DEWS2006 ミニサーベイ／平成18年3月2日

ストレージシステムの高度化技術に関する研究動向

東京大学 生産技術研究所
産学官連携研究員 合田和生
kgoda@tkl.iis.u-tokyo.ac.jp

データ管理

- 情報の爆発
  - 24x10^{18}B (2003)
  - 指数関数的増大
- 高度情報化社会
  - 電子政府
  - 電子商取引

安心、快適、容易な
データ管理

ストレージ技術

More New Information Over Next 2 Years Than in All Previous History

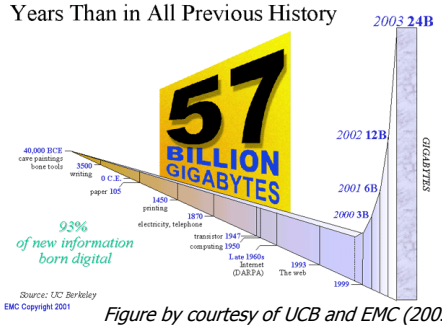
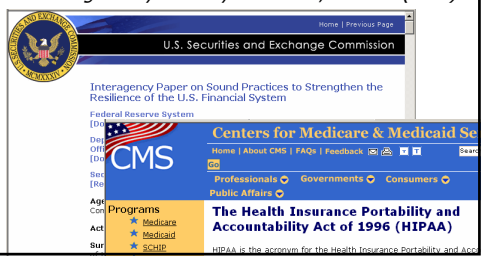


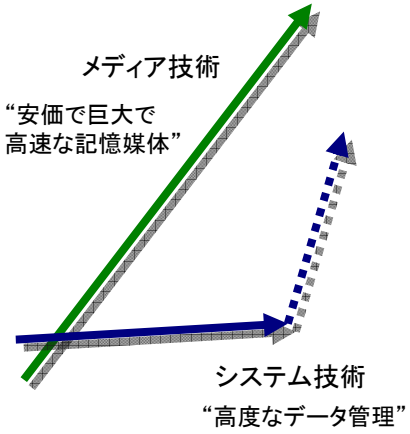
Figure by courtesy of UCB and EMC (2003)

Figures by courtesy of US-SEC, US-CMS (2004)



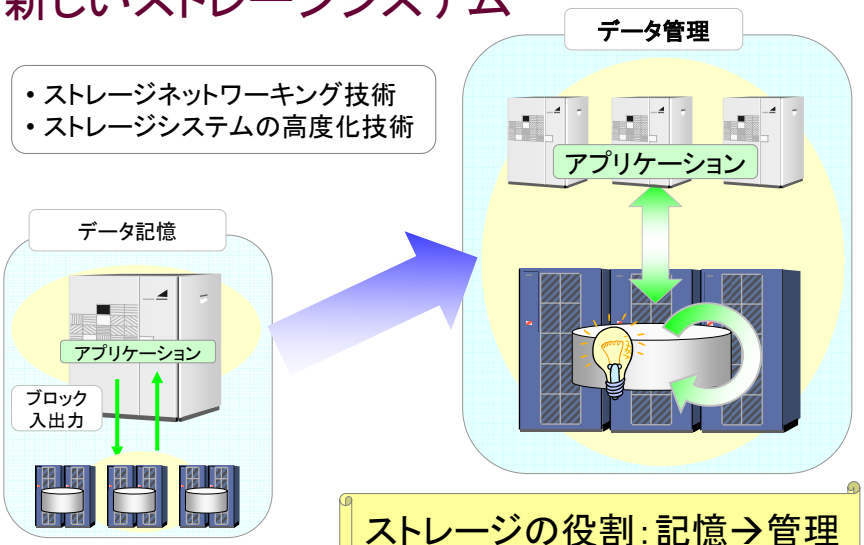
ストレージ技術の進展

- メディア技術
  - 過去約50年における飛躍的な進展
    - 10^8倍の高密度化
  - 更に加速（1990~）
    - 密度：+100% CGR
    - 価格：-30% CGR
- システム技術
  - cf. RAID, ファイルサーバ
  - 基本的なアーキテクチャは不変
    - ブロックレベル入出力命令（READ, WRITE）
  - → 技術革新の兆し



新しいストレージシステム

- ストレージネットワーキング技術
- ストレージシステムの高度化技術



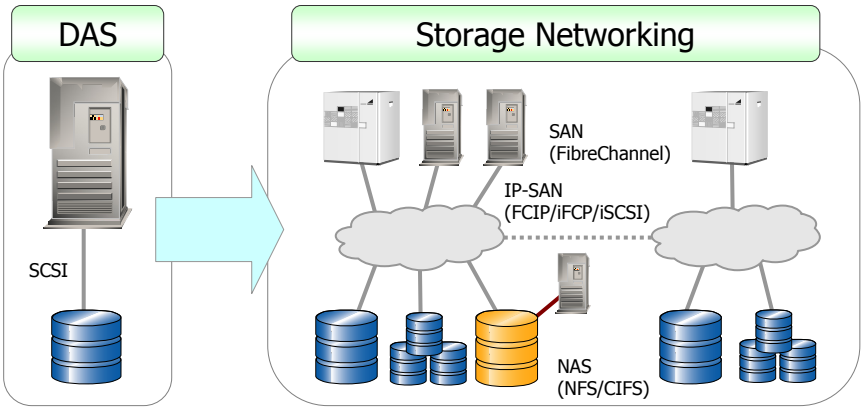
ストレージの役割：記憶→管理

アジェンダ

- ストレージネットワーキング
- 最近のストレージシステム高度化技術
- ストレージフュージョン技術

ストレージネットワーキング

ストレージネットワーキング



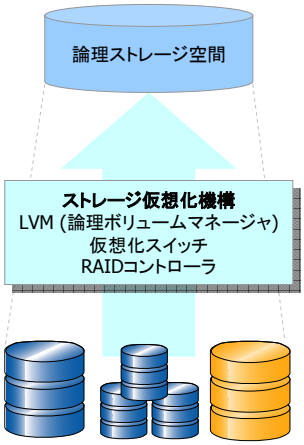
DAS: Direct Attached Storage
SAN: Storage Area Network
NAS: Network Attached Storage

ストレージネットワークの登場により、
ストレージの機能がシステムとして主体化

平成18年3月2日 DEWS2006 ミニサーベイ

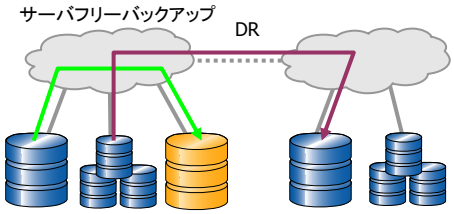
ストレージ仮想化と資源プロビジョニング

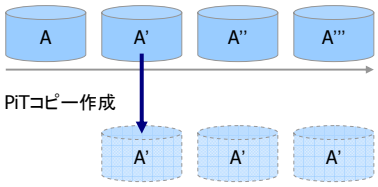
- ストレージ仮想化
  - ストレージ空間アドレスの論理・物理変換
  - 物理ストレージ装置の隠蔽
  - 論理ストレージ空間の構成
    - LU（論理ユニット）
    - LV（論理ボリューム）
- 資源プロビジョニング
  - 論理ストレージ空間への資源割り当ての調整
    - 静的、動的
    - 容量、性能（ストライプ度）
- ストレージアプリケーション基盤技術
  - 高度コピー技術
  - 高度データ保護技術
  - ストレージ資源管理技術



高度コピー技術

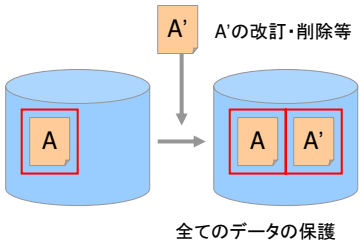
- LAN・サーバフリーバックアップ
  - LAN・サーバを経由せず、SAN内で完了するバックアップ
  - バックアップタスクの独立化
  - サードパーティコピー命令等を利用
- DR (Disaster Recovery)
  - 広域災害対策のための遠隔コピー
  - 同期転送方式、非同期転送方式
- PiTコピー
  - 対象論理空間の一時点イメージを維持する論理コピー、物理コピーを生成
  - システム試験や一貫性のあるバックアップ生成等に利用
  - ビットマップ方式、コピーオンライト方式

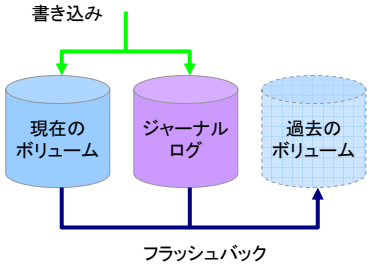




高度データ保護技術

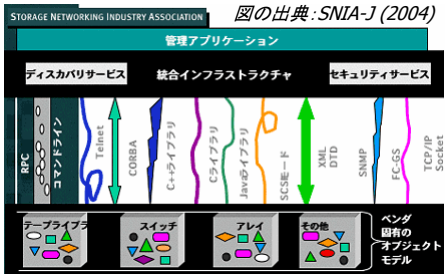
- WORM (Write Once Read Many)
  - 固定コンテンツ記憶
  - コピーオンライト方式利用
  - 情報保護規制への対応
    - 企業改革法（米）
    - e-文書法、個人情報保護法（日）
- CDP (Continuous Data Protection)
  - ジャーナルログを利用したフラッシュバック機能を実現
  - 非データベースアプリケーションに対するREDO/UNDO
  - システム試験や、ヒューマンエラーからの普及等に利用



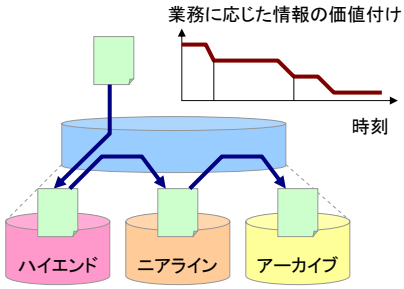


ストレージ資源管理技術

- SMI-S (Storage Management Initiative - Specification)
  - SNIAによるストレージ管理プロトコル
  - マルチベンダ、マルチレイヤ
- ILM (Information Lifecycle Management)
  - 業務に基づく情報の価値付けの時間的な変化に応じたストレージ資源の割り当て・管理
  - 情報を持つことのTCOを削減
  - cf. HSM, DLM



図の出典：SNIA-J (2004)



ストレージネットワーキング

- SAN、NAS等のストレージネットワーク
- ストレージ仮想化とプロビジョニング
- 高度なストレージアプリケーション

- → ストレージがシステムとして機能

最近のストレージシステム高度化技術

ストレージシステムの高度化技術

- データベースマシン（1970~80年代）
  - “Logic per track” (D. L. Slotnick, 1980)
  - 専用ハードウェアにより高性能処理を目指す
  - プロセッサ性能が不十分
  - 性能利得によるコスト正当化が不十分
- 近年
  - 情報爆発時代における巨大データ管理の社会的要請
  - ストレージ装置の計算機資源の向上
  - ストレージネットワーキングによる環境整備
- → 再び、ストレージシステムの高度化の流れ

最近のストレージシステム高度化技術

- アクティブストレージ技術
- コラボラティブストレージ技術
- オートノミックストレージ技術

アクティブストレージ技術

- ストレージ上のプロセッサを用いてデータインテンシブアプリケーションを実行
  - データになるべく近いところでコードを実行
  - 高い内部帯域を有効活用
- 高いディスクアクセス並列度
  - → 高いIO帯域
  - → 高いアプリケーション処理スループット
- Active Disks (UCSB, 1998)
  Active Storage (CMU, 1998)
  iDisks (UCB, 1998)
  - 基本的なアイデアはデータベースマシンに類似
  - プロセッサ技術の向上によって支持

アクティブストレージ技術
Active Disks, Active Storage

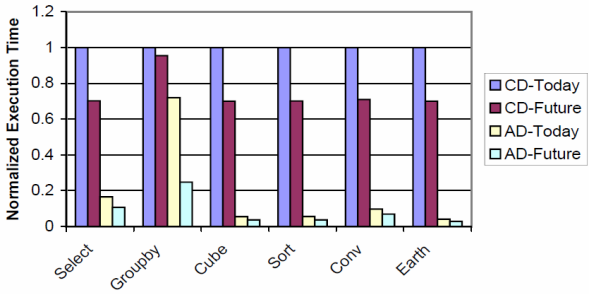


Figure by courtesy of A. Acharya et al. (1998)

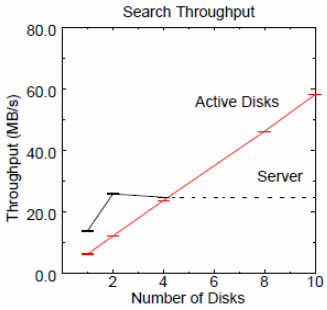


Figure by courtesy of E. Riedel et al. (1998)

サーバ実行と比較して、アクティブストレージは
飛躍的な性能向上を達成

アクティブストレージ技術
Freeblock scheduling

- アクティブストレージ上におけるIOスケジューリング技術
  - ディスクドライブの物理構造を意識した緻密スケジューリング
  - ディスクドライブの性能を最大限に活用
- Freeblock scheduling
  - ディスクドライブの回転遅延、シーク時間を活用
  - データインテンシブなアプリケーションを実行

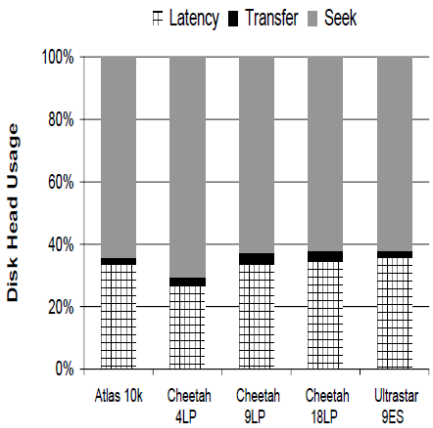


Figure by courtesy of C. Lumb et al.. (2000)

アクティブストレージ技術
Freeblock scheduling

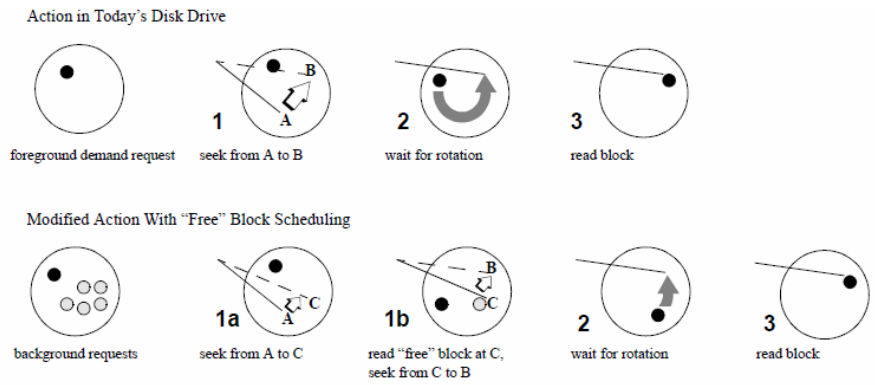
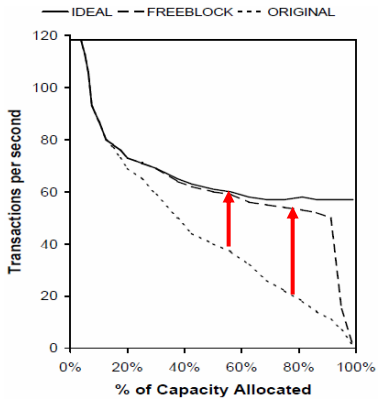
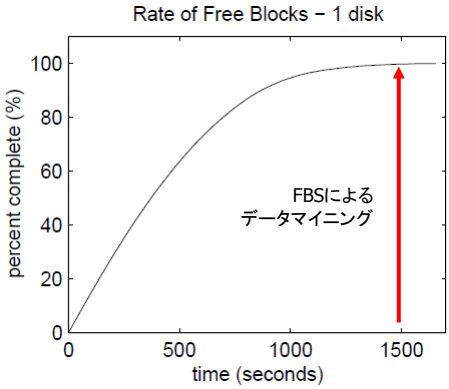


Figure by courtesy of E. Riedel et al. (2000)

平成18年3月2日 DEWS2006 ミニサーベイ

アクティブストレージ技術
Freeblock scheduling





PostMark実行下でのLFSセグメンテーションクリーニング

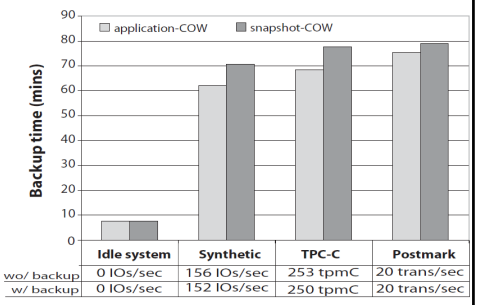
OLTP実行下でのデータマイニング

Figure by courtesy of C. Lumb et al. (2000)

Figure by courtesy of E. Riedel et al. (2000)

アクティブストレージ技術 Freeblock scheduling

- 一般フレームワーク化
  - フォアグランドアプリケーション
    → 同期IO
  - バックグラウンドアプリケーション
    → 非同期IO（FBS対象）
- データ管理ジョブに利用
  - PiTコピー作成
  - バッファクリーニング
  - レイアウト再編成

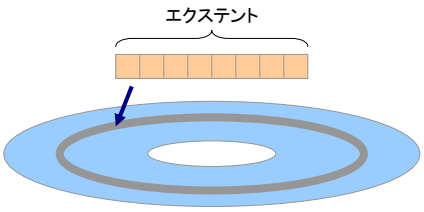


Freeblock scheduling APIを用いた
スナップショット作成

Figure by courtesy of E. Thereska et al. (2003)

アクティブストレージ技術 Traxtents

- アクティブストレージ上におけるIOスケジューリング技術
  - ディスクドライブの物理構造を意識したレイアウト方式
  - ディスクドライブの性能を最大限に活用
- Traxtents (Track-aligned extents)
  - ファイルシステムのエクステント（隣接ブロック群）を、トラックを跨がないように配置



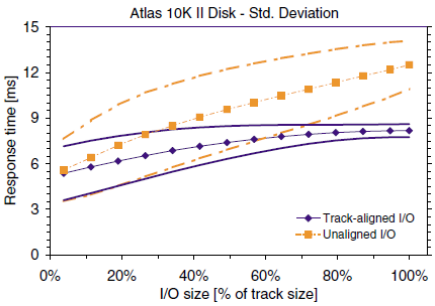
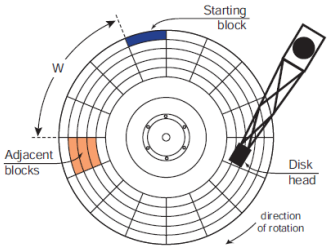
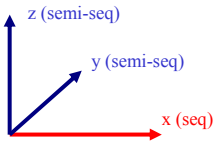


Figure by courtesy of J. Schindler et al. (2002)

アクティブストレージ技術 Adjacent blocks

- Adjacent blocks
  - 回転遅延・シーク時間から一定時間で到達可能なブロック候補群
  - 余計な回転遅延を削減
- 多次元データの格納に利用
  - ディスクドライブの物理構造を意識した多次元データ→ブロックのマップ（MultiMap）
  - 各次元方向の走査に「程々の」スループットを期待





4D earthquake dataset.

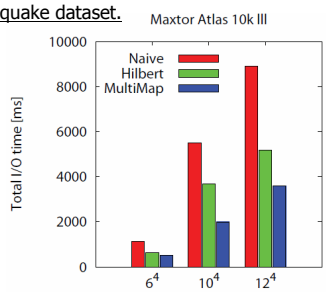


Figure by courtesy of S. Scholosser et al. (2003)

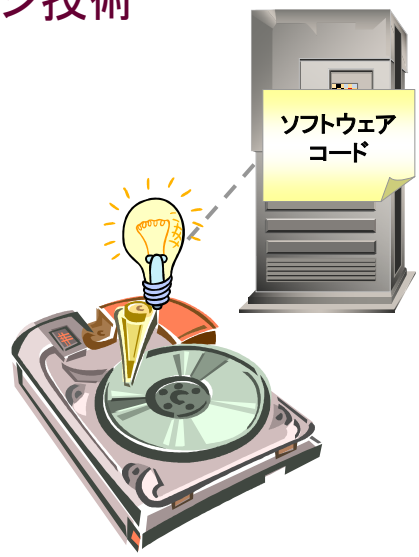
Figure by courtesy of S. Scholosser et al. (2003)

アクティブストレージ技術 まとめ

- ストレージ上のプロセッサを用いてデータインテンシブアプリケーションを実行
  - 高いディスク並列を以って高いIO内部帯域を活用
  - 物理情報を利用したIOスケジューリングとデータ配置技術
  - → 高いアプリケーション処理スループットを目指す
- 課題
  - プロセッサ性能
  - インターフェースの標準化
  - 専用ハードウェアの開発コスト

コラボラティブストレージ技術

- アプリケーションはサーバ上で実行
- ストレージがサーバ上のアプリケーションを意識して動作
  - ヒント情報
    専用のインターフェースを通じて情報交換
  - アプリケーション予測
    ストレージ側でアプリケーション動作を高度に予測
- ストレージアーキテクチャの大幅な改編を回避

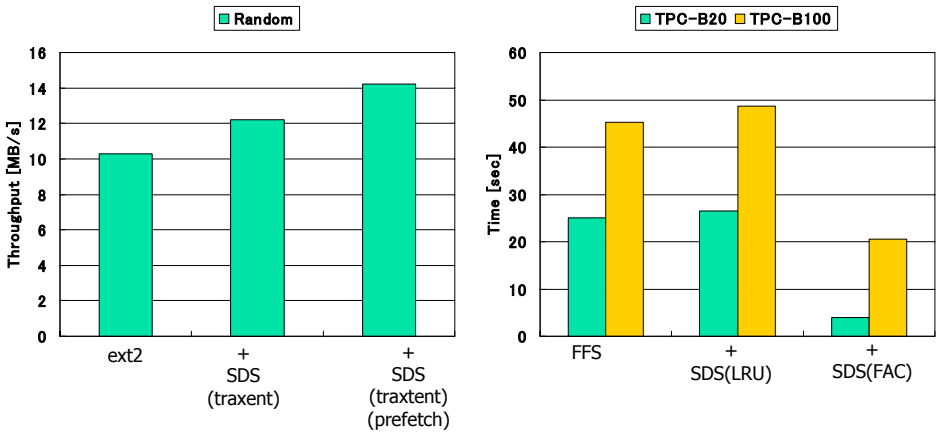


平成18年3月2日 DEWS2006 ミニサーベイ

コラボラティブストレージ技術 Semantically-smart disk

- 上位のファイルシステムを理解して動作する「賢い」ディスク
  - アプリケーションコード SCSIインターフェースは不変
  - ディスクのプロセッサ、メモリを活用
  - ファイルシステムを補助
- → 高い性能、高い可用性、高い保安性
- データベースにも拡張
- ファイルシステムの構造・操作を類推
  - FFS, ext2fsを仮定
  - ブロックの同定（データ、iノード、ビットマップ、ジャーナル、メタデータ）
    - 直接的ブロック分類：プローブプロセスを利用
    - 間接的ブロック分類：オンラインでのブロック監視
  - ブロック間関連の抽出
  - 高レベルファイル操作の類推
- ファイルシステムの操作を補助

コラボラティブストレージ技術
Semantically-smart disk

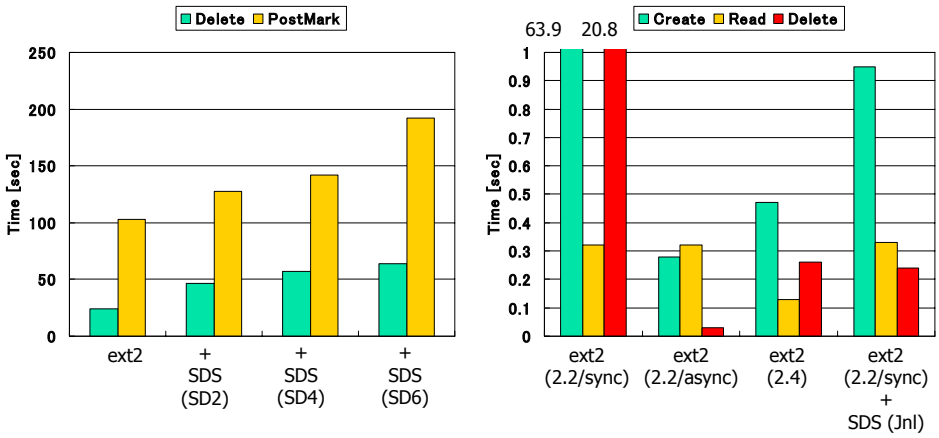


Track-Aligned Extents **Structural Caching**

Experimental results by courtesy of M. Sivathanu et al. (2003)

平成18年3月2日 DEWS2006 ミニサーベイ

コラボラティブストレージ技術
Semantically-smart disk



Secure Deletion **Batch Commitment**

Experimental results by courtesy of M. Sivathanu et al. (2003)

コラボラティブストレージ技術
Second-tier cache management

- 2つのキャッシュが競合することの無駄を排除
  - バッファキャッシュ（サーバ）
  - ディスクキャッシュ（ストレージ）
- サーバ側でWRITEコマンドにヒント付け
  - sync. write
  - async. replacement write
  - async. recoverability write
- TQ (type queue) キャッシュ管理アルゴリズム
  - ストレージ側でWRITEヒントを活用

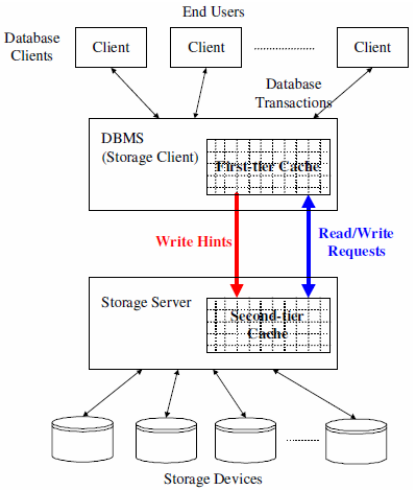


Figure by courtesy of X. Li et al. (2005)

平成18年3月2日 DEWS2006 ミニサーベイ

コラボラティブストレージ技術
Second-tier cache management

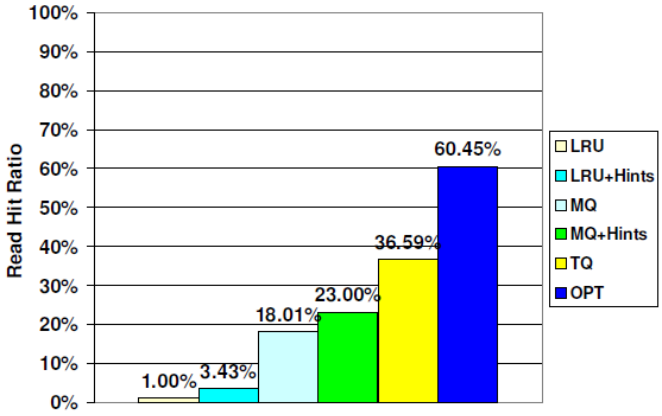


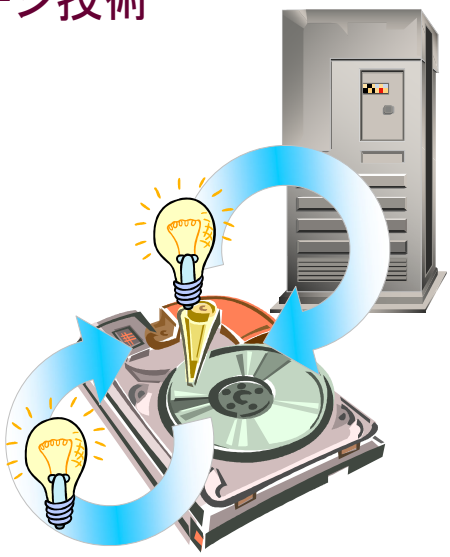
Figure by courtesy of X. Li et al. (2005)

コラボラティブストレージのまとめ

- ホスト上のアプリケーションとの緩やかな連携
  - ストレージシステムが、アプリケーションを意識することにより自己調整
- ストレージアーキテクチャの大幅な改編を回避
- 課題：
  - インターフェースの共通化
  - ヒント情報 vs. アプリケーション予測

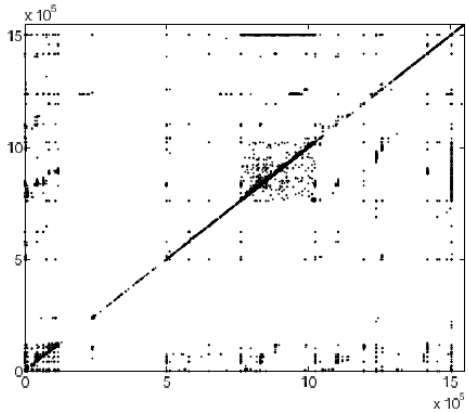
オートノミックストレージ技術

- IOの振る舞いを詳細に観察し、解析
  - システムチューニング
  - ソフトウェアデバッグ
- ストレージ管理に係るコストの大幅な低減
- 業務の変化に即応可能なシステム構成を実現



オートノミックストレージ技術
C-Miner

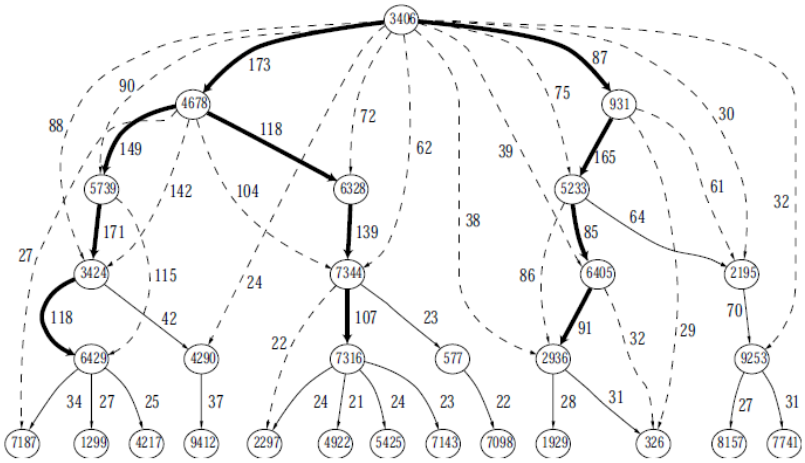
- ブロックレベルの典型的なIO振舞を抽出し、IO制御に利用
- IOトレースベースの学習方式
  - ブロック間相関抽出マイニング技術 “C-Miner” (frequent sequence mining)
- 頻出IOパターンをキャッシュ管理に活用
  - CDP (Correlation-Directed Prefetching)
  - Layout reorganization



ブロック間相関 (cello-96)

Figure by courtesy of Z. Li et al. (2004)

オートノミックストレージ技術
C-Miner



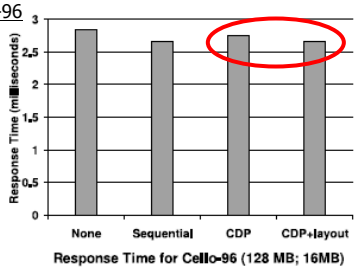
マイニングされたブロック間シーケンス

Figure by courtesy of Z. Li et al. (2004)

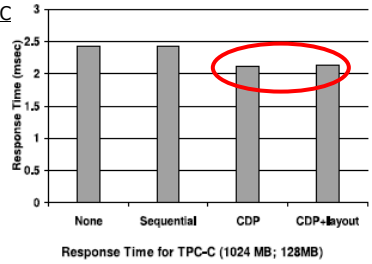
オートノミックストレージ技術 C-Miner

- IO応答性能の削減効果（Cello92/96, TPC-C, OLTP）
  - 線形プリフェッチによる改善：7-20%
  - CDP+Layoutによる改善：12-25%

Cello-96



TPC-C



Figures by courtesy of Z. Li et al. (2004)

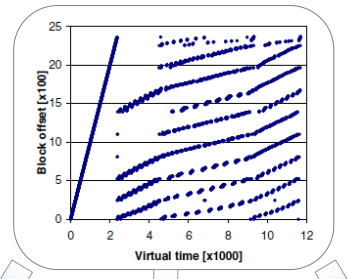
平成18年3月2日 DEWS2006 ミニサーベイ

オートノミックストレージ技術 Program-counter-based classification

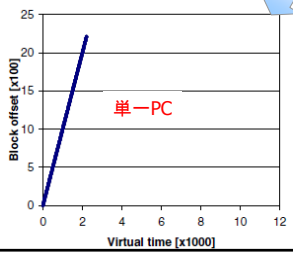
- サーバからのIO発行情報を活用
  - PC（プログラムカウンタ）を利用してIOパターンを解析
    - PCからソフトウェアサブルーチン（IOが発行されたソフトウェア文脈）を類推
    - 頻出IOパターンを抽出
      - sequential, loop, others
- → バッファキャッシュ管理に利用

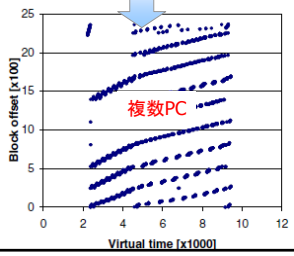
オートノミックストレージ技術
Program-counter-based classification

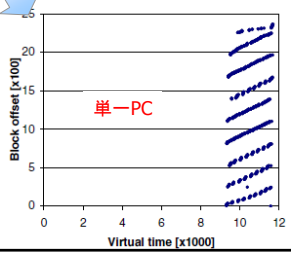
取得されたIOトレース



Figures by courtesy of C. Gniady et al. (2004)

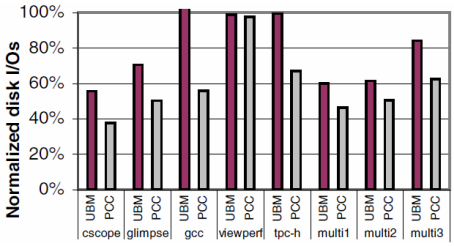


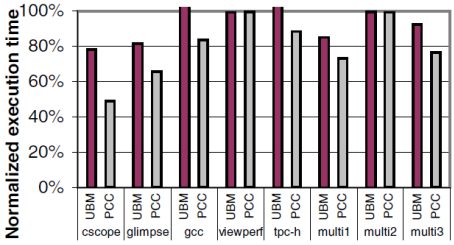




オートノミックストレージ技術
Program-counter-based classification

- UBM（ファイル毎キャッシュ管理）との比較
  - ヒット率の向上（平均：12.8%, 最大：29.3%）
  - 実行時間の短縮（平均：13.7%, 最大：29.0%）





Figures by courtesy of C. Gniady et al. (2004)

オートノミックストレージ技術
Model-driven anomaly detection

- IO負荷とスループットモデルに基づく変則検出
  - 4層のモデル化
    - OSキャッシュモデル
    - 先読みモデル
    - スケジューリングモデル
    - ストレージ装置モデル
  - → IOシステム全体を捉えるデバッギング方式

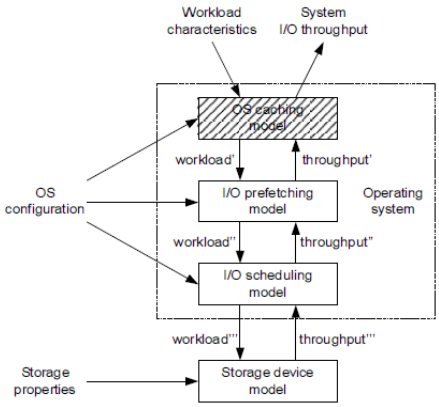
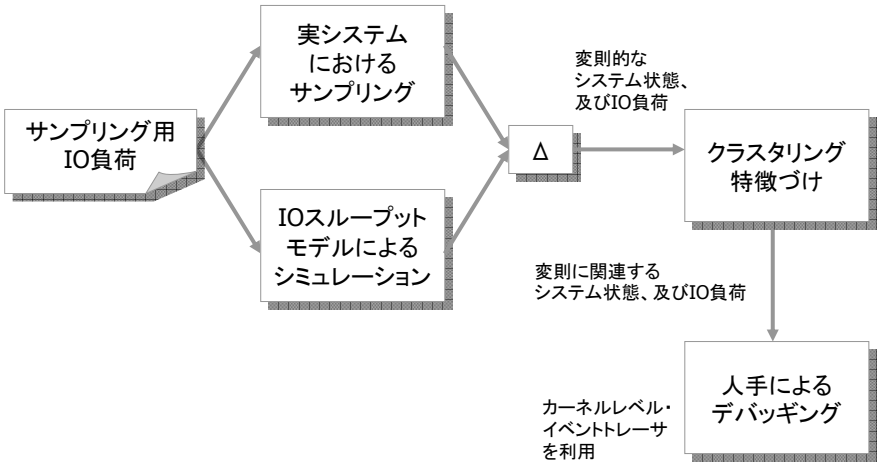


Figure by courtesy of K. Shen et al. (2005)

オートノミックストレージ技術
Model-driven anomaly detection



オートノミックストレージ技術
Model-driven anomaly detection

- Linuxカーネルに実在するバグを発見
  - 先読み(1)
  - 予測スケジューリング(2)
  - エレベーションスケジューリング(1)
- 全工程
  - サンプル取得：6分／サンプル
    → 40時間／400サンプル
  - クラスタリング、特徴付け：＜1分
  - 人手によるデバッグ：1~2日／バグ
    → ~6日／全4バグ

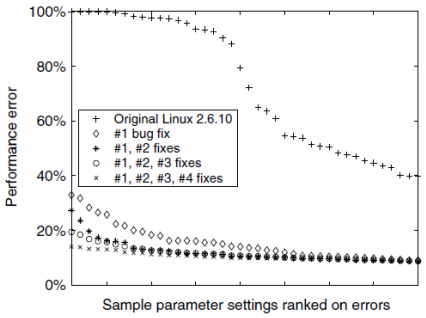


Figure by courtesy of K. Shen et al. (2005)

オートノミックストレージ技術
まとめ

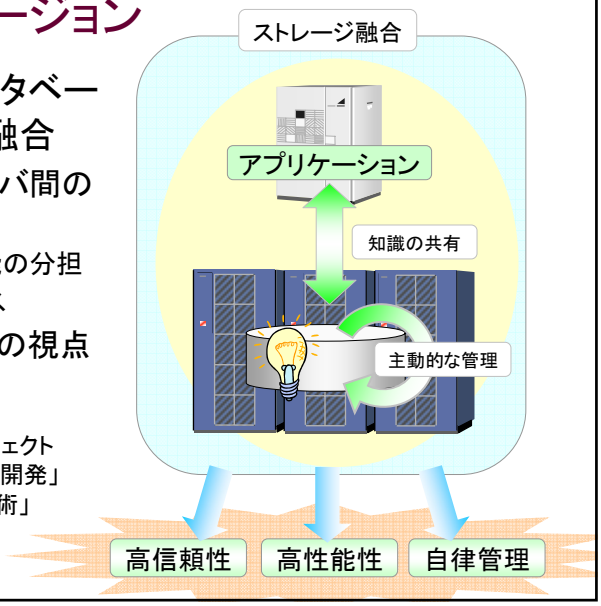
- IOトレースを広範囲に活用
- 新しい解析技術
  - データマイニング技術
  - ソフトウェアフックによる文脈予測技術
  - モデル依存変則検出技術
- 課題
  - 実システムに対する検証

ストレージフュージョン技術

ストレージフュージョン

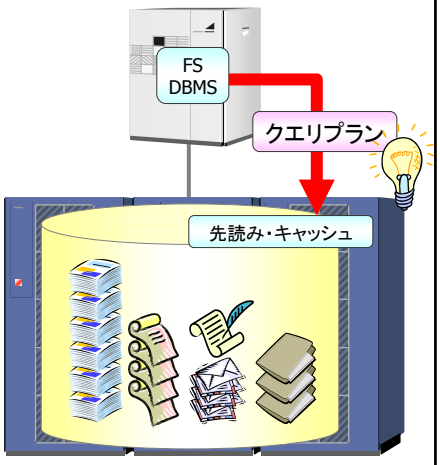
- ストレージとデータベースのしなやかな融合
  - ストレージ・サーバ間の効率的な切り口
    - データ管理機能の分担
    - インターフェース
  - 大学、企業双方の視点

文部科学省リーディングプロジェクト
「e-Society 基盤ソフトウェアの開発」
に係る「先進的なストレージ技術」



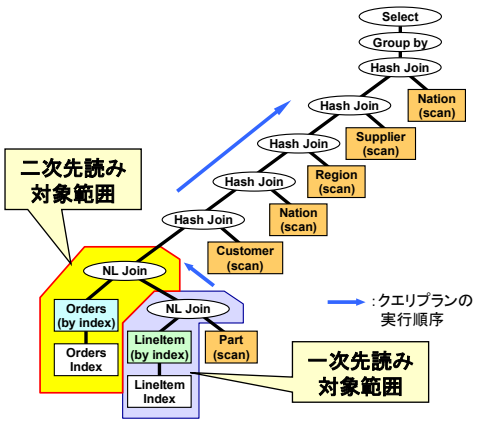
ストレージフュージョン技術 Query-plan-based prefetching

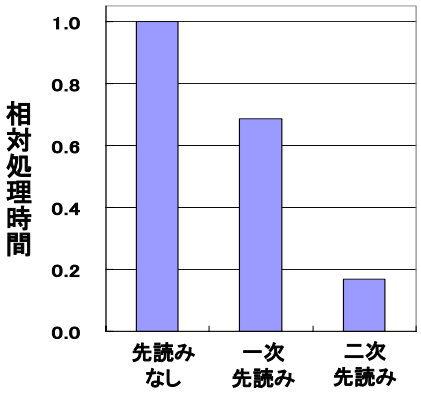
- 問い合わせ実行計画を利用して、データベース構造に基づいた先読みを実現
  - QPをストレージに注入し、索引構造に基づき先読み
  - ストレージ装置の有するキャッシュを活用
- cf. 線形先読み (sequential prefetch)



平成18年3月2日 DEWS2006 ミニサーベイ

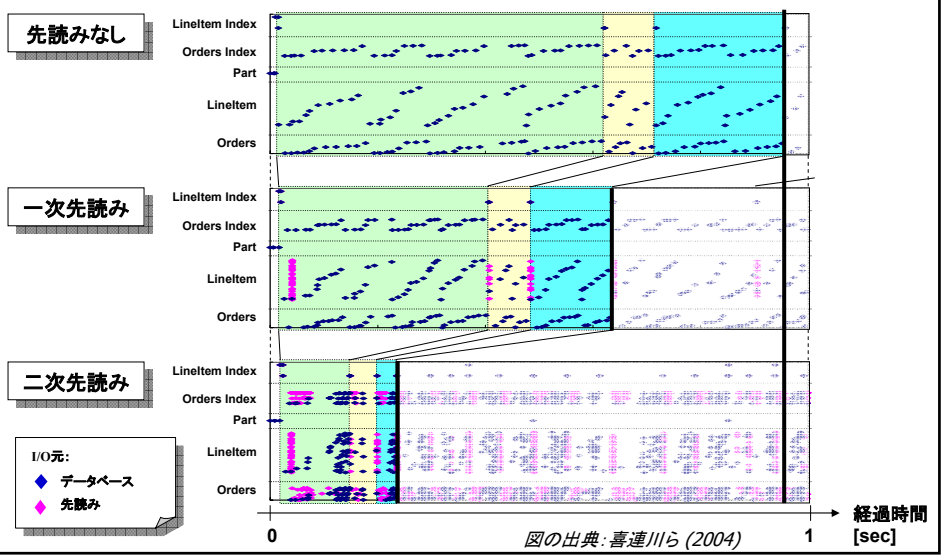
ストレージフュージョン技術 Query-plan-based prefetching





図の出典:喜連川ら (2004)

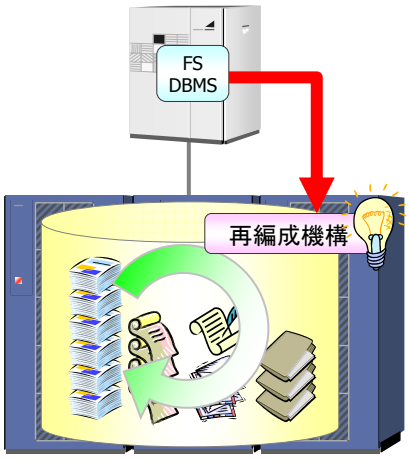
ストレージフュージョン技術
Query-plan-based prefetching



図の出典：喜連川ら (2004)

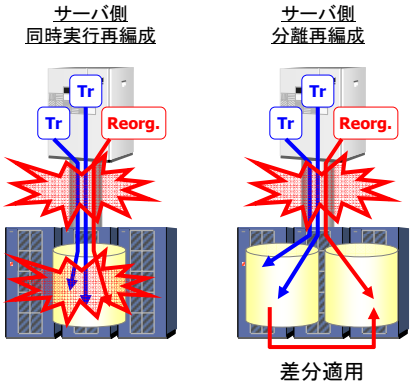
ストレージフュージョン技術
Self-reorganizing storage

- DBMSのデータ格納構造の管理機能（データベース再編成）をストレージシステムに移管
  - （高レベル）問い合わせ処理 →サーバ
  - （低レベル）データ格納構造の管理 →ストレージ
- 新しいオンラインデータベース再編成アーキテクチャ
  - 低オーバヘッド
  - 超高スループット



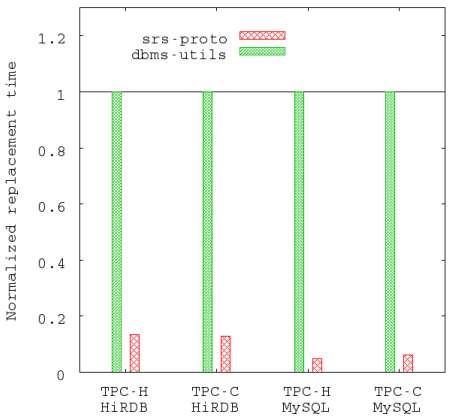
ストレージフュージョン技術
Self-reorganizing storage

- 10TB級の巨大データベースの再編成を想定
  - サーバ-ストレージ間IO帯域がボトルネック
- ストレージシステムの高い内部IO帯域を活用
  - ストレージ側分離再編成戦略
  - IO指向オンライン再編成処理
    - 物理アドレスレベルIOスケジューリング
    - パイプライン化データベース再編成
    - 高速ログ適用処理

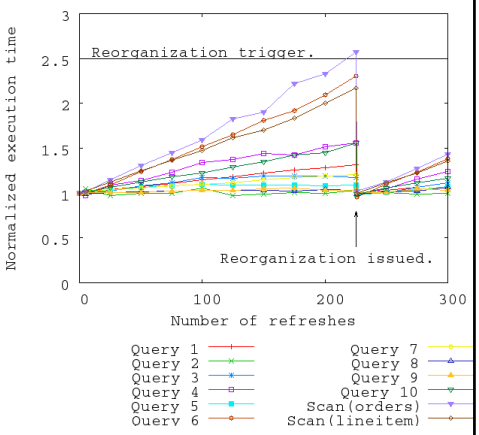


平成18年3月2日 DEWS2006 ミニサーベイ

ストレージフュージョン技術
Self-reorganizing storage



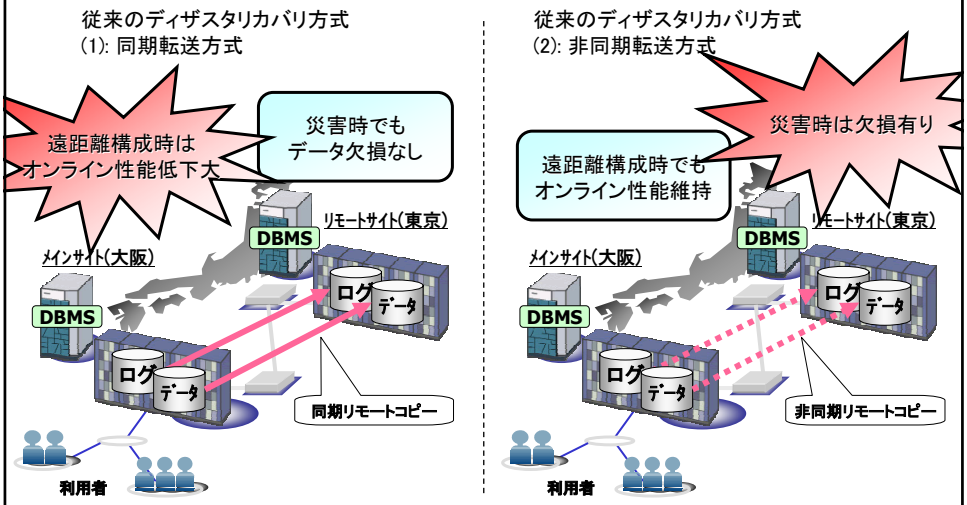
従来型データベース再編成との比較



TPC-Hにおける構造劣化とデータベース再編成の一例

図の出典：合田ら (2004)

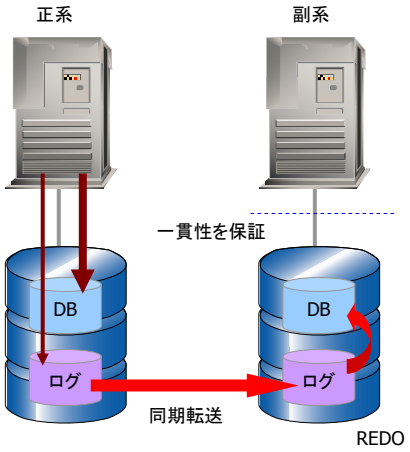
ストレージフュージョン技術
Log-based remote replication



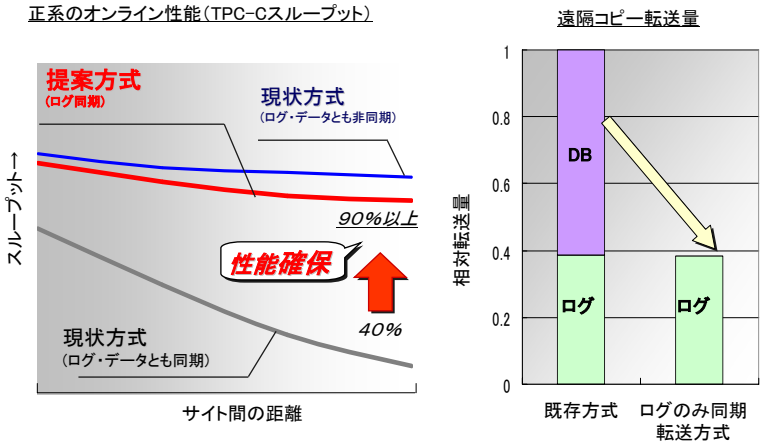
図の出典:喜連川ら (2004)

ストレージフュージョン技術
Log-based remote replication

- データベースの書き込み特性を遠隔コピー操作に活用
- ログ同期転送方式
  - 正系のオンライン性能を維持
    - 大規模災害に対する長距離転送を実現
  - 常に副系の一貫性を維持 “Transactional Storage”
    - → 災害復旧時間を大幅に短縮



ストレージフュージョン技術
Log-based remote replication



図の出典：喜連川ら (2004)

ストレージフュージョン技術のまとめ

- サーバ上のアプリケーションとしなやかに融合するストレージの高度化技術
  - 大学、企業双方の視点
  - 新しいストレージ融合技術を提案、検証
- 新しいアプローチによる高性能、高可用、低管理コストの実現
  - Query plan based prefetching
    - データベース中の索引情報をストレージのキャッシュ管理に活用
  - Self-reorganizing storage
    - データベースの格納構造管理をストレージに移管
  - Log-based remote replication
    - データベースの書き込み特性を遠隔コピー操作に活用

その他の話題

その他の話題
ストレージの自律分散化

- OceanStore (UCB)
  - 分散ハッシュに基づくP2Pストレージシステム
  - 冗長符号化、ビザンティンプロトコル
- 自律ディスク（東工大）
  - ディスクプロセッサを利用した高性能分散ファイルシステム
  - ログによる自律的複製取得
  - Fat-B木によるデータ空間配置の自律的バランス化
- FAB (HP Lab.)
  - ブリック(PC)を統合した大規模ディスクアレイ
  - 多数決原理に基づく一貫性管理プロトコル

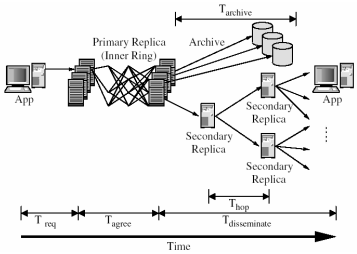


Figure by courtesy of S. Rhea et al. (2003)

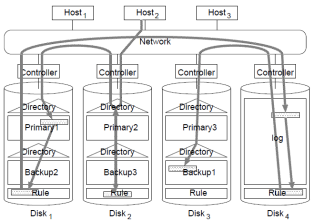
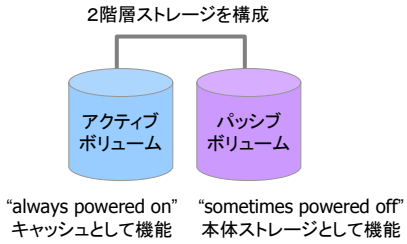


Figure by courtesy of H. Yokota. (1999)

その他の話題
ストレージの省電力化

- MAID (Copan Syst.)
  - アクティブボリュームとパッシブボリュームの2階層ストレージ構成
  - 常時通電のアクティブボリュームをキャッシュとして利用
  - 低負荷時にパッシブボリュームの給電を停止
  - 主に、アーカイブストレージとして利用
- AutoMAID (Nexsan Tech.)
  - 低頻度アクセス時に、全ディスクをスタンバイ化
    - 複数のスタンバイモードをサポート
  - アーカイブストレージ以外に、ニアラインストレージ、オンラインストレージとしても利用可能
- Hibernator (HP Lab.)
- Dynamic RPM (Pa. State Univ.)
  - 動的に回転速度を変更可能な次世代ディスクドライブの活用を目指す



60% savings – after 5 minutes of idle time, drives are programmed to unload heads and slow down to 4000 RPM. Very fast recovery time.
80% savings – after 5 minutes of idle time, drives are programmed to spin down. Recovery time on the order of 30 seconds

Configuration examples by courtesy of Nexsan. (2005)

まとめ

- 安心、快適、容易なデータ管理の要請
- ストレージシステムにおける技術革新
  - ストレージネットワーキング
  - 最近のストレージシステムの高度化技術

参考文献

参考文献
（ストレージネットワーキング）

- 喜連川 優（編）. ストレージネットワーキング. オーム社. 2002.
- 喜連川 優（編）. ストレージネットワーキング技術—SNIAストレージ技術者認定プログラム準拠. オーム社. 2005.
- SNIA. A Dictionary of Storage Networking Terminology. http://www.snia.org/education/dictionary/.
- SNIA-J. SNIA用語集. http://www.snia-j.org/dictionary/.
- SNIA Technical Council. Shared Storage Model: A framework for describing storage architectures. 2003.
- Rob Peglar. Storage Virtualization I - What, Why, Where and How? SNIA Tutorials. 2005.
- Rob Peglar. Storage Virtualization II - Effective Use of Virtualization. SNIA Tutorials. 2005.
- Nik Simpson. ILM: Tiered Storage and the Need for Data Classification. SNIA Tutorials. 2005.
- Storage Management Initiative, SNIA. SMI-S: A Standard For Managing Storage. 2005.
- ファイバチャネル協議会（編）. ファイバチャネル技術解説書. 論創社. 2001.
- JDSFファイバチャネル技術部会. ファイバチャネル技術解説書 II. 論創社. 2003.

参考文献
（データベースマシン）

- D. Slotnick. Logic per Track Devices. Advances in Computers 10, pp. 291-296. 1970.
- S. Su and G. Lipovski. CASSM: a cellular system for very large database. Proc. VLDB 1975, pp. 456-472. 1975.
- E. Ozkarahan, S. Schuster and K. Smith. RAP - Associative Processor for Database Management. Proc. AFIPS 1975 , pp. 379-387. 1975.
- D. DeWitt and P. Hawthorn. A Performance Evaluation of Database Machine Architectures. Proc. VLDB 1981, pp. 199-214. 1981.

参考文献
（アクティブストレージ技術）

- E. Riedel, G. Gibson and C. Faloutsos. Active Storage for Large-Scale Data Mining and Multimedia. Proc. 24th VLDB 1998, pp. 62-73. 1998.
- A. Acharya, M. Uysal and J. Saltz. Active Disks: Programming Model, Algorithm and Evaluation. Proc. 8th ASPLOS 1998, pp. 81-91. 1998.
- K. Keeton, D. Patterson and J. Hellerstein. A Case for Intelligent Disks (IDISKs). SIGMOD Record 27(3), pp. 42-52. 1998.
- Erik Riedel, Christos Faloutsos, Gregory R. Ganger, David Nagle. Data Mining on an OLTP System (Nearly) for Free. Proc. SIGMOD Conference 2000: pp. 13-21. 2000.
- Christopher R. Lumb, Jiri Schindler, Gregory R. Ganger, David Nagle, Erik Riedel. Towards Higher Disk Head Utilization: Extracting "Free" Bandwidth from Busy Disk Drives. Proc. OSDI 2000, pp. 87-102. 2000.
- Christopher R. Lumb, Jiri Schindler, Gregory R. Ganger. Freeblock Scheduling Outside of Disk Firmware. Proc. FAST 2002, pp. 275-288. 2002.
- Jiri Schindler, John Linwood Griffin, Christopher R. Lumb, Gregory R. Ganger:. Track-Aligned Extents: Matching Access Patterns to Disk Drive Characteristics. Proc. FAST 2002, pp. 259-274. 2002.
- Steven W. Schlosser, Jiri Schindler, Stratos Papadomanolakis, Minglong Shao, Anastassia Ailamaki, Christos Faloutsos, Gregory R. Ganger. On Multidimensional Data and Modern Disks. Proc. FAST 2005, pp. 225-238. 2005.

参考文献
（コラボラティブストレージ技術）

- Xuhui Li, Ashraf Aboulnaga, Kenneth Salem, Aamer Sachedina, Shaobo Gao. Second-Tier Cache Management Using Write Hints. Proc. FAST, pp. 115–128. 2005
- Muthian Sivathanu, Vijayan Prabhakaran, Florentina I. Popovici, Timothy E. Denehy, Andrea C. Arpaci-Dusseau, Remzi H. Arpaci-Dusseau. Semantically-Smart Disk Systems. Proc. FAST 2003. 2003.
- Muthian Sivathanu, Lakshmi N. Bairavasundaram, Andrea C. Arpaci-Dusseau, Remzi H. Arpaci-Dusseau: Database-Aware Semantically-Smart Storage. Proc. FAST 2005, pp. 239–252. 2005.

平成18年3月2日 DEWS2006 ミニサーベイ

参考文献
（オートノミックストレージ技術）

- Zhenmin Li, Zhifeng Chen, Sudarshan M. Srinivasan, Yuanyuan Zhou. C-Miner: Mining Block Correlations in Storage Systems. Proc. FAST 2004, pp. 173-186. 2004.
- Zhenmin Li, Zhifeng Chen, Yuanyuan Zhou. Mining block correlations to improve storage performance. TOS 1(2), 213-245. 2005.
- Chris Gniady, Ali Raza Butt, Y. Charlie Hu. Program-Counter-Based Pattern Classification in Buffer Caching. Proc. OSDI 2004, pp.395-408. 2004.
- Ali Raza Butt, Chris Gniady, Y. Charlie Hu. The performance impact of kernel prefetching on buffer cache replacement algorithms. Proc. SIGMETRICS 2005, pp. 157-168. 2005.
- Kai Shen, Ming Zhong, Chuanpeng Li. I/O System Performance Debugging Using Model-driven Anomaly Characterization. Proc. FAST 2005, pp. 309–322. 2005.

参考文献
（ストレージフュージョン）

- 向井 景洋，根本 利弘，喜連川 優．高機能ディスクにおけるアクセスプランを用いたプリフェッチ機構に関する評価．電子情報通信学会第１１回データ工学ワークショップ (DEWS2000)，3B-2，2000.
- 出射英臣，茂木和彦，西川記史，大枝高．クエリプランを利用した先読み技術の開発と初期評価．電子情報通信学会第１６回データ工学ワークショップ／第３回DBSJ年次大会 (DEWS2005)，5B-o1，2005.
- 河島徹，河村信男，山口浩太，藤原真二，鈴木芳生，大枝高．ストレージのリモートコピー機能を利用したDBのディザスタリカバリ方式，第６６回情報処理学会全国大会，2004.
- 合田和生，喜連川優．データベース再編成機構を有するストレージシステム．情報処理学会論文誌データベース，46(SIG 8,TOD 26)，pp. 130-147. 2006.
- 喜連川優．先進的なストレージ技術およびWeb解析技術．e-Societyシンポジウム2003講演資料集．2003.
- 喜連川優．先進的なストレージ技術およびWeb解析技術．e-Societyシンポジウム2004講演資料集．2004.

平成18年3月2日 DEWS2006 ミニサーベイ

参考文献
（自律分散技術）

- Ben Y. Zhao, John Kubiatowicz, Anthony D. Joseph. Tapestry: a fault-tolerant wide-area application infrastructure. Computer Communication Review 32(1), p. 81. 2002.
- Sean C. Rhea, Patrick R. Eaton, Dennis Geels, Hakim Weatherspoon, Ben Y. Zhao, John Kubiatowicz. Pond: The OceanStore Prototype. Proc. FAST 2003. 2003.
- Haruo Yokota. Autonomous Disks for Advanced Database Applications. Proc. DANTE 1999, pp. 435-442. 1999.
- Daisuke Ito, Haruo Yokota: Automatic Reconfiguration of an Autonomous Disk Cluster. Proc. PRDC 2001, pp. 169-172. 2001.
- Svend Frølund, Arif Merchant, Yasushi Saito, Susan Spence, Alistair C. Veitch. FAB: Enterprise Storage Systems on a Shoestring. Proc. HotOS 2003, pp. 169-174. 2003.
- Yasushi Saito, Svend Frølund, Alistair C. Veitch, Arif Merchant, Susan Spence. FAB: building distributed enterprise disk arrays from commodity components. Proc. ASPLOS 2004, pp. 48-58. 2004.

参考文献
（省電力化技術）

- Dennis Colarelli, Dirk Grunwald, Michael Neufeld. The Case for Massive Arrays of Idle Disks. Proc. FAST 2002, WiP report. 2002.
- Copan Systems. http://www.copansys.com/.
- Fred Moore and Aloke Guha. Introducing COPAN Systems MAID architecture (Massive Arrays of Idle Disks). White Paper, Copan Systems and Horison Information Strategies. 2004.
- Nexan Technologies. http://www.nexsan.com/.
- Qingbo Zhu, Zhifeng Chen, Lin Tan, Yuanyuan Zhou, Kimberly Keeton, John Wilkes:. Hibernator: helping disk arrays sleep through the winter. Proc. SOSP 2005, pp. 177-190. 2005.
- E. Pinheiro, and R. Bianchini. Energy conservation techniques for disk array-based servers. Proc. ISC 2004, pp. 68-78. 2004.
- Sudhanva Gurumurthi, Anand Sivasubramaniam, Mahmut Kandemir, and Hubertus Franke. Reducing Disk Power Consumption in Servers with DRPM. Computer 36(12), pp.59-66. 2003.
- H. Yada, H. Ishioka, T. Yamakoshi, Y. Onuki, Y. Shimano, M. Uchida, H. Kanno, and N. Hayashi. Head positioning servo and data channel for HDDs with multiple spindle speeds. IEEE Transactions on Magnetics, 36(5), pp. 2213–2215. 2000.

謝辞

- 本講演資料を作成するにあたり、東京大学生産技術研究所の喜連川優教授にご指導をいただきました。
- 本講演資料において紹介する「ストレージフュージョン技術」は、文部科学省リーディングプロジェクト「e-Society 基盤ソフトウェアの開発」に係る「先進的なストレージ技術」よるものです。当該プロジェクトの推進にあたっては、協力企業である株式会社日立製作所より多くの有益なコメントを頂戴しました。