

追記型データベースのスケラビリティ評価

加藤 守† 山岸 義徳† 中村 隆顕† 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

ログ、センサデータ、メールなどの、時系列的に追加される大量の追記型データの高速処理を目的として、スケラブルインテリジェントストレージ方式 SISA(Scalable Intelligent Storage Architecture)を適用した追記型データベースを開発した。

本稿では、SISA による追記型ログデータベースについて検索ノード追加による検索性能のスケラビリティを検証しその有効性を実証する。

2. 背景と課題

2.1. 大規模追記型データ管理の課題

ペタバイト級におよぶ大量の追記型データの長期保存を行うシステムでは、ストレージコストの増大や、多額の初期導入コストが課題となる。そのため、データ圧縮によるストレージ削減や、段階的なシステム規模拡張による初期導入コストの削減が求められている。

2.2. SISA 概要

これらの課題に対して、我々は、ストレージやプロセッサの並列処理、効率的なデータアクセスおよびデータ圧縮により、データ転送や処理のボトルネックを解消しスケラブルな性能を実現するSISA共通基盤ソフトウェア、およびSISA上の各目的別の追記型データベースの開発に取り組んできた(図1)。全文検索エンジンFTSへのSISA適用効果は[1]で、高速集計検索エンジンAQLによるデータウェアハウスへの適用効果は[2]で報告した。

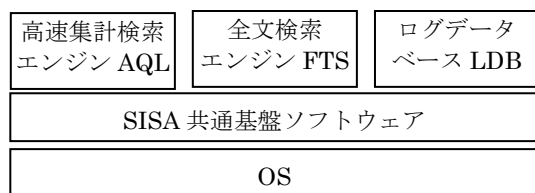


図1 SISA と追記型データベース

SISA は、汎用のハードウェア品により構成可能であり、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-Processing / Simultaneous 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) により並列処理を行う構成とする。

2.3. ログデータベースLDB

ログデータベースLDB[3]は、大量に発生する多様な形式のログを一元管理するためのデータベース管理システムであり、以下の特長を持つ。

- ・ ログを形式によらずそのまま蓄積保存。
- ・ 正規表現指定による高速検索。
- ・ ログを圧縮保存することでストレージ容量を概ね 1/10 以下に削減。

LDBのSMP構成によるスケラビリティは[4]で報告した。

3. LDBのスケラビリティ拡大

3.1. LDBへのISN適用方式

ペタバイト級の大量ログの長期保存を行うために、LDBにISNを適用することで、以下を実現する。

- ・ ストレージ削減のためのデータ圧縮により増大するプロセッサ負荷の分散
- ・ 段階的なサーバ追加による性能向上

ISNの適用は以下の方法による。

- (1) データの書き込み時には、複数のレコードからなるまとまったサイズのブロック単位で圧縮し、圧縮ブロックの形式で各ISNのディスクに均等に格納する。

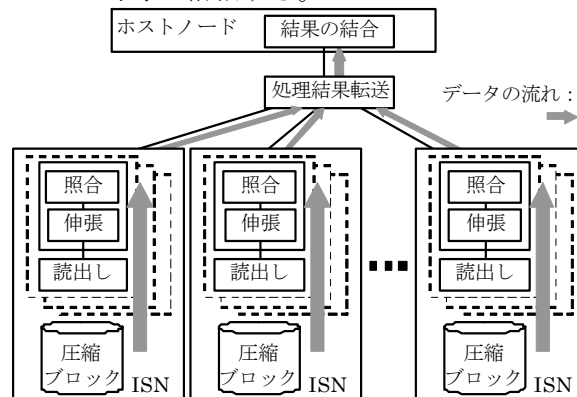


図2 ISN 並列検索処理のデータの流

- (2) データの検索時には、圧縮ブロックのメモリ上への伸張と、メモリ上での正規表現の検索条件の照合をISNに分散して並列処理する。さらに各ISNにおいては複数のプロセッサに

Scalability Evaluation of a Write-Once Database

†Mamoru Kato, Yoshinori Yamagishi, Takaaki Nakamura, Mitsunori Kori

Information Technology R&D Center, Mitsubishi Electric Corporation

よりこれらの処理を圧縮ブロック単位で並列に実行することで、LDBにおいて処理負荷の大きい伸張処理の高速化を実現する(図 2)。

3.2. ISN構成LDBの性能モデル

ISN 構成の LDB の 1 回の検索にかかる全検索処理時間 T は式(1)で表される。ここで、 T_0 は起動処理時間、 T_n は ISN-ホストノード間のデータ転送時間、 T_d は ISN での圧縮データ伸張および照合処理時間、 T_r は圧縮ブロック読出し時間、 T_w は検索結果の書込み処理時間である。

$$T = T_0 + \text{MAX}(T_n, T_d, T_r, T_w) \quad (1)$$

T_n 、 T_d 、 T_r 、 T_w は以下の式(2), (3), (4), (5)で表される。但し、 S はログデータサイズ、 C は圧縮率、 P は選択率、 V_n 、 V_d 、 V_r 、 V_w はそれぞれデータ転送、伸張・照合、読出し、書込みの速度、 N_d は ISN のディスク総数、 N_p は ISN のプロセッサ総数である。

$$T_d = S / (V_d \times N_p) \quad (2) \quad T_n = P \times S / V_n \quad (4)$$

$$T_r = (1 - C) \times S / (V_r \times N_d) \quad (3) \quad T_w = P \times S / V_w \quad (5)$$

圧縮率 C が大きく、選択率 P が小さい場合、ログデータサイズ S が十分大きければ T は概ね T_d で近似でき、検索速度は N_p に比例する。

4. 評価

ISNによるログデータベースを商用システム [5] 上に実装し、実際のログデータを用いて検索処理のスケラビリティの評価を行った。

4.1. 評価環境

評価には表 1 に示す PC サーバを用いた。ISN は同一構成の PC サーバを 1 台~8 台使用した。

表 1 評価環境

ホストノード	OS	Windows Server 2003 R2 (64 bit)
	CPU	Xeon 3060 2.40GHz (2 core) × 1
	Memory	4 GB
	HDD	250GB, 7,200rpm, SATA2 × 4 (RAID 5)
	LAN	1000Base-T
ISN (1 台 当り)	OS	Linux Fedora 11 x64
	CPU	Xeon E3110 3GHz (2 core) × 1
	Memory	8GB
	HDD	300GB, 15,000rpm, SAS × 3 (RAID 5)
	LAN	1000Base-T

表 2 評価パラメータ

データサイズ S [GB]	409.5
圧縮率 C [%]	94.78
選択率 P [%]	0.12
ISN ディスク総数 N_d	1, 2, 4, 6, 8
ISN プロセッサ総数 N_p	2, 4, 8, 12, 16
データ転送速度 V_n [MB/sec]	80
伸張および照合処理速度 V_d [MB/sec]	233
圧縮ブロック読出し速度 V_r [MB/sec]	250
検索結果書込み速度 V_w [MB/sec]	24
起動処理時間 T_0 [sec]	0.60

4.2. 評価パラメータ

評価システムにおいて、式(1)のパラメータは表 2 に示す通りである。ログデータとしてはセキュリティログ(607, 516, 100 件)を使用した。

4.3. 検索の並列処理性能

ISNのプロセッサ総数 N_p に対する検索速度(ログデータサイズ/検索処理時間)を図 3に示す。 $N_p=16$ まで N_p にほぼ比例した検索性能向上が確認された。

4.4. 性能モデルによる検証

検索処理時間 T の式(1)による予測性能と実測値を表 3 に示す。概ね一致しており、妥当な性能が得られている。

5. おわりに

SISA による追記型ログデータベースを商用システム上に実装し、検索ノード追加による検索性能のスケラビリティを確認した。

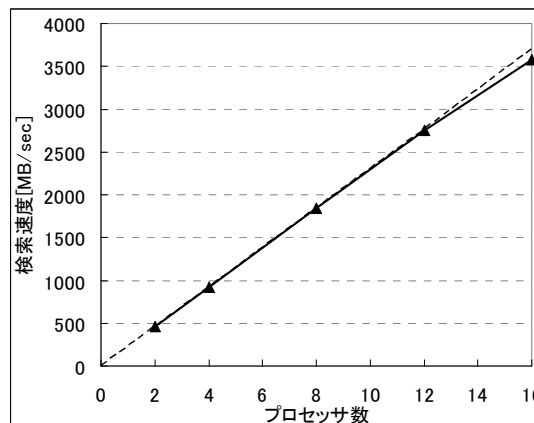


図 3 検索速度

表 3 検索処理時間の予測性能と実測値

N_d	N_p	T_d [秒]	T_n [秒]	T_r [秒]	T_w [秒]	T 予測 [秒]	T 実測 [秒]
1	2	902	6.3	94.8	22.9	902	908
8	16	113	6.3	11.8	22.9	113	117

参考文献

- [1] 金子, 他: スケ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評価, 情報処理学会全国大会第 64 回, 4ZA-5 (2002)
- [2] 清水, 他: スケ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, 情報処理学会全国大会第 68 回, 5D-1 (2006)
- [3] 中村, 他: 大規模ログデータベースの実現, 情報処理学会全国大会第 68 回, 1D-2 (2006)
- [4] 竹内, 他: 大規模ログデータベースの評価, 情報処理学会全国大会第 68 回, 1D-1 (2006)
- [5] 平井, 他: 統合ログ管理技術と履歴追跡型ログデータベース, 三菱電機技報, Vol. 83, No. 12, pp. 15-18 (2009)