三菱“DIAPRISM”の高速化技術

平岡精一* 郡 光則*
藤原聡子* 東 辰輔*
山岸義徳* 三角武嗣**

要 旨

今回，三菱“DIAPRISM”では，PCサーバ／Windows NT上で大容量データを活用したデータウェアハウスシステムやデータマートシステムを構築するための高速化技術を開発した。

データウェアハウスシステム及びデータマートは，

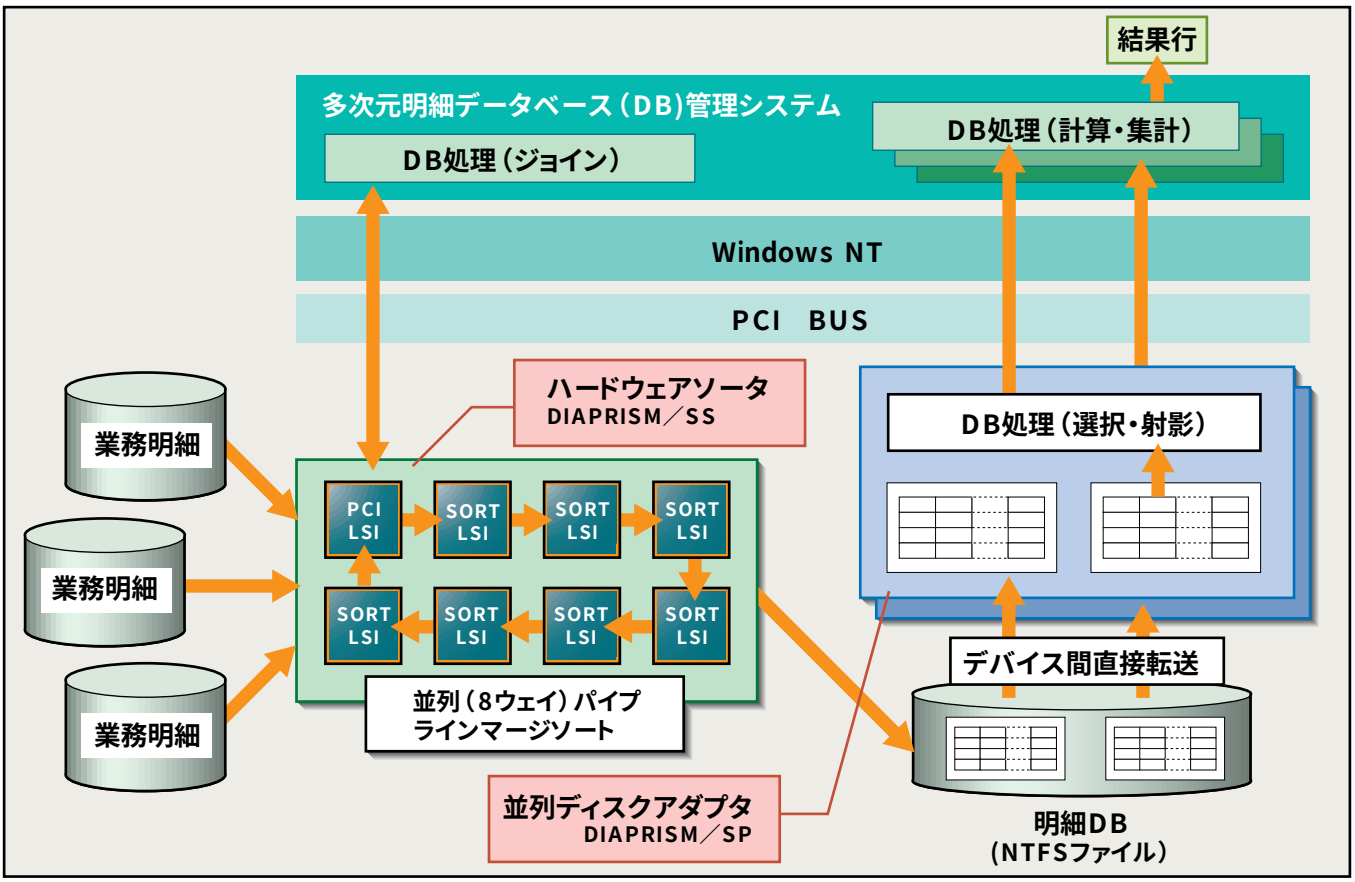
(1) 基幹業務における複数のデータベースからデータを抽出し，これを目的別に編集し，時系列に蓄積する処理

(2) 時系列に蓄積された大量情報から検索集計処理を行い，分析する処理

で構成される。

(1)のデータを時系列に蓄積する処理では，複数の表データを結合（ジョイン）して情報活用の目的に合わせて明細データ形式を形成してから，ロード蓄積する必要がある。DIAPRISMでは，抽出データに対する結合処理に対する高速化を，並列（８ウェイ）パイプラインマージソートというアルゴリズムを実装したハードウェア“DIAPRISM／SS”によって実現した。

(2)の大量情報を検索集約する処理では，ディスクと主メモリ間のデータ転送が性能上のボトルネックになる。DIAPRISMでは，処理に必要なデータをディスクから読み出すことでデータ転送量そのものを減らすデータ配置技術と，データベース処理のうち選択及び射影などの前半処理部分を主メモリに入力する前にハードウェア“DIAPRISM／SP”によって並列的に実行するＩ／Ｏプロセッサ技術という二つの基本技術によって，PCサーバ／Windows NTにおいて並列機などに匹敵する高速性を実現している。



DIAPRISM基本処理方式

DIAPRISMでは，データを蓄積する処理においてはハードウェアソータ(DIAPRISM／SS)によるソート機能で，データを検索集計する処理においてはデータ配置技術と並列ディスクアダプタ(DIAPRISM／SP)によるＩ／Ｏプロセッサ技術で高速化処理を実現している。

 *情報通信システム開発センター **情報システム製作所