

# 5D-1 スケーラブルインテリジェントストレージアーキテクチャによる 並列データウェアハウス

清水 英弘 郡 光則  
三菱電機株式会社 情報技術総合研究所

## 1. はじめに

大規模なデータを対象とする情報検索処理では、一般にはデータ転送がボトルネックとなることが多い。我々は、ストレージの並列処理により、データ転送ボトルネックを解消しスケーラブルな性能を実現する方式の開発に取り組んできた [1][2]。この方式を SISA (Scalable Intelligent Storage Architecture) と呼ぶ。

本稿では、SISA によるデータウェアハウスの実装および、実業務データを用いた評価を行うことによりデータウェアハウスにおいて SISA が有効であることを実証する。

## 2. スケーラブルインテリジェントストレージアーキテクチャ (SISA)

### 2.1. 特長

SISA は、スケーラブルな性能を出すことを設計目標としている [3]。全文検索システムにおける SISA の適用効果については、[2]および[3]で報告している。

### 2.2. 構成

SISA は、汎用のハードウェア品により実現可能である。データ規模が小さく並列度があまり必要ない場合は、PC サーバ 1 台で主記憶共有型の SMP/SMT (Symmetric Multi-Threading) により並列処理を行う構成とし、データ規模が多い場合は複数のインテリジェントストレージノード (ISN) [2][3] により並列処理を行う構成とすることが可能である。

### 2.3. データウェアハウス処理への適用

SISA を以下に示す方法でデータウェアハウスに適用し高速化を実現する。

- (1) 選択・射影・集計のデータベース処理を並列に実行し、CPU 処理ボトルネックを解消。
- (2) テーブルをトランスポートし特定のカラムだけアクセスし、ブロック化によりシーケンシャルなディスクアクセスを実現。 [1]
- (3) 高速な伸張が可能な独自の圧縮方式により、テーブルを圧縮して I/O 負荷を削減。 [4]

図 1 に SMP/SMT 並列処理のデータの流を示す。

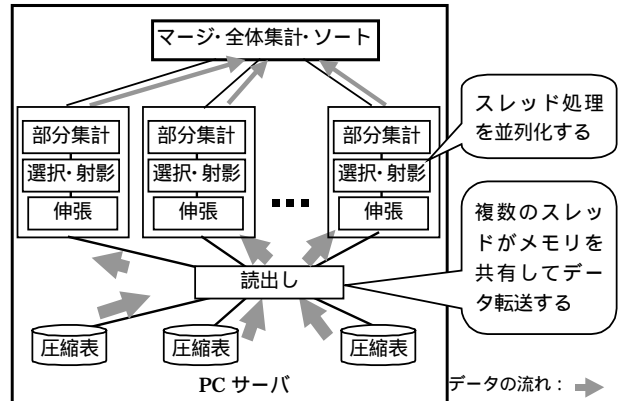


図 1 SMP/SMT 並列処理のデータの流れ

問合せ実行時に、圧縮されたテーブルをディスクから読み出し、伸張、選択、射影、部分集計などの前処理を、複数のスレッドにより並列処理する。最後に前処理結果をマージ、全体集計、ソートする。この一連のデータ転送の流れをデータドリブンのスケジューリングを行うことで、途切れない並列処理を行う。

図 2 に、ISN 並列処理のデータの流を示す。

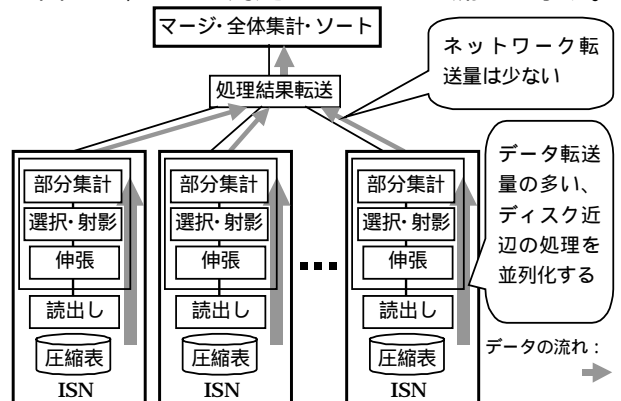


図 2 ISN 並列処理のデータの流れ

それぞれの ISN で圧縮されたテーブルをディスクから読み出し、伸張、選択、射影、部分集計処理を行う。データ転送量の多いディスク近くの処理を並列化することで、データ転送ボトルネックを解消する。また、前処理結果だけネットワーク転送しネットワーク転送量を削減する。

## 3. 性能評価

SISA による並列データウェアハウスシステムを実装し、実際の業務で利用しているデータを

A Data Warehouse by Parallel Processing with the Scalable Intelligent Storage Architecture.  
Hidehiro Shimizu, Mitsunori Kori  
Information Technology R&D Center, Mitsubishi Electric Corporation

用いた評価を行う。処理対象とするデータ量を一定として、PC サーバ 1 台の SMP/SMT 構成と、複数台の ISN を使う ISN 構成のそれぞれについて、スケーラビリティを評価した。

### 3.1. 評価条件

表 1 に、データの種類と Query の数を示す。

No	内容	Query	レコード件数	データ量
1	Webログデータ	6個	708Bytes	2億件 133.5GByte
2	商品分析データ	4個	128Bytes	2億件 24.0GBytes
3	仕分明細データ	3個	692Bytes	2億件 128.8GByte

表 1 データの種類

SMP/SMT 構成については、表 2 のハードウェア構成で OS の起動パラメータを指定することにより稼動 CPU 数を変化させる。ISN 構成については、表 3 のハードウェア構成で ISN 数を変化させる。

表 2 SMP/SMT 並列処理ハードウェア構成

ホスト ノード	OS	Windows2003 Server x64 Enterprise Edition
	CPU	Xeon 3.66GHz × 4
	Memory	16GB
	HDD	15000rpm 137GB × 8
	SCSI	Ultra320 × 1

表 3 ISN 並列処理ハードウェア構成

ホスト ノード	OS	Windows2003 Server
	CPU	Xeon 3.20GHz × 2
	Memory	2GB
	LAN	1000Base-T
インテリジェ ントストレ ージノード (1 台あたり)	OS	Linux Fedra Core 3
	CPU	Pentium4 3.06GHz
	Memory	1GB
	HDD	15000rpm 73GB × 3
	SCSI	Ultra320 × 1
	LAN	1000Base-T

### 3.2. SMP/SMT による並列処理性能

表 1 の 13 個の Query について、表 2 のハードウェア構成で稼動 CPU 数を 1~4 に変化させた、CPU 数 1 を基点とした応答性能の比を図 3 に示す。

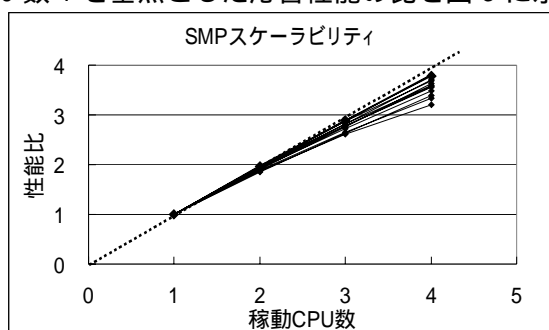


図 3 SMP 並列処理のスケーラビリティ

稼動 CPU 数 4 個までほぼ比例して検索性能の向上が確認された。

### 3.3. ISN による並列処理性能

表 1 の 13 個の Query について、表 3 のハードウェア構成で ISN 数を 2~8 に変化させた、ISN 数 2 を基点とした応答性能の比を図 4 に示す。

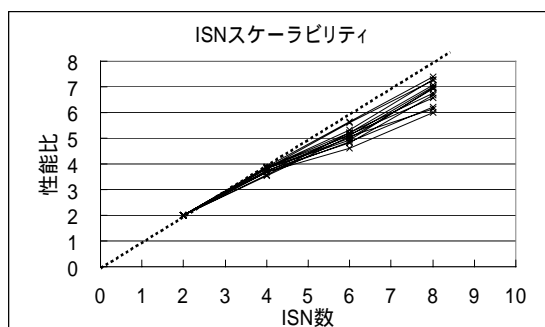


図 4 ISN 並列処理のスケーラビリティ

ISN 数 8 までほぼ比例して検索性能の向上が確認された。ISN 並列処理では、前処理結果のネットワーク転送負荷がボトルネックとなる可能性がある。ISN 数 8 の場合について前処理結果のネットワーク転送スループットを実測した結果、ほとんどが 10MB/sec 以下であった(図 5)。したがって ISN 数 100 個程度まで 1000Base-T の転送性能限界とならず、十分なスケーラビリティが得られると予想できる。

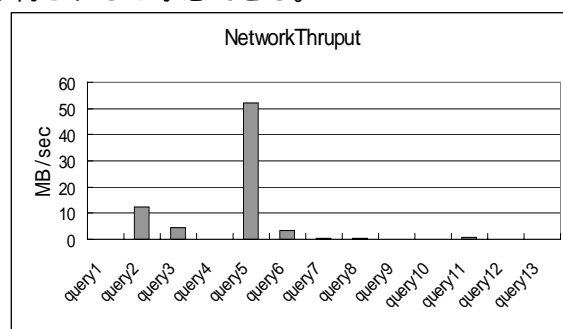


図 5 ネットワーク転送スループット

## 4. まとめ

実業務データを用いた評価を行い、CPU 数にほぼ比例したスケーラビリティが得られていることを確認し、SISA がデータウェアハウスに対しても有効であることを実証した。ISN 並列処理においては、ISN 数が 100 個程度まで対応可能であると推定される。

## 参考文献

- [1] 道下学、他、”データベースプロセッサ DIAPRISM(2)データ管理方式”, 情処全大第 57 回, 4K-08 Oct. 1998
- [2] 清水英弘、他、”スケーラブルインテリジェントストレージによる大規模並列全文検索の実現”, 情処全国大会第 64 回, 4ZA-4, Mar. 2002
- [3] 郡光則、他、”検索機能を備えたストレージシステムによる大規模並列全文検索”, 信学技報、CPSY-2002-47, Aug 2002
- [4] 郡光則、他、”大規模データ処理向け超並列可逆データ圧縮伸長処理技術の開発”, 2001 年度 IPA 先端的情報化推進基盤整備事業報告