

## 要 約

## A. システム要件

70億件のトランザクションデータを保管・更新する大規模データベースシステムの構築事例を紹介します。

それまでのシステムは、「国税関係帳簿の電磁的記録等による保存等の承認申請を行うための法律要件(電子帳簿保存法)」を満たしていなかったため、大量の紙の帳票を法定期限まで廃棄できず、倉庫コスト、人的稼働の面で早急な改善が求められていました。これは、対象とするデータが大き過ぎたため、それまでのシステムでは、処理できなかったためです。

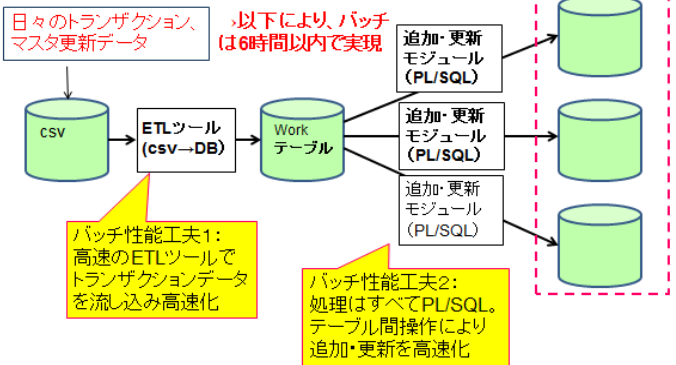
ここでは、データベースの構成の工夫とテーブル間演算操作の導入によって上記問題を解決するシステムを開発しましたので、その概要を報告します。

- (1) 対象データをカバーするシステムの構築。
  - ・トランザクションデータ**10年分**で**70億件**
  - ・日々250万件、**月5800万件**のデータ
  - ・しかも、約10%以上は数年に及ぶ更新データ
- (2) 電子帳簿保存法の法律要件を満たすこと
  - ・全件データの保存と見読機能を充足。性能保証。
  - ・特別に大量の検索でない限り、ストレスの生じない範囲内に検索を可能とする(**数秒~数分**)  
(**国税調査官**が操作することもある)
- (3) データは直近**3年間**のものと、それ以降のもの2段階で保存。**10年経過**したデータは削除

## B. 適用技術(その1): バッチ性能向上

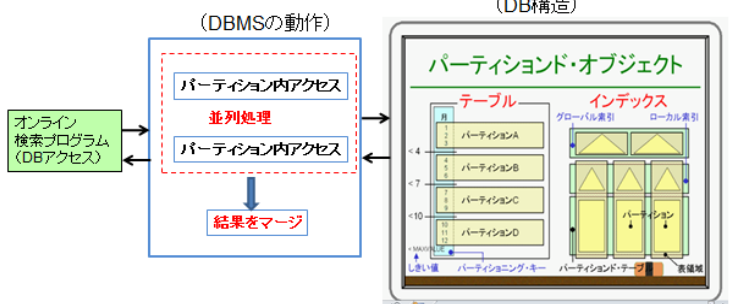
目標: バッチは、7時間のバッチ時間帯で、すべて追加・更新を完了すること

更新先保存データ(16種)



## C. 適用技術(その2): オンライン性能向上

データのパーティション分割方式を採用 → パーティション分割方式の最適設計を進めた!!  
DBアクセスは、複数パーティションを並列に実行させ、結果はマージして出力



## D. データベースのパーティション設計

- (1) 大量月次データ(明細は、2000万件/月)のパーティション設計  
→ 月単位でパーティションを構成。Indexは、運用性、性能面からlocal indexを採用。

売掛金伝票	(Global)	(Local)	年月	伝票番号	登録日	伝票種	データ
			当年度(4月)	年月	伝票番号	登録日	伝票種
当年度(3月)	年月	伝票番号	登録日	伝票種	データ		
前3年度(4月)	年月	伝票番号	登録日	伝票種	データ		
前3年度(3月)	年月	伝票番号	登録日	伝票種	データ		
復元年度(4月)	年月	伝票番号	登録日	伝票種	データ		
復元年度(3月)	年月	伝票番号	登録日	伝票種	データ		

- (2) 年次データ(明細は、数百万件/年)のパーティション設計  
→ 年単位でパーティションを構成。Indexは、(1)と同様 Local indexを採用。

返品実績	年月	伝票番号	伝票種	伝票ID	伝票ID	データ
当年度(3月)	年月	伝票番号	伝票種	伝票ID	伝票ID	データ
全3年度(4月)	年月	伝票番号	伝票種	伝票ID	伝票ID	データ
全3年度(3月)	年月	伝票番号	伝票種	伝票ID	伝票ID	データ
復元年度(4月)	年月	伝票番号	伝票種	伝票ID	伝票ID	データ
復元年度(3月)	年月	伝票番号	伝票種	伝票ID	伝票ID	データ

- (3) マスタ類 → パーティション化せず

## E. 性能測定結果

- (1) バッチ処理
  - ① 運用時間帯(夜間7時間)で、すべてのバッチ処理を修了。
  - ② 日々の差分バックアップ、週次の全件バックアップも問題なし。
- (2) オンライン検索時間
  - ・レスポンス時間(測定例から)
    - ① 結果が1000件未満なら3秒以内で検索可能
    - ② 結果が10万件で21秒で検索可能
    - ③ 結果が40万件で90秒で検索可能

以上の通り、バッチ、オンラインともに要件を満たす「大規模データベースシステムの構築」ができました。

本システムは、現在8年目を迎えますが、日々問題なく現役で運用されております。