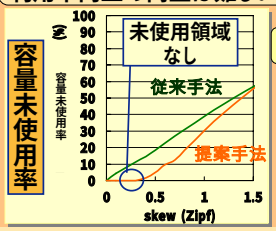
これまでの成果

アクセス負荷分散と容量利用率向上のための複製管理戦略

Adaptive Overlapped Declustering

負荷分散 / 従来手法

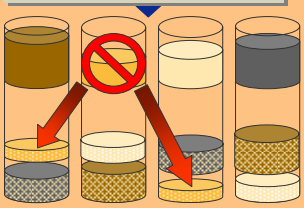
アクセス負荷分散と容量利用率向上の両立は難しい



プライマリでアクセス負荷分散

分割バックアップの配分調整で容量利用率の向上

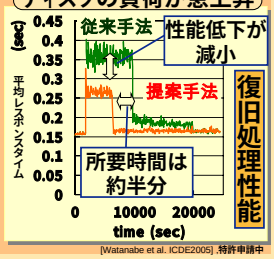
同時復旧で所要時間半減



分割バックアップに負荷が分散され性能低下が減小

従来手法 / 障害復旧

バックアップのあるディスクの負荷が急上昇



[Watanabe et al. ICDE2005] ,特許申請中

バージョン間の利用頻度の差を利用した負荷分散と容量分散の同時実現

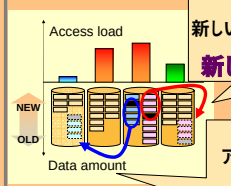
COBALT

新しいバージョンには高速にアクセスできるほうが良い

差分を利用したバージョンファイルへのアクセス頻度を考慮してディスク装置の**アクセス負荷分散**と**容量分散**を同時に実現

古いバージョンも必要時に取り出せるようにすることが必要

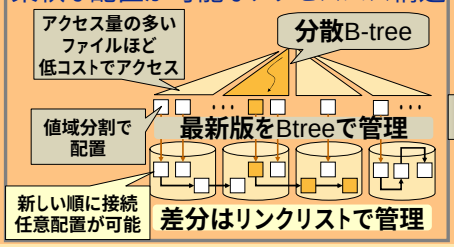
アクセス頻度の差を用いたデータ配置戦略

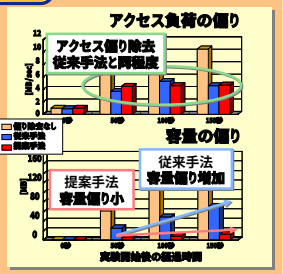


アクセス量の分散
新しいファイルへのアクセス量は多い
新しいファイルを移動させる

データ格納量の分散
アクセス分布に影響を与えない
古い差分情報を移動させる

柔軟な配置が可能なアクセスパス構造



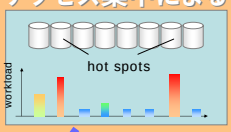


[NAKANO et al. DEWS2005] ,特許申請中

管理コスト削減を実現する自律的データ再配置管理と通常処理の両立と性能保証

Replica-assisted Migration

アクセス集中による性能低下



長期的HOTSPOTの除去

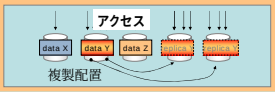
データマイグレーション

ストレージノード間のデータ再配置によりアクセス負荷を分散

負荷を減らすべきストレードノードでデータ管理による負荷が発生
→通常処理の性能保証が困難

一時的 HOTSPOTの除去

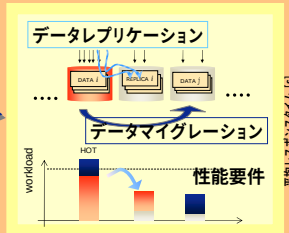
データレプリケーション



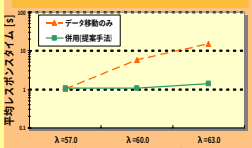
複製データを複数のストレージノードへ分散配置し，アクセスを回送

併用＆協調スケジューリング

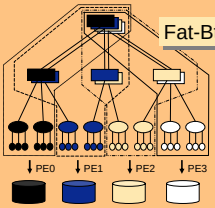
負荷変動予測とシステム性能要件情報を利用し，データマイグレーションによる上昇負荷をデータレプリケーションにより他のストレージノードへ分散



通常処理の性能劣化を抑えることが可能．



大規模ストレージシステム構成のための分散Btree並行性制御



並列Btree

無共有並列計算機上の並列DBに適した効率的なインデックス構造

並行性制御が必須

Btree構造更新時の一貫性を保証システム処理性能に影響大

MARK-OPT

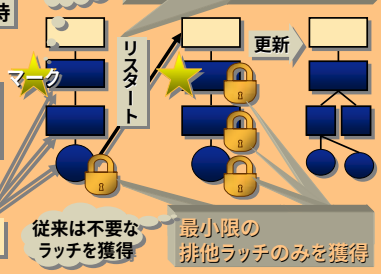
目的

更新発生時のシステム処理性能の維持

アプローチ

楽観的処理*中のマーキングによる更新発生範囲の先読み

*ラッチ範囲不足時にラッチの全解放再獲得(リスタート)が必要



従来は複数回リスタートが必要

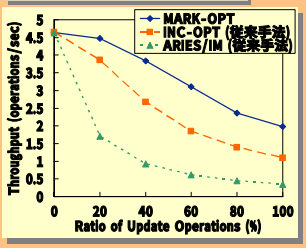
1回のリスタートのみでたいていの更新処理を完了

実験

従来手法の性能を常に改善

大規模高更新環境では大幅改善

マーキングの有効性を実証



[Yoshihara et al. DEWS2005], 特許申請中