

超高性能データベースシステムへの期待

東京大学
生産技術研究所

教授 喜連川 優

言うまでもなくコンピュータ関連技術の進展は著しい。コンピューティングパワーとストレージキャパシティのいずれをとっても、2年しないうちに大きく色あせる時代となりつつあるが、データベース技術に対しどちらがより大きなインパクトを与えているのであろうか？

データウェアハウスなる用語は、当初バズワードと言われていたが、今では、学会においてもすっかり定着しており、トランザクション処理を主体とする勘定系データベースから分離した、非定型的な検索を主体とする新たなデータベースの構築が進められている。これは、分散から集中へのある意味では逆方向の流れともとらえられるが、そのドライビングフォースは、二次記憶コストの低減によるところが大きい。データウェアハウスに加えて、更にアプリに特化したデータマートを形成することが通例のようであるが、そうなるとう全部で三重のコピーを維持することにもなるからである。

別の見方として、View Materialization(ビュー実体化)なる技術が注目されている。通常、複雑な情報系問合せでは、多数のリレーションのジョインが必要であり、極めて負荷の重い処理と考えられてきた。このため、マルチカラムインデックス、スタージョイン、ビットマップインデックスなど、幾つかの手法が生み出されてきたが、更なる高速化技術としてView Materializationが注目されている。必要なViewをあらかじめ作っておくと、大きく高速化可能となる。ビューの実体化にはベースリレーションに比べて大きな空間が必要となるため、むやみに作ることはできないものの、二次記憶のビット単価が大幅に低下してきたことが大きなけん(牽)引力となっていることは間違いない。

データウェアハウスと同様な発想の下、Web Warehousingなる言葉も現れつつある。よく利用される問合せは大幅な性能向上が期待される。ここでもDiskの容量増加に期待するところがある。

また、データウェアハウスのフロントエンドとしてのOLAPもその地位を確立しつつある。高速なユーザーイン

タフェースを提案するMOLAPもユーザーニーズに適合した柔軟なビュー生成とみなせるが、現状では、MOLAPのデータ構造生成は膨大な時間を必要としている。オンラインとはいえ、オフラインでのキューブ生成処理が不可欠となっている。

こうして眺めてみると、安いディスクスペースで不足するCPUパワーを補っているというのが現状のハイエンドデータベースの世界と言えよう。すなわち、データベース処理の高性能化はますます重要な研究開発課題と言える。データベースミドルウェアは米国に比べて大きく差をつけられてしまっているのが現状であり、何とか巻き返しを図りたいものである。

データベースは、今後リレーショナルモデルからユーザー定義可能なメソッドの取り込みを可能とするオブジェクトリレーショナルモデルへとより一層高度なアプリケーションの支援が求められており、処理負荷はますます大きくなる。さらに、ERPを始めとするデータベースを中心技術とする統合情報管理システムもCPUインテンシブであり、当面、CPUパワーが十分であるという状況にはならないと予想される。

米国では、ASCIなる超並列スーパーコンピュータプロジェクトが活気づいているようであるが、数年前からはデータベースにおける並列処理の方が、科学技術計算に比べ、より広く利用されるに至っている。データベース演算の並列化はユーザープログラムを全く変更することなく実現可能であり、並列プロセッサはCPUパワーを補う極めて有効な手法である。一方で、ハードウェアの開発環境の整備が進み、容易にASICを作成可能になったことから、専用エンジンは小さなコストで大きな付加価値を与え得るポテンシャルを秘めている。

超大規模データベースを対象とした並列プロセッサ・専用コンピュータによるデータベース演算性能のブースティングは極めて興味深い重要な領域と言えよう。