなブリッジシステムが必要になります。仮にこれをアダプターと命名します。このアダプターと超長期保管メモリチップとの間の物理インターフェースが何時の時代でも必ず構成できるものであれば、アダプター自身は長期の信頼性保障の必要はありません。また、アダプター自身の物理形態は、個人使用の場合、現在のメモリカードの様なものから、データセンターであれば、巨大なラックのようなものまで、様々なものが考えられます。さらに、ファイルシステムのスケラビリティを確保するためにメモリ内の ID、メタデータ領域の定義などを検討しなければなりません。このことは、超長期保管システムのメタデータスキーマの検討へ繋がります。

■ 超長期保管メモリのインパクトについて

多くの課題はありますが、課題が解決し、2 チップを張り合わせ、1mm 厚、1cm 角の超長期保管メモリが実現できたとします。容量は 16GB 程度を想定します。このメモリ媒体の保管に対するインパクトを考えて見ます。2007 年に全世界で生成されたと思われるデジタル情報量は、281EB（エクサバイト）と言われています。このデータを 1mm 厚、1cm 角、16GB の超長期保管メモリに全て保管したとすると、50m プール 1 つ分の容積で保管できます。しかも、保管温度は、150℃程度まで大丈夫なので、空調などの電力コストが不要になります。

さらに、半導体チップビジネスへのインパクトについて考えてみます。ITRS（国際半導体技術ロードマップ）によれば先端の半導体技術は 32nm から 28nm へ、2020 年には 14nm まで微細化が進むと予想されています。同時に、シリコンウェハの大口径化も進み、現在 450mm への移行が始まろうとしています。今回の超長期保管メモリチップが実用となれば、その需要は、人類の知的生産活動に比例し延び続けます。全世界で生み出されるデジタルデータ量の 1%に何らかの保管が必要となると仮定すると、2014 年には、月産 5 万枚の 450mm 工場、10 棟分の需要が生まれます。さらに、アダプターという新たなセットビジネスも生まれます。

■ まとめ

大きな財政負担なしにデジタルデータを長期に保存し、意味理解を保障するシステムと記憶媒体への要求は確実に増大しています。現在、長期保存と意味理解の技術的な保証の無いまま、様々な情報の電子化、デジタル化が進行しています。既に幾つもの機関がこの問題について検討を進めていますが、抜本的な進展は有りません。その様な状況の中で、100℃以上の保管温度を許容でき、故障率の保障が可能な半導体技術には大きな可能性があります。この分野には、大きな意味と価値があり、半導体技術は大きな役割を果たす事ができると考えます。

アジェンダ

- はじめに： 危機意識と幾つかの消失事例
- 長期デジタル保管を必要とする人々と課題に取り組んでいる団体、組織
- デジタル保管の課題と保管戦略
- ストレージの寿命[信頼性]
- 保管コスト
- 期待される超長期保管システムとは
- 半導体千年メモリの可能性について
- まとめ

1

危機意識

スウェーデン国立公文書館の報告書： デジタルブラックホール

デジタルへの過剰な熱意は危険が伴う。デジタル変換プロセスを開始するにはかなりの初期投資が必要である。その投資は、デジタルのメンテナンス経費を継続的に保証する構造的な資金調達手段が無かったり、将来計画が廃止されたりした場合は無駄な投資になってしまう。

このような**長期計画の無いデジタル化プロジェクトは宇宙のブラックホール**のようなものである。

http://www.jiima.or.jp/pdf/Palm_Black_Hole.pdf

2

危機意識

2010年電子情報通信学会誌1月号

情報がどのくらい長く正確に保存できるのか、そしてそれをどのくらい長く利用できるのかが問われている。

中略

デジタルコンテンツのSustainabilityに懸念が出ており、「デジタルジレンマ」とも呼ばれる。現在のデジタル情報を蓄積するストレージの寿命は10年オーダー、したがって、情報にエラーが生じる前にコピーをするマイグレーションが行われる。しかし、指数関数的に増加しているデジタル情報を例えば10年に一度コピーをすると、そのコピー量はより急速に増大しいずれ破たんする。

中略

現代の情報がすべてではなくとも1,000年後に見ることができるのであろうか。また、技術進歩により蓄積システムが新方式に置き換わり、旧方式が利用できなくなることはよく経験することである。このようにデジタル情報の恒久的保存と活用が可能な技術が確立されなければ人類の歴史が失われてしまうこととなる。

巻頭言

電子情報通信技術の発展を 二次元から四次元へ

会長 青山友紀



これまで電子情報通信の分野における技術の発展は主として情報の量と質の二次元を視野に向上がなされてきた。例えば量の次元では、より多くの情報を処理できるコンピュータ、より高速で大容量のネットワーク、より多くの情報を蓄積できるストレージ、それらを実現するデバイスの研究開発がなされてきた。質の次元では音声や映像などのアナログ情報の忠実性の向上、ステレオ化、立体化、高解像度化、デジタル情報の正確性、正規性、信頼度の向上、情報の違法使用耐性向上などである。量と質の次元は独立しているわけではなく、質の向上が量の増大を生じさせる。また量の増大が質の劣化をもたらす可能性が生じる。したがって、量を機械、質を機械とする二次元平面の第1象限を右上方に技術の発展がなされてきたと考えられる。そしてデジタル技術がそれを加速してきたのである。

しかし、この二次元だけを視野にした技術の発展では根本的な問題があることが明らかになってきた。追加すべき次元の一つは時間軸である。情報がどのくらい長く正確に保存できるのか、そしてそれをどのくらい長く利用できるのかが問われている。デジタルには多くのメリットがあり、本、書籍、設計図、カルタ、写真、映画フィルムなどをデジタル化して保存する動きが図書館、博物館、病院、設計事務所、フィルムセンターなどで開始あるいは検討されている。そのデジタルコンテンツのSustainabilityに懸念が出ており、「デジタルジレンマ」とも呼ばれる。現在のデジタル情報を蓄積するストレージの寿命は10年オーダーといわれており、したがって、情報にエラーが生じる前にコピーをするマイグレーションが行われる。しかし、指数関数的に増加しているデジタル情報を例えば10年に一度コピーをすると、そのコピー量はより急速に増大しいずれ破たんすることは明らかである。マイグレーションの予算も確保できなくなる。マイグレーションができなくなればデジタル情報が失われていくことになる。我々はパピルスや紙によって、数千年前のエジプトの記録や1,000年前の源氏物語を読むことができる。現代の情報がすべてではなくとも1,000年後に見ることができるのであろうか。また、技術進歩により蓄積システムが新方式に置き換わり、旧方式が利用できなくなることはよく経験することである。このようにデジタル情報の恒久的保存と活用が可能な技術が確立されなければ人類の歴史が失われてしまうこととなる。

最後の軸はグリーンICTを実現するエネルギーの次元である。地球上のすべての人々が平等に現代技術に支えられた高い生活レベルを享受することが目に見える具体的な目標としてすべての国々の人々に実感され、それに向けて急速に進んでいることが明白となり、人類の最重要課題として突き付けられることになった。1997年の京都議定書以降、国際政治もこの問題を取り上げるなど一部の知識人が問題視するだけでなく、政治が動き出したことは事象の切迫感を物語っている。この課題はすべての理工学分野において対応する必要があるが、電子情報通信の領域においても、量・質・時間の三軸が進展しても、必要なエネルギーがそれに応じて増大することは許容されない状況となった。すなわち、ビット当りの処理、伝送、蓄積に要するエネルギー効率を高める技術が必要である。

以上の四次元空間をバランス良く拡大させる電子情報通信技術の発展が期待されており、電子情報通信学会においても、2010年代にこの四次元を総合的に考える研究を進展させるよう期待したい。

3

過去の消失事例

1975年に打ち上げられた**NASAのバイキング探査機**(火星探査機)の初期のデジタルデータは、磁気データテープ上に記録され、アーカイブされていた。しかし、1999年に南カリフォルニア大学のジョセフミラーが、古いバイキングのデータをチェックしようとし、テープは探し出せたものの、それを読む方法を見つけられなかった。**わずか25年前のデータであるにもかかわらず、そのフォーマットをNASAがすっかり忘れてしまっていた。**幸いなことに、3分の1ほどのデータを繋ぎ合わせることに成功し、いくつか有用な結果を引き出すことができた。これは、部分的な参照資料と記録が紙に印刷され、テープと共に保管されていたおかげである。この事件は、「データの絶滅」と呼ぶ危険に対する重要な警鐘となった。その後、系統だったデータ移行を行うことにより米国連邦政府内のデータ資産を保護することを企図して作られた、**Open Archival Information System (OAIS、開放型アーカイブ情報システム)**と呼ばれる**データ参照モデルの開発の契機となった。**

慶応大学DMC 訳 ザ・デジタル・ジレンマより

4

過去の消失事例

国立国会図書館におけるデジタル資料の長期保存に関する調査

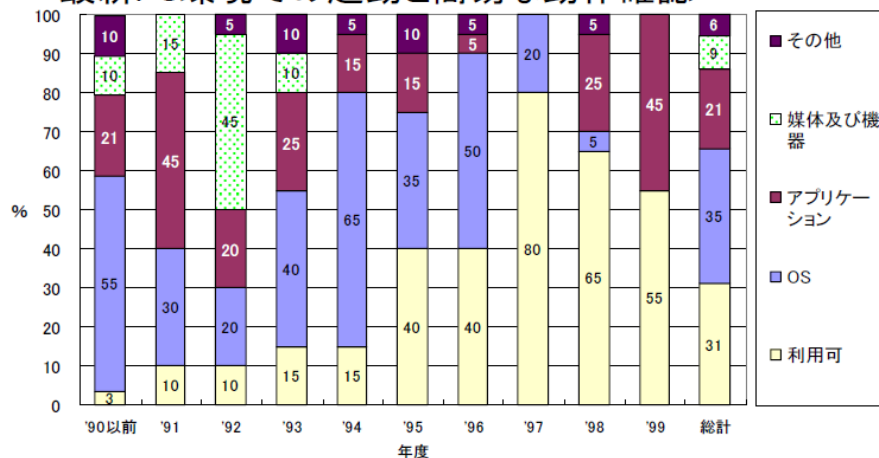
2007年2月 国立国会図書館関西館

デジタルアーカイブシンポジウムーデジタルアーカイブとその長期利用に関する研究会ー資料より

当館所蔵資料の再生可能性調査

PC用電子資料200点(1990~1999年度受入分)

最新PC環境での起動と簡易な動作確認



■ 全体の7割弱の資料の利用に問題

5

個人のデータ(デジカメ写真)

国立国会図書館-National Diet Library

電子情報の長期的な保存と利用

Q10: デジタルカメラの画像を孫の代まで保存するにはどうすればよいですか？

画像をデジタルデータとして保存し、孫の代まで長期にわたって見ることを可能とするには、そのファイル形式が将来にわたって使用できることが必要です。現在、一般的なデジタルカメラの画像の保存形式は、**いつ時代遅れのファイル形式になるかもしれません。**

中略

他に考えられる方法として、**デジタルカメラの画像を写真店等で銀塩写真用の印画紙にプリントして保管する方法があります。**銀塩写真とは、フィルムに記録する旧来からのカメラ写真のことをさします。銀塩写真用の印画紙の寿命は、適切に保管すれば100~150年ともいわれています。ただ、この方法ではどうしても**物理的な劣化は避けられません。**

まだまだ、個人レベルでデジタルカメラの画像を保存するための最良の方法は確立していないといえます

6

アジェンダ

- はじめに： 危機意識と幾つかの消失事例
- 長期デジタル保管を必要とする人々と課題に取り組んでいる団体、組織
- デジタル保管の課題と保管戦略
- ストレージの寿命[信頼性]
- 保管コスト
- 期待される超長期保管システムとは
- 半導体千年メモリの可能性について
- まとめ

7

アメリカ政府と公文書館

●アメリカ議会図書館 (Library of Congress)

デジタル情報インフラおよび保管の国家計画 (NDIIPP) を推進。議会は2000年に、NDIIPPの共同プロジェクトに1億ドルの予算。

○国立音声映像保存センター (NAVCC) :

アメリカ議会図書館の映画、放送及びレコード音声部門の全ての収蔵物を保管。年間8ペタバイト以上の情報を取り込む。著作権法により120年もしくはそれ以上の長さの寿命が要求される。

●国立公文書記録管理局 (National Archives and Records Administration: NARA)

アメリカ政府の全公式記録の保管を担当。何千というフォーマットで作られた何兆にもなるデータを「永久に」保存。このため、長期的な計画の立案のため Electronic Records Archive (ERA、電子記録アーカイブ) 計画を立ち上げる。

医療

- 1996年に米国で施行された健康保険の可搬性と責任に関する法律 (HIPAA) が契機となり、医療用データのアーカイブが生まれる。米国で最も規模の大きな病院は、2007年当時1ペタバイトのデジタルデータを保存。アーカイブは、1週間に3テラバイトの割合で増え続けており、2008年にはデータ量が2倍に。データの増加傾向は今後も続く。

8

科学技術計算

●SAN DIEGO SUPERCOMPUTER CENTER (SDSC、サンディエゴスーパーコンピュータセンター)：

スーパーコンピュータを、国家的研究を行うグループ向けに運用。5ペタバイト以上の容量を持つハイブリッドストレージシステムを保有。SDSCのストレージ容量は、14ヶ月ごとに2倍に増えており、2008年の終わりには10ペタバイトに。SDSCの移行間隔は5年。デジタルアーカイブの紛失を引き起こす主要な原因は、人間のミスとディスクハードウェアの故障。

もう一つの問題、情報爆発

米EMCの予測(2008年):2007年における世界のデジタル情報量は合計281E(エクサ)バイト(2810億Gバイト)に。今後情報量は年平均約1.6倍で増加し、2011年には1.8Z(ゼッタ)バイト(1800Eバイト)に拡大する。

9

世界の活動団体

NDIIPP

National Digital Information Infrastructure and Preservation Program

2000年にLibrary of Congressによって設立、米国を中心に184のパートナーが参加。

<http://www.digitalpreservation.gov/>

<http://www.loc.gov/library/digitalpreservation.html>

2000年には総額1億ドルの予算措置が行われていましたが、2007年年度で4,700万ドルが削減。



世界の活動団体

ERCIM

the European Research Consortium for Informatics and Mathematics



ERCIM - the European Research Consortium for Informatics and Mathematics - aims to foster collaborative work within the European research community and to increase cooperation with European industry. Leading research institutes from nineteen European countries are members of ERCIM.
<http://www.ercim.eu/>

iPRES

International Conference on Preservation of Digital Objects

2004年に中国-ヨーロッパのDigital Preservationに関するワークショップとして開催。その後、1年ごとに中国、ヨーロッパ、アメリカで開催。2011年に日本(筑波)で始めて開催。現在は、デジタル情報の保存に関する世界最大規模の国際会議。

11

日本の活動団体

国立国会図書館:

多くの調査活動。

- 「電子情報保存に係る調査研究報告書」平成15年3月
 - 「電子情報の長期的保存とアクセス手段の確保のための調査報告書」H16年3月
 - 「電子情報の長期的保存とアクセス手段の確保のための調査報告書」H17年3月
 - 「パッケージ系電子出版物の長期的な再生可能性について」図書館調査研究リポート No.6(H18年3月)
 - 「電子情報の長期利用保証に関する平成18年度調査報告書」H19年3月
- http://www.ndl.go.jp/jp/aboutus/preservation_02.html#

国立公文書館:

電子記録について調査活動。

- アーカイブズの観点から見る電子記録管理ガイド(1997年ICA報告書8)日本語版
 - 電子記録:アーキビストのためのワークブック(2005年ICA報告書16)日本語版
 - 電子媒体による公文書等の適切な移管・管理・利用に向けてー調査研究報告書ー (H18年3月)(国立公文書館)
 - 電子公文書等の作成時又は作成前からの評価選別」に係る調査研究報告書(H21年3月)
 - 電子公文書等の管理・移管・保存・利用システムに関する調査 報告書(H21年3月)(内閣府)
- http://www.archives.go.jp/law/report.html#Repo_02

12

日本の活動団体

社団法人日本画像情報マネジメント協会（略称：JIIMA）：

1962(昭37) 日本マイクロ写真協会が社団法人として認可

1995(平 7) 協会名称社団法人日本画像情報マネジメント協会（略称：JIIMA）に改称。 <http://www.jiima.or.jp/>

情報保存研究会(JHK)：

2000年に設立。元藤沢市公文書館長 高野修氏を会長に、情報保存機関の勉強をすることを目的として会員セミナーを実施。企業紹介を中心に企業ダイレクトリを発行。 <http://www.e-jhk.com/html/index.html>

財団法人 デジタルコンテンツ協会（Digital Content Association of Japan／DCAJ）：

1991年4月1日 設立

情報化社会をリードする良質なデジタルコンテンツの制作、流通、利活用を推進することにより、これに係わる産業の健全な発展を促すとともに、文化の向上と快適かつ心豊かな国民生活の実現及び国際貢献に資すること。

関連財団からの委託を受けて調査報告事業を展開。

<http://www.dcaj.org/report/index.html>

13

アジェンダ

- はじめに： 危機意識と幾つかの消失事例
- 長期デジタル保管を必要とする人々と課題に取り組んでいる団体、組織
- デジタル保管の課題と保管戦略
- ストレージの寿命[信頼性]
- 保管コスト
- 期待される超長期保管システムとは
- 半導体千年メモリの可能性について
- まとめ

14

デジタルデータ特有の課題

メタデータ

メタデータというのはデータについてのデータ。長期保存のためには、保存用のメタデータや管理用のメタデータといったものも必要になる。

コンテンツ・データ・オブジェクトがそもそもの保存対象のデジタル情報で、それ以外は全てメタデータ。

デジタル情報では、特に重要。ビット列が絵の情報なのか、テキストなのか、数値データなのか、フォーマットは、何時、誰が、目的は、履歴はなどなど。

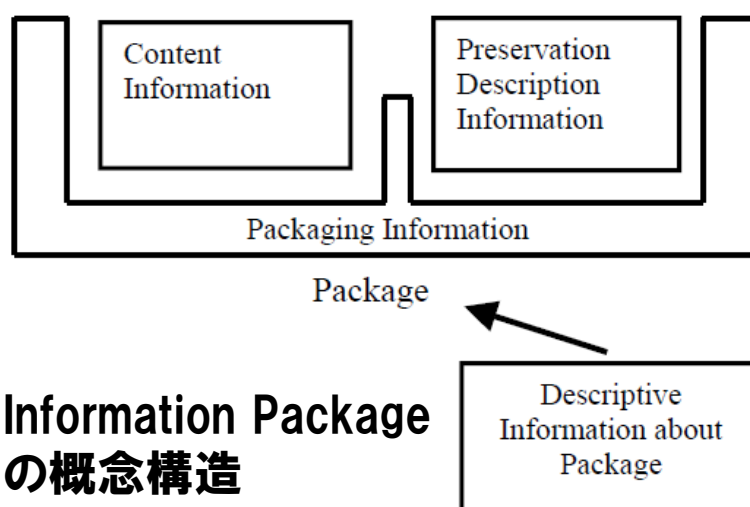
メタデータの記述要素や構造を定義するものを**メタデータスキーマ**という。デジタル情報の長期保存に使えるようなメタデータスキーマが様々な機関で検討されている。

15

デジタルデータ特有の課題

Open Archival Information System (OAIS) の参照モデル

OAIS参照モデルでは、情報オブジェクトを保存するために**情報パッケージ**という概念を導入。情報パッケージには、内容情報(Content Information)、保存記述情報(Preservation Description Information, PDI)、およびパッケージ



Information Package
の概念構造

化情報(Packaging Information)からなる。

また、情報パッケージに関する記述(Descriptive Information about Package)も必要とされる。この4 要素のうち内容情報以外はメタデータと言える。

16

マイグレーションとエミュレーション

デジタルデータを保管するための手順とデータシステムイメージはOAISとして標準化された。しかし、実体としてのデジタルデータには**システムのハード、ソフトの寿命**という問題が残る。

寿命	ハードウェア	ソフトウェア
3～5年	ホスト・コンピュータ	・アプリケーションソフトウェア ・オペレーティングシステム ・デバイスドライバ
5～10年	物理インターフェース	・インターフェース用ファームウェア
3～5年	メディア・ドライブ	・ドライブコントロール用ファームウェア
5～10年	メディア	・ファイルシステム ・データファイルフォーマット ・物理レコードフォーマット

17

マイグレーションとエミュレーション

Research Libraries Group (RLG; デジタル情報の保存に関して活動している団体) の報告書 (1996年5月) より

●**リフレッシュ**: 単に情報を別の新しいメディアにコピー。

●**エミュレーション**: オリジナルのアプリケーションやOSの振舞いを真似るエミュレータを用意する。オリジナルの媒体が生きていることが条件。

●**マイグレーション(移行)** (migration): 包括的な手段。マイグレーションはデジタル資料のあるハード／ソフト環境から別の環境への、あるいはある世代のコンピュータ技術から次世代の技術への定期的な移動。

しかし、これは時間も金もかかる複雑な作業。

ハリウッドの問題意識

現在、フィルムアーカイブを置き換え得る技術は存在しないとの認識。

現在100年以上にわたって保安が保障されると認識されている手段はフィルム保管のみ

フィルムシステムの利点とは：

- 世界的な標準方式
- 品質の劣化がなく長期的なアクセス（最低でも100年間）が保証
- 将来の（そして未知の）配給の機会やニーズを満たすため、複製したマスターが作れる
- オリジナル・カメラ・ネガや制作段階で録画、録音された品質と同等か、それ以上の品質
- テクノロジープラットフォームの変化と無関係
- 運用互換性がある
- 保管コストの増大が少なく、財政変動の影響を受けにくい
（棚に置いておくだけで保管ができる）

慶応大学DMC 訳 ザ・デジタル・シネマより

19

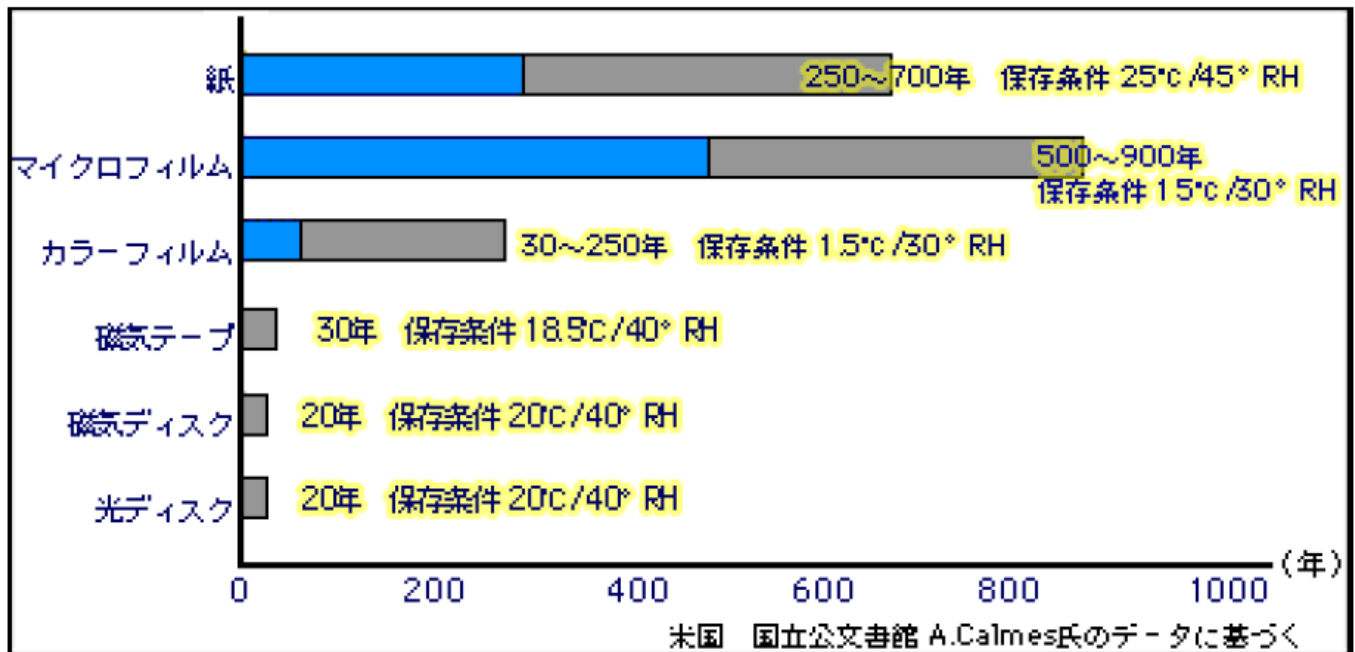
アジェンダ

- はじめに： 危機意識と幾つかの消失事例
- 長期デジタル保管を必要とする人々と課題に取り組んでいる団体、組織
- デジタル保管の課題と保管戦略
- ストレージの寿命[信頼性]
- 保管コスト
- 期待される超長期保管システムとは
- 半導体千年メモリの可能性について
- まとめ

20

ストレージ寿命[信頼性]

国立国会図書館 H15電子情報保存に係る調査研究報告書より



出典：株式会社極東マイクロの Web サイト (<http://www.kyokuto-micro.co.jp/media.ht>)

21

ストレージ寿命[信頼性]

アーカイブ仕様フィルム 21°C、50%RH 期待寿命500年

パッケージラベルの製品情報

ボックス底面
8桁の意味

3 5 5 1 0 1 1 2

スプーリング機番
(1~3)
スリット番号
(1~85)
スプール番号
(1~29)
ジュリアンデット
(001~365)



ISO 18901
LE 500の保
存環境

乳剤番号

ワイドロール
番号

連番

カタログ番号

期待寿命
500年の媒体

製品番号

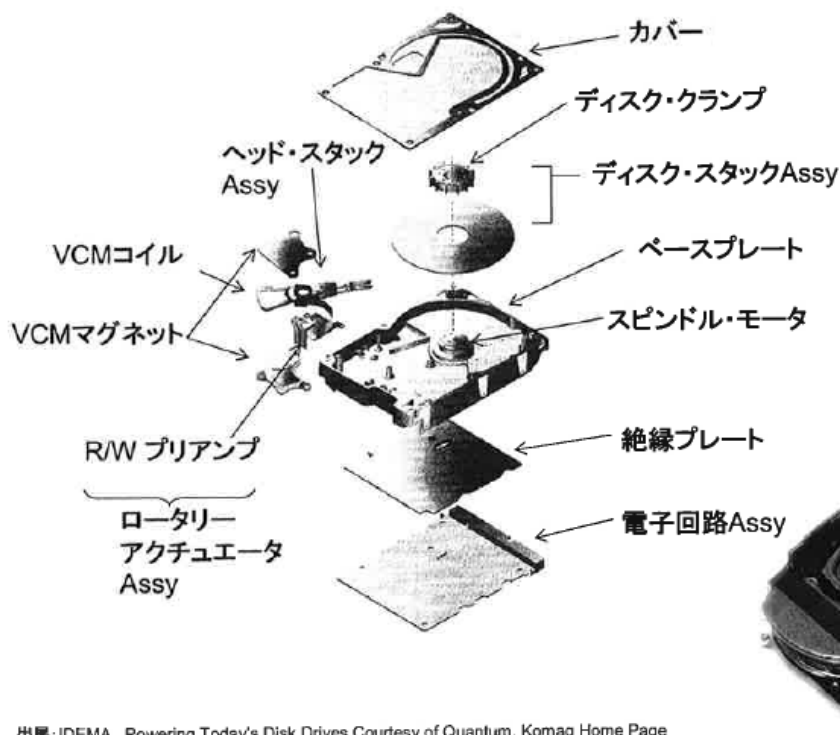
22

ストレージ寿命[信頼性]

ハードディスク

HDDの構成とメカニズム

ディスク装置の構造



多くの機構部分

どの部品が寿命を決めるのか？

産業用の製品で保証は5年程度と言われる

出展: IDEMA Powering Today's Disk Drives Courtesy of Quantum, Komag Home Page

23

Appears in the Proceedings of the 5th USENIX Conference on File and Storage Technologies (FAST'07), February 2007

Failure Trends in a Large Disk Drive Population

Eduardo Pinheiro, Wolf-Dietrich Weber and Luiz André Barroso

Google Inc.

1600 Amphitheatre Pkwy

Mountain View, CA 94043

{edpin, wolf, luiz}@google.com

10万台以上の民生用ハードディスクドライブを使用した調査を行った結果、温度やアクセス頻度に関係なく故障する。



5th USENIX Conference on File and Storage Technologies - Paper

Pp. 1-16 of the *Proceedings*

Disk failures in the real world: What does an MTTF of 1,000,000 hours mean to you?

Bianca Schroeder Garth A. Gibson
Computer Science Department
Carnegie Mellon University
{bianca, garth}@cs.cmu.edu

異なるベンダー製の10万個以上のSCSI, ファイバチャネル, Serial ATA(SATA)ドライブを調査。長期に渡って計算された年間故障率(annualized failure rate: AFR)のデータを紹介。故障率を表す従来の「バスタブ曲線」モデルに乗らない。ハードドライブの交換率は、初期の通電テスト期間後でも低いレベルで安定しない。

24

ストレージ寿命[信頼性]

Optical Storage の信頼性について

Volume 109, Number 5, September-October 2004
Journal of Research of the National Institute of Standards and Technology

[J. Res. Natl. Inst. Stand. Technol. 109, 517-524 (2004)]

Stability Comparison of Recordable Optical Discs—A Study of Error Rates in Harsh Conditions

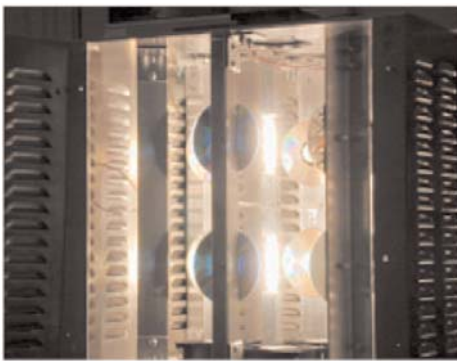


Fig. 2. Light Chamber.

It is demonstrated here that CD-R and DVD-R media can be very stable (sample S4 for CD-R and sample D2 for DVD-R). Results suggest that these media types will ensure data is available **for several tens of years** and therefore may be suitable for archival uses. Unfortunately, it is very difficult for customers to identify these more stable media.

25

ストレージ寿命[信頼性]

Optical Storage の信頼性について 業界の取り組み

●(財)デジタルコンテンツ協会：長期保存のための光ディスク媒体の開発に関するフィジビリティスタデ(平成18年3月)
DVDの評価データの拡充を実施し寿命推定法をより確実なものにするための基礎実験を行い、その結果を標準試験規格として提案し、光ディスク媒体の長期保存化を目指す。

●社団法人日本記録メディア工業会：「光ディスクの現状と将来」(2009年、パナソニック、長谷川氏講演から)
記録型光ディスクとして年間200億枚弱も消費される状況のなかで、Blu-ray Disc(BD)も成長の兆しが出始めており、2009年は6000万枚程度の市場を形成すると思われます。
最近では、光ディスクに関して、アイリングモデルを採用した加速試験による寿命推定法のISO化が完了しました。BDは、アレニウス法により100年以上の寿命が推定されています。

26

ストレージ寿命[信頼性]

幾つかの事実

- Linear Tape-Open (LTO) : **磁気テープの平均寿命は30年**。しかし、主要テープベンダーも、またNARAも、磁気テープに記録したデジタル資産は、**5年から10年おきでデータを移行**するように推奨。
- ハードディスク: グーグルとカーネギーメロン大学の最近の研究は、メーカーが主張しているほどの信頼性はなく、一般に認められている「**バスタブカーブ**」の故障特性にも従っていない。**多数のドライブがメーカーが定めた「平均故障時間」(MTBF) よりはるかに短い時間で故障**し、一般に故障の予兆指標だと考えられている温度と故障率との間には、低い相関性しかないことを示している。大規模なデジタル画像アーカイブの管理者は、**製品の品質むら**を指摘。
- ストレージテクノロジーのサプライヤーとエンドユーザの両者とも、デジタルストレージシステムの全てについて、**発表されている平均寿命の数字を大幅に割り引いて見ている**。

慶応大学DMC 訳 ザ・デジタル・ジレンマより

27

ストレージ寿命[信頼性]

半導体の長期保管ストレージ

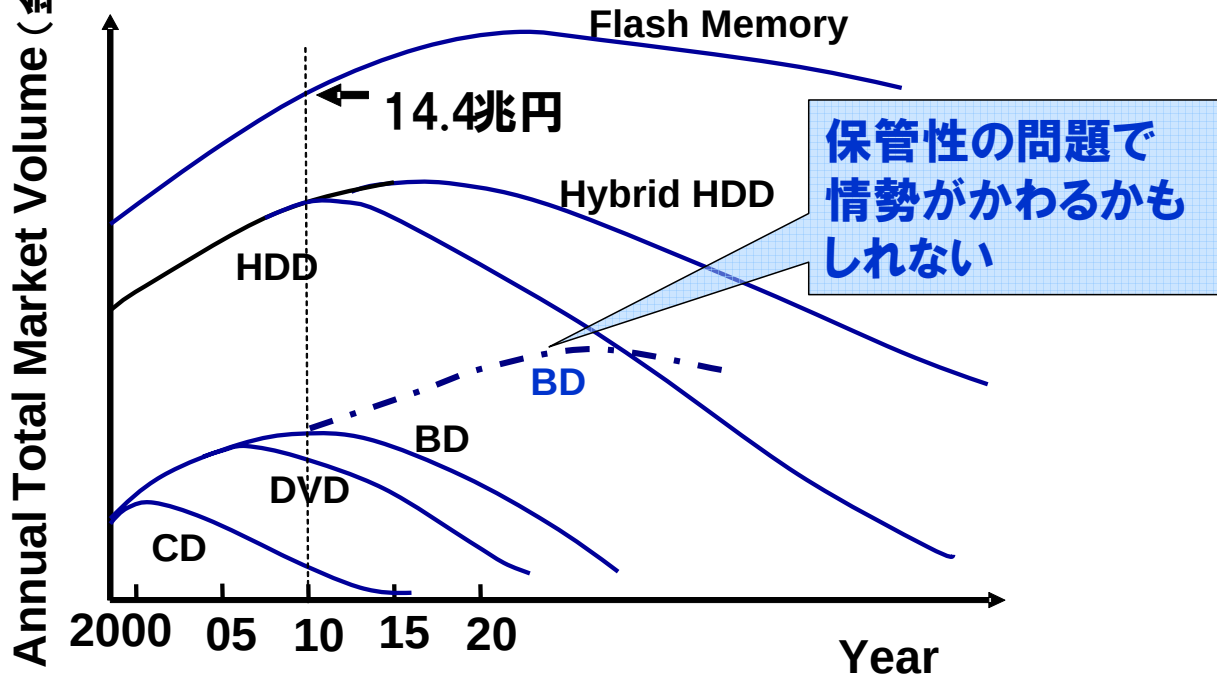
サンディスク新製品、ライトワンスメモリーカードSD™WORM

消去や改ざんが不可能な追記型メモリーカード8月発売開始
2008年7月15日、サンディスク・コーポレーション(NASDAQ: SNDK) はデータファイルの改ざんや消去に対するセキュリティが重視される業種向けのデジタルメモリー カードSD™WORM「WriteOnce Read Many」カードを発表。
静電気や湿気から守られた適切な室温の環境においては100年の寿命を持つ。

しかし、根拠が不明。コストが高く大規模アーカイブ、個人タ対応の商品ではない。

ストレージ寿命[信頼性]

主力デジタルストレージデバイスの市場予測
市場は半導体へと予測**[保管性は考慮外]**
しかしコストの点で今のところアーカイブ用としては考慮外



29

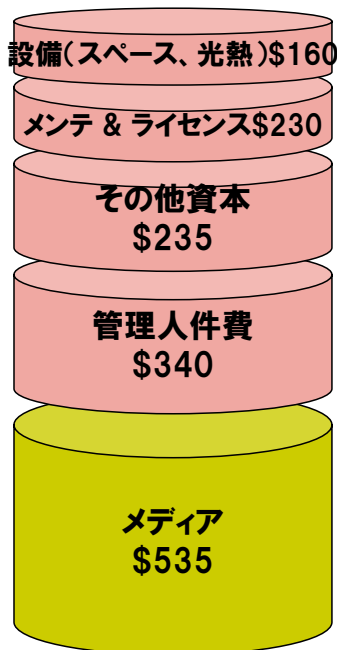
アジェンダ

- はじめに： 危機意識と幾つかの消失事例
- 長期デジタル保管を必要とする人々と課題に取り組んでいる団体、組織
- デジタル保管の課題と保管戦略
- ストレージの寿命[信頼性]
- 保管コスト
- 期待される超長期保管システムとは
- 半導体千年メモリの可能性について
- まとめ

30

保管コスト 幾つかの事例

SAN DIEGO SUPERCOMPUTER CENTER (SDSC、サンディエゴスーパーコンピュータセンター)



ハードディスクの場合:

科学技術計算ディスクストレージ費用

- \$1,500/(年・Tbyte) (2006年)
- **維持管理コストはハードディスクメディア本体の約3倍**

小規模なアーカイブでは、維持管理コストの比率は、さらに増大

慶応大学DMC 沢・デジタル・ジレンマより

31

保管コスト

保管コストを正しく見積もるのは難しい

デジタルストレージシステムを構築する際の総所有コストの内訳。

- ハードウェア、オペレーティングシステム、およびアプリケーションソフトウェアの初期費用
- ソフトウェアおよびハードウェアのメンテナンス契約
- **ハードウェア、オペレーティングシステム、およびソフトウェアアプリケーションの更新費用**
- 分散システムの外部ネットワークアクセス費用
- **メディアの初期費用および更新費用**
- 継続的トレーニングを含む人的費用
- 電力及び冷房費用
- 設備及び不動産の費用、税金、保険
- デジタル資産の収蔵量の増大にともなう費用の増加
- データの取り込みおよびアクセス費用

3～5年毎の
マイグレーション時に発生

慶応大学DMC 沢・デジタル・ジレンマより

32

アジェンダ

- はじめに： 危機意識と幾つかの消失事例
- 長期デジタル保管を必要とする人々と課題に取り組んでいる団体、組織
- デジタル保管の課題と保管戦略
- ストレージの寿命[信頼性]
- 保管コスト
- 期待される超長期保管システムとは
- 半導体千年メモリの可能性について
- まとめ

33

期待される超長期保管システムとは

コスト分析の結果と今後の規模の増大から期待されるメモリ媒体の要件

- 長期間放置されたとしても、財政的な困難が生じたとしても、影響を受けないことが理想。棚に置いておくだけ“保管”可能なこと。**環境、メンテナンスフリー。**
- 媒体自身のコストが安いのみならず、コスト構造を変える媒体。すなわち**読み出し装置、システムコストが安くできる媒体。**
 - ⇒ **物理インターフェースが単純なもの**
【電気的な接点が最も単純】
- 拡張性(スケーラビリティ)の保証

34

コスト分析の結果と今後の規模の増大から期待されるメモリ媒体の要件 [つづき]

さらに考慮すべき事項

- ・ データ爆発:そもそもマイグレーションは将来も可能か？
データの転送速度の問題。
- ・ 今後、デジタルボーンのコンテンツが主流に

●コストが妥当で、原本=アーカイブ=ライブラリが実現できること

アーカイブ=ライブラリが実現できれば、アクセス手段の確保はエミュレーションが主流に。

⇒ 低コスト化。

35

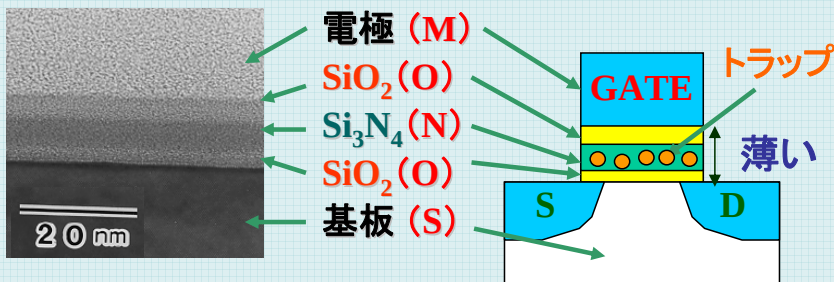
アジェンダ

- はじめに: 危機意識と幾つかの消失事例
- 長期デジタル保管を必要とする人々と課題に取り組んでいる団体、組織
- デジタル保管の課題と保管戦略
- ストレージの寿命[信頼性]
- 保管コスト
- 期待される超長期保管システムとは
- 半導体千年メモリの可能性について
- まとめ

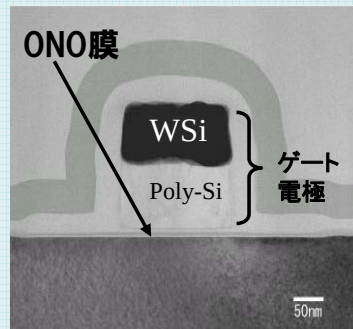
36

MONOS型不揮発性メモリ

MONOS構造



構造が簡単
製造が容易
高い歩留まり
高い信頼性
信頼性評価手法確立



MONOS型不揮発性メモリは、日立、Spansionで実用化されている。シチズン、ソニーにおいても使用されている。

MONOS型不揮発性メモリは、SiO₂薄膜で挟まれた窒化シリコン膜(ONO膜構造)に電荷をため、記憶を保持するメモリ。

ONO膜構造を最適化し、使用方法を規定する事によって、極めて高い記憶保持力と信頼性を得られる。

37

意味のあるものとするための提案 チップとしての信頼性について

- 配線に金属は用いない。
少なくとも腐食性の高い金属を避ける。
⇒ 究極はSiとSiの酸化物、窒化物のみとする
- 読み出しPadの工夫。
耐磨耗、耐腐食。 ⇒ 究極は非接触
- 信頼性の高いパッケージ
⇒ 究極はパッケージなし。チップが最終商品

38

その他の課題

1000年動作保証されたNVMメモリチップを作ることが可能になったとしてさらに必要な考慮事項:

■ 1000年動作保証するNVMメモリチップの仕様をどうするのか

- Write once
- 基本“紙と同じ機能”
- メモリとして最小限の回路: 書き込み、読み出し、保護回路

■ データ形式、読み出しの方式は、意味理解の保障は

- 永続性を保障できるデータフォーマットとは
- 物理的な読み出し方式は
- メタデータスキーマは、ハード的なメタデータは
- 永続性のあるメタデータ記述言語は、“ロゼッタストーン”は必要か

この問題の整理、解決のために“**アダプター**”という概念を提案。
超長期保管メモリチップとアダプターとで時間軸上の役割、機能の分担をさせる。

39

千年読み出しを維持できるシステム アダプターとメモリ媒体チップ

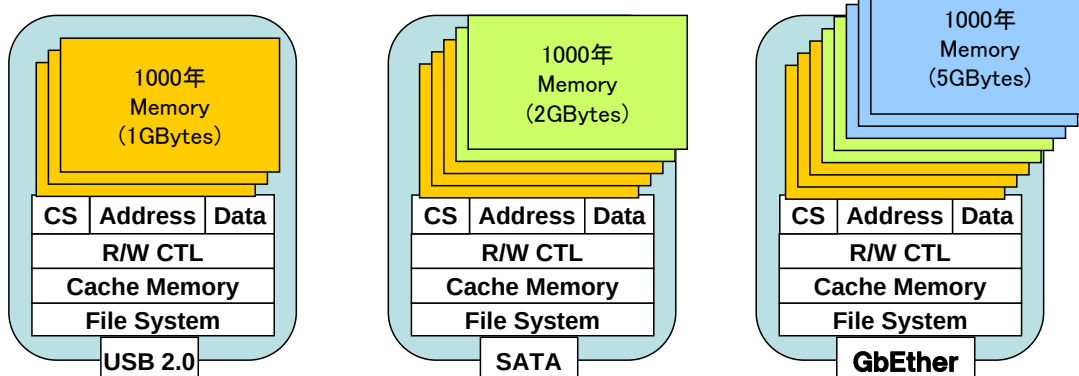
・ アダプター

- 超長期保管メモリチップのインターフェースがシンプルで上位互換を維持(標準化)
- 複数のメモリからなる単一ファイルシステムを構築できる
- 記録フォーマットの変換

・ 超長期保管メモリチップ

- I/Fはシンプルで上位互換を維持
- 複雑なコントローラなどはいれない

- ・ メタデータスキーマは
- ・ メタデータは何処に持たせるのか

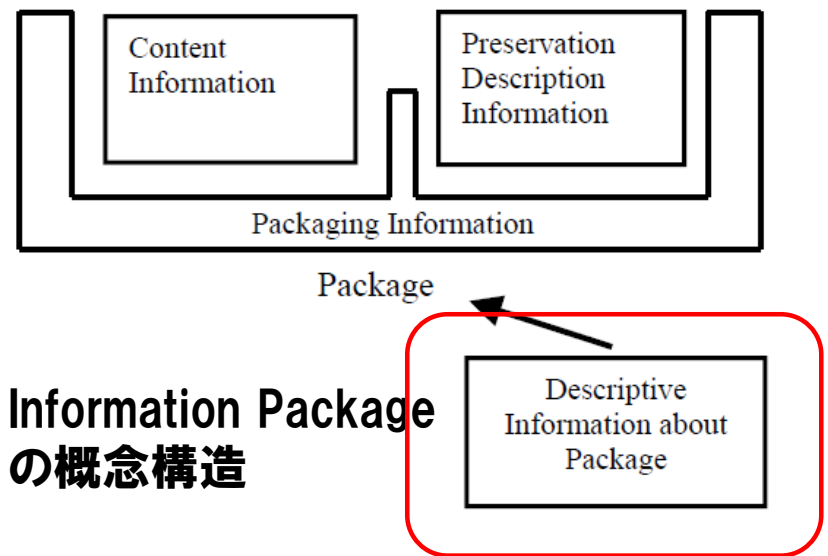


40

デジタルデータ特有の課題

Open Archival Information System (OAIS) の参照モデル

OAIS参照モデルでは、情報オブジェクトを保存するために**情報パッケージ**という概念を導入。情報パッケージには、内容情報(Content Information)、保存記述情報(Preservation Description Information, PDI)、およびパッケージ化情報(Packaging Information)からなる。また、情報パッケージに関する記述(Descriptive Information about Package)も必要とされる。この4 要素のうち内容情報以外はメタデータと言える。



41

もう一つの課題： コスト

●そもそも、チップの価格は？（ロジックチップの場合） 32nm、450mm、2万枚工場のコスト(STRJの試算)

- ・大口径化は1.5倍のコストアップ
- ・世代毎の工場建設費用は1.2倍のコストアップ
- ・加工費として、32nm450mmは72万円

300mm(65nm)ではおよそ576個

$330000 / 576 = 573\text{円}$ (NAND 1GB/チップ)

450mm(32nm)ではおよそ1296個

$720000 / 1296 = 555\text{円}$ (NAND 8GB/チップ)

●チップ収率の増大とプロセスコストの低減

基本的には、メモリセルアレーのみのチップに
コントローラ等、直接周辺以外のメモリ回路

⇒アダプターへ

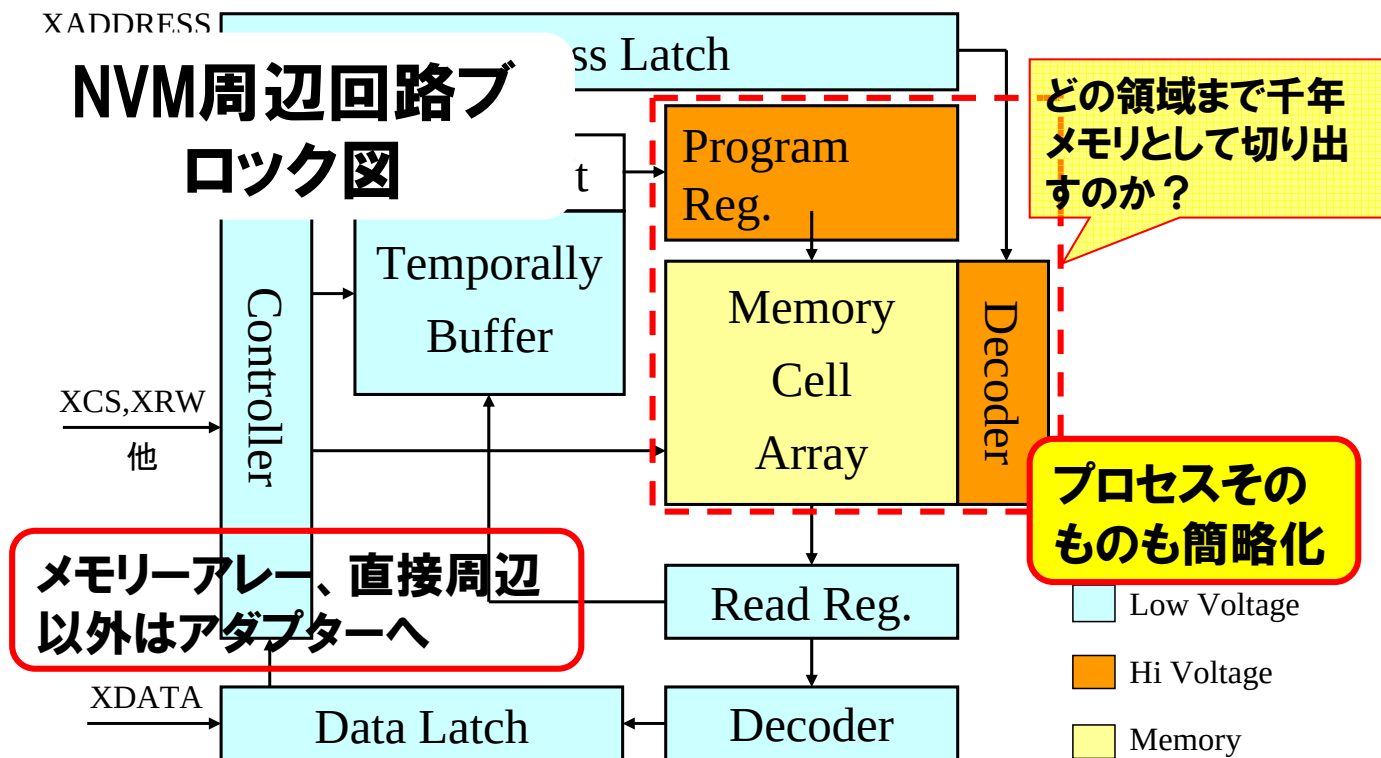
42

もう一つの課題：コスト

1cm²のダイコストは標準
ロジックチップで500円程
度(STRJ見積)

チップ収率の増大とプロセスコストの低減

⇒ 基本的には、メモリセルアレーのみのチップに



43

半導体産業と超長期保管メモリの 相性について(半導体産業へのインパクト)

STRJ(JEITA半導体技術ロードマップ委員会)
経済性検討小委員会より投資効率のよくなる
工場規模として以下の数字が報告されている

投資効率のよくなる工場規模(出展:STRJ)

150mm 14,000枚／月

200mm 20,000～24,000枚／月

300mm 32,000～48,000枚／月

450mm 50,000～75,000枚／月

44

450mmFabの課題 いったいなにをつくるのか

450mmウェーハになれば;

- 面積は159,043mm²で140mm²のチップは908チップ (歩留80%)
- 2万枚/月の工場で年間1.8億個を供給できる。これは年間のパソコン出荷台数1.75億台(2004年)とほぼ同数。
- インテルのMPU年間出荷個数は2.4億個
- 単一品種でこの工場を埋めることは不可能。複数品種を流す必要がある。

32nmデバイス

微細化、消費電力、バラツキの問題はあるものの以下のデバイスがある

- MPU
- 16GFALSH
- 8GDRAM、NAND
- ASSP (MediaProcessorなど)
- ASICは? マスク費用(開発費)の問題もあり色々なICを作れない

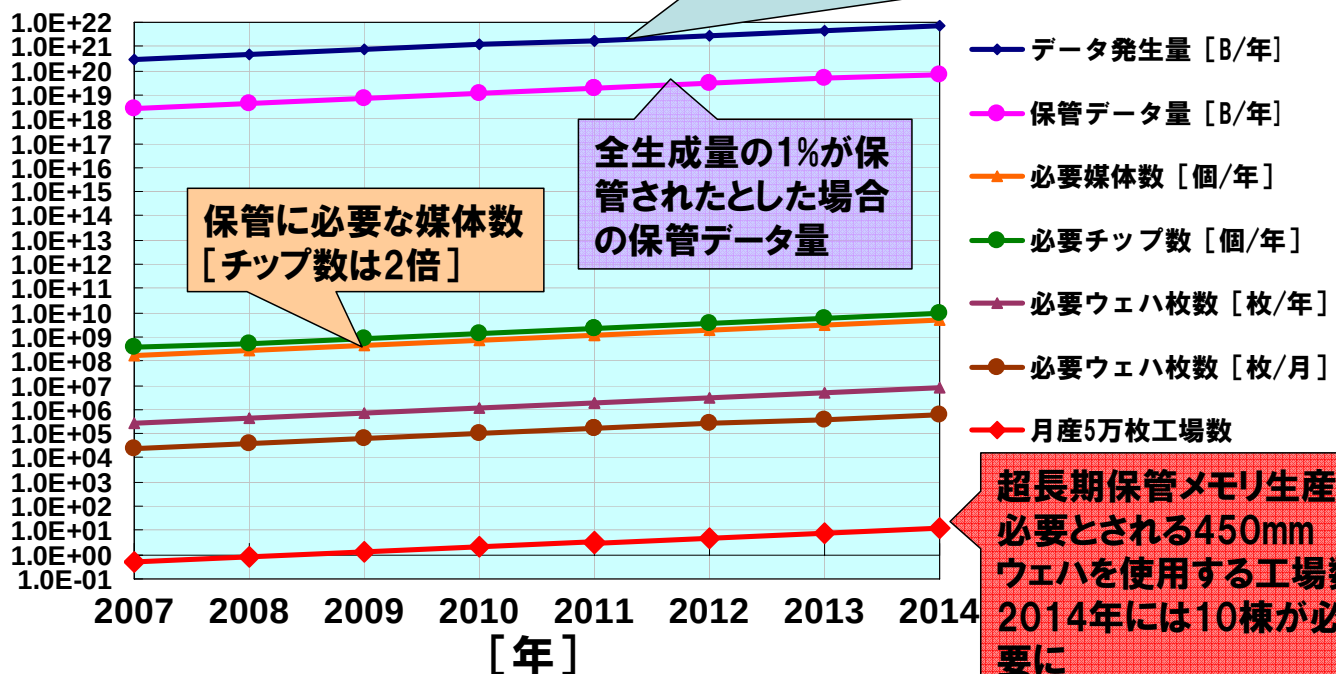
---\$24Mのマスク費用

45

商品としての超長期保管メモリが実現した場合、世界中で必要とされる450mmLSI生産工場数の予測

【前提】 工場規模 : 月産5万枚
媒体容量(2チップ構成) : 16GB

EMCの報告を元に予測した全世界でのデジタルデータ全生成量



46

アジェンダ

- はじめに： 危機意識と幾つかの消失事例
- 長期デジタル保管を必要とする人々と課題に取り組んでいる団体、組織
- デジタル保管の課題と保管戦略
- ストレージの寿命[信頼性]
- 保管コスト
- 期待される超長期保管システムとは
- 半導体千年メモリの可能性について
- まとめ

47

まとめ

- 100年以上、1000年近く、妥当なコストで意味理解を保障する記憶システムは必須
- エレクトロニクスはこの問題に対して大きな責任がある
- 半導体技術、特に不揮発性メモリ技術が解を提供できる
- 意味理解を保障するためにはハード、ソフトにまたがるメタデータとメタデータスキーマの検討が必要
- 標準化を達成し、商品としての超長期保管メモリ(千年メモリ)が実現できれば、半導体産業に大きな恩恵をもたらす

48