

# 大規模シミュレーションにおけるインコア型ストレージの概念の提案

尾西 翔平<sup>†</sup> 橋川 伸吾<sup>††</sup> 鈴木 敬久<sup>†</sup>

## 1. はじめに

近年, HPC (High Performance Computing) 技術の発展に伴い, 大規模な数値解析が行われるようになってきた. これにより, 大規模データのストレージデバイスへの入出力がボトルネックとなっている. 上記のボトルネックを解消するために様々な研究<sup>1)2)</sup>が行われており, 文献<sup>1)</sup>は, データベース分野のインメモリ型データベースに関する研究である. この方式では, 従来のファイルシステムにおいてボトルネックであったディスクアクセスが生じないため, 高速なデータ入出力が可能である. しかし, 大規模数値シミュレーションデータ入出力を目的として, この方式を用いた研究はほとんど行われていない.

本研究の目的は, ファイルシステムに依存しないデータ入出力の新たな概念として, インコア型ストレージを提案することである. その基礎検討として, インコア型ストレージへのデータ入出力性能の評価を行った. 更に, 数値シミュレーション用入出力 API (Application Programming Interface) の設計を行った.

## 2. インコア型ストレージの概念

図 1 は, 本研究で提案するインコア型ストレージの概念図である. 本提案方式では, ファイルシステムに依存しないデータ入出力を実現するために, MRAM (Magnetoresistive Random Access Memory)<sup>3)</sup>を用いる. この MRAM は現在メインメモリとして用いられている DRAM と同等の高速アクセスが可能であり, 加えて不揮発性を有している. この MRAM を用いることで, 保存データをデータそのものとして保存することが可能となる. また, ネットワークを介しているため CPU や GPU (Graphics Processing Unit) など様々なプロセッサを搭載した複数のノードからデータ入出力を行うことが可能となる.

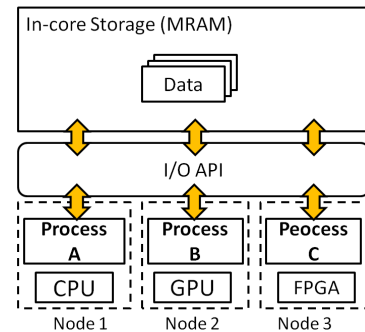


図 1 インコア型ストレージの概念図

## 3. インコア型ストレージの実装

図 2 は, 今回実装したインコア型ストレージのシステム図である.

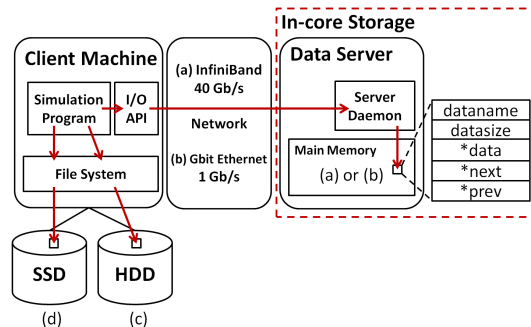


図 2 実装したインコア型ストレージのシステム図

現在 MRAM をシステムに実装することは困難であるため, 今回はインコア型ストレージとして DRAM を内蔵したデータサーバ 1 台を用いて, 概念の基礎的部分に関して予備的な実装を行った. クライアントマシン内で行った数値シミュレーションの出力データは, 入出力 API を用いてデータサーバ内のメインメモリ上に構造体として保存される. また, 本予備的な実装では Gbit Ethernet を用いた場合と, より広帯域な InfiniBand を用いた場合の 2 つの条件下でインコ

<sup>†</sup> 首都大学東京  
Tokyo Metropolitan University  
<sup>††</sup> 株式会社 エルザジャパン  
ELSA JAPAN Inc.

ア型ストレージへネットワーク接続し入出力を行う。

本研究では、表 1、表 2 に示す環境下で、1 GB ～ 10 GB の配列データを、(a) InfiniBand を介してインコア型ストレージへ入出力した場合、(b) Gbit Ethernet を介してインコア型ストレージへ入出力した場合、(c) ローカルな HDD へ入出力した場合、(d) ローカルな SSD へ入出力した場合の各入出力時間を 10 回ずつ測定し、入出力データサイズと平均入出力時間を元にスループット性能を求めた。

表 1 測定環境 (クライアントマシン)

CPU	AMD Opteron Processor 6172
MEMORY	DDR3-1600 12.8 GB/s
HDD	SATA 6 Gb/s
SSD	SATA 6 Gb/s
OS	CentOS 5.7
Network Interface	InfiniBand or Gbit Ethernet

表 2 測定環境 (データサーバ)

CPU	Intel Xeon CPU X5670
MEMORY	DDR3-1600 12.8 GB/s
OS	CentOS 5.7
Network Interface	InfiniBand or Gbit Ethernet

#### 4. 測定結果

図 3、図 4 に、(a)～(d) のスループット性能を示す。

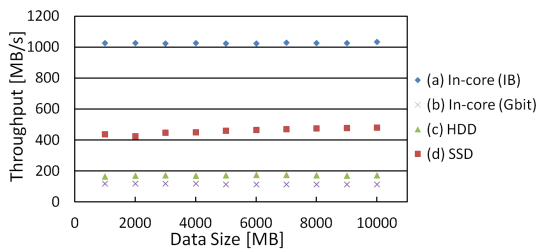


図 3 データ出力性能

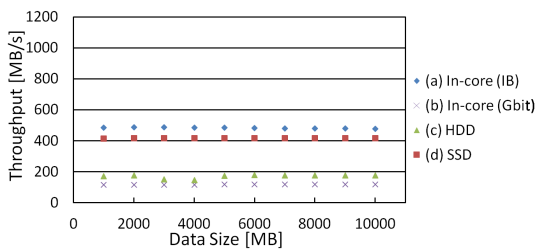


図 4 データ入力性能

図 3 より、データ出力性能に関しては、(a) の場合は約 1.0 GB/s、(b) の場合は約 117 MB/s、(c) の場合は約 190 MB/s、(d) の場合は約 470 MB/s であることが確認された。また図 4 より、データ入力性能に関しては、(a) の場合は約 480 MB/s、(b) の場合約 113 MB/s、(c) の場合は約 180 MB/s、(d) の場合は約 420 MB/s であることが確認された。本提案方式を用いることで、データ出力性能ではローカルな HDD の約 5.2 倍、ローカルな SSD の約 2.1 倍、データ入力性能ではローカルな HDD の約 2.7 倍、ローカルな SSD の約 1.1 倍のスループット性能が得られることが確認できた。

#### 5. 考察

本提案方式で InfiniBand を用いてデータ入出力を行った場合、ローカルなストレージへの入出力性能を上回るスループット性能が得られた。また、Gbit Ethernet を用いた場合では、InfiniBand を用いた場合よりも低いスループット性能となった。このことから、本提案方式の入出力性能はネットワークの通信帯域に大きく依存することが明らかとなった。

#### 6. まとめ

ファイルシステムに依存しないデータ入出力を実現するインコア型ストレージの概念の提案を行った。また、入出力性能の評価を行い、出力性能は約 1.0 GB/s、入力性能は約 480 MB/s のスループット性能であることが確認された。本提案方式で InfiniBand を用いた場合、ローカルなストレージへの入出力性能を上回ることが確認できた。今後は、複数の入出力要求を並列に処理するアルゴリズムを開発する。また、MRAM のハードウェア実装の可能性について検討する。

#### 参考文献

- 1) Lina Lan, *et al*, "A Novelweight Main Memory Database for Telecom Network Performance Management System," *JOURNAL OF NETWORKS*, VOL. 7, No.4, pp.667-674, 2012.
- 2) Arifa Nisar, *et al*, "Delegation-Based I/O Mechanism for High Performance Computing Systems," *IEEE TRANSACTIONS OF PARALLEL AND DISTRIBUTED SYSTEMS*, VOL. 23, NO. 2, FEBRURY 2012.
- 3) 中山 昌彦, *et al*, "垂直磁化方式を用いたスピン注入磁化反転による MRAM スケーラビリティの展望," 電子情報通信学会技術報告, VOL. 108, NO. 346, pp. 37-41, 2008.