

クラウドストレージにおける HDD と SSD の特性比較

建部 大輔[†] 中村 俊介[†] 首藤 一幸[†]

[†]東京工業大学

1 はじめに

1.1 クラウドストレージ

クラウドストレージは、多くのマシン、ペタバイトを超えるデータを扱うことを想定し、スケールできるようにデザインされたデータストアである。スケーラビリティ、ハイパフォーマンス、高可用性などを目的としている。

1.2 HDD(ハードディスクドライブ) と SSD(ソリッドステートドライブ) の性能の違い

HDD には、ディスクの目的のエリアにヘッドを動かす時間、ディスク上のデータが回転してくる時間があるため、データを読み出し際に時間がかかるが、SSD と比べると容量は大きくなっている。また SSD は、フラッシュメモリにデータを記録しているために HDD に比べて高速なデータの読み書きが可能である。

1.3 目的

クラウドストレージである Cassandra における、HDD と SSD でどのくらいの性能差が出るかということを比較、評価する。

2 Cassandra

Cassandra はログベースの書き込み性能重視のクラウドストレージである。Cassandra のストレージは、commit log と memtable と SSTable の3つから構成される。書き込み処理は、まずディスク上の commit log にシーケンシャルに書き込まれ、次にメモリ上の memtable に map 形式で書き込まれる。しかし、memtable の指定容量を超えると、入りきらなくなり古いデータから非同期で順にディスク上の SSTable に内容が書き込まれる。読み出し処理は、まず memtable 上のデータと SSTable 上のデータを複数回読み出しそれらをマージしてデータを返す。

利点として、書き込みはシーケンシャルに行わ

れランダム I/O が発生せず、ライトロックを必要としないため、書き込み性能重視となっており、高速で処理できる。しかし、読み出し時は、データを複数の SSTable から読み出すためディスクランダム I/O が生じてしまうため、性能が損なわれる。HDD と SSD ではクラウドストレージでの読み書き性能にどのくらいの差があるのかを、実験を行って調べる。

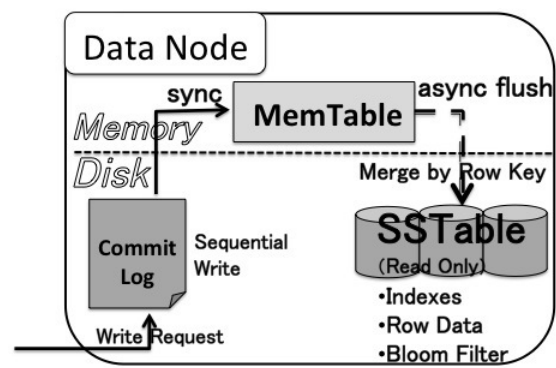


図 1: 書き込みの流れ

3 実験・評価

3.1 YCSB

Yahoo!クラウドサービングベンチマーク(YCSB)は、Yahoo! Research がさまざまなクラウドストレージがを公平に評価することを目的として実装したオープンソースのベンチマークフレームワークであり、実アプリに近いコアワークロードが用意されている。本研究では、これらのワークロードの read と update の比率を変え、実験を行った。

3.2 実験環境

HDD 2 台と SSD 2 台を搭載したサーバで Cassandra を動作させ、YCSB を用いて 1 ノードの性能を測った。

- HDD: Western Digital 社 1TB
- SSD: Crucial 社 128GB
- OS: Linux 2.6.35.6-48.fc14.x86-64

A Comparison Study of HDD and SSD with a Cloud Storage

[†] Daisuke TATEBE

[†] Shunsuke NAKAMURA

[†] Kazuyoshi SHUDO

- Java SE 6 Update 21
- CPU: Intel(R) Xeon(R) 24MHz
- Memory: 32GB
- JVM: 3GB
- MemtableThroughputInMB: 64MB
- MemtableOperationsInMillions: 30 万個

MemtableThroughputInMB は memtable のカラムファミリー単位のデータ容量であり MemtableOperationsInMillions はメモリが保持するカラムファミリー単位のカラム数である。

3.3 実験内容

上記の環境で、1 件 1KB のデータを 1000 万件ロードする。表 1 の 4 つのワークロードに対し、スループットを等しくしたときのそれぞれのレイテンシを測定、またスレッド数を等しくした場合の最大スループットを測定する。

表 1: YCSB ワークロード

Workload	Read	Update	Selection	App. Ex.
Write-Only	0%	100%	Zipfian	Log
Write-Heavy	50%	50%	Zipfian	Session
Read-Heavy	95%	5%	Zipfian	tagging
Read-Only	100%	0%	Zipfian	Cache

4 評価・考察

図 1 は、1 秒あたり 3,000 回のクエリを発行したとき読み出しと書き込みの遅延を、ワークロードごとに示したグラフである。もっとも遅延が減少した Read-only のワークロードでは HDD の遅延時間の 59.3 % の遅延時間まで減少した。

図 2 は、クライアント数を 100 として負荷をかけたときのスループットをワークロードごとに示したグラフである。SSD のスループットは、最大で HDD のスループットの 1.44 倍になった。

書き込み処理だけの write-only ワークロードでもスループットの差は出るが、遅延時間の差はあまりでない。これは Cassandra が書き込み性能重視のストレージであるため、遅延時間自体が小さくなっており、書き込み性能はあまり変わっていない。

読み出し処理だけの read-only ワークロードでのスループットは、書き込み処理と比べる

と小さい。しかし、読み出し処理自体の遅延時間が大きいので、スループットの差が小さくても、遅延時間に読み出し性能は大きく改善された。

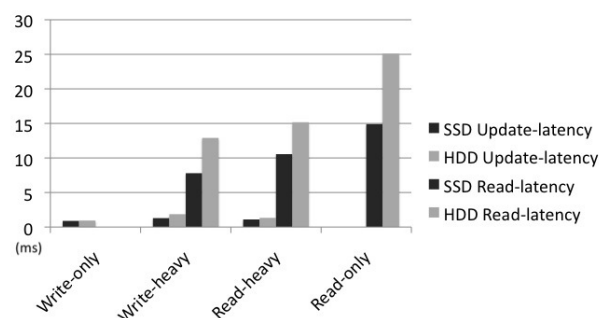


図 2: 遅延

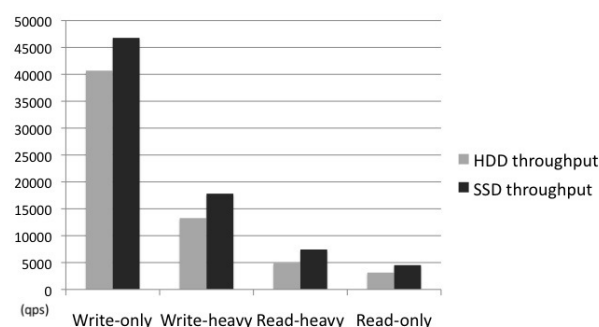


図 3: スループット

5 まとめ

本研究でクラウドストレージにおいて、HDD と SSD の性能を比較し、読み出し性能に大きな差があることがわかった。

参考文献

- [1] Cooper B. F., Silberstein Adam, Tam Erwin, Ramakrishnan Raghu and Sears Russel "Benchmarking Cloud Serving Systems with YCSB," SOCC'10, 2010
- [2] Chang Fay, Dean Jeffrey, Ghemawat Sanjay, Hsieh Wilson C., Walach Deborah A., Burrows Mike, Chandra Tushar, Fikes Andrew and Gruber Robert E. "Bigtable: A Distributed Storage System for Structured Data," OSDI'06, 2006
- [3] Lakshman, Avinash and Malik, Prashant "Cassandra - A Decentralized Structured Storage System," LADIS'09, 2009