

大規模センサネットワークに好適なデータベースの比較検討

末松 直樹[†] 岩田 寛生^{††} 宮崎 敏明^{††}

会津大学コンピュータ理工学部[†]

会津大学大学院コンピュータ理工学研究科^{††}

1. 序論

センサネットワークは物理環境の観測に幅広く普及しており、観測データを蓄積するために様々なデータベース(DB)が使われている。しかし、センサネットワークの規模が大きくなると、観測データも膨大となり、それらをうまく扱えるDBが必要となる。近年、いわゆるビッグデータを扱うための分散DBが幾つか提案され、実用化されている。本稿では、複数種類のリレーショナルDBと分散DBを、大規模センサネットワークで使用するために必要な側面から比較し、分散DBの1つであるCassandra[1]が有用であることを示す。

2. センサネットワークにおけるデータベース

本稿では、日本各地に配置した複数のセンサネットワークを、インターネットを介して統合し、1つのシステムとした大規模センサネットワークを想定する。本システムにおいて、ユーザは、過去のデータはDBから、リアルタイムデータはDBを介さずセンサネットワークからインターネットへのゲートウェイを経由して直接取得出来るものとする。

センサネットワークは短い時間間隔で小容量のデータを生成するため、DBは小容量のデータを高速で書き込み出来る必要がある。また、想定環境において、リアルタイムデータはDBを介さないため、読み出しに関しての厳密な一貫性を必要ない。また、ネットワーク分断があった場合の読み取り可能なデータを考慮すると、データはセンサネットワーク毎にローカルなDBに保存されていることが望ましい。

文献[2]では、サーバ1台とクライアント1台を用いたセンサネットワーク用のDBの比較を行っているが、複数台を用いた分散DB環境の比較は実施していない。本稿では、DBの基本性能を1台のマシンで検証比較した後、データを保有するマシンを複数台設けた同一環境で分散DBの比較を行う。

3. 読み書き速度の計測比較

ここでは、広く普及しているリレーショナルDB(RDB)であるMySQL[3]およびPostgreSQL[4]、分散DBであるCassandra、Hadoop[5]分散ファイルシステムを用いたHBase[6](以下HBase/Hadoop)の、計4種類のDBに関して、読み書きの性能を比較する。用いたデータは、センサネットワークで生成されるデータ1件が小容量であるという特徴を考慮し、表1の通りとする。まず、4種類のDBを1台のマシンで性能比較し、その後分散DBをマシンの台数を増やして性能比較する。

表1 扱うデータ構造

主キー	Integer (4 byte)
値	Text (固定半角数字 200 文字 = 200 byte)

3.1 マシン1台を用いたDBの比較

1台のマシンにサーバとクライアントを起動し、10,000件、100,000件、1,000,000件のデータの読み書きを行い、要した時間を測定した。MySQLについてはディスク上の実データへのアクセス処理を司るストレージエンジンの相異による2種類(MyISAMとInnoDB)、Cassandraについては同期・非同期読み書きについてそれぞれ計測した。計測に用いたハードウェアおよびソフトウェア環境を表2および表3に示す。また、書き込み計測結果を表4に、読み出し計測結果を表5に示す。

表2 ハードウェア環境

CPU	Intel® Core™ i5-2400 CPU @ 3.10GHz
メモリ	6 GB
ディスク	サイズ: 500 GB 回転速度: 7200 RPM

表3 ソフトウェア環境

OS	CentOS release 6.4 (Final) x86_64
JDK	OpenJDK 1.7.0_45
MySQL	サーバ: 5.1.69 ストレージエンジン: MyISAM, InnoDB クライアント: connector-java-5.1.26
PostgreSQL	Server: 8.4.18 Client: postgresql-jdbc-8.4.701
Cassandra	Server: 2.0.x Client: cassandra-driver-core 1.0.3
Hadoop	2.0.0-cdh4.4.0
HBase	Server: 0.94.6-cdh4.4.0 Client: hbase-0.94.6

表4 マシン1台での書き込み計測結果(ms)

データ件数	10,000	100,000	1,000,000
MySQL MyISAM	786	5,666	49,080
MySQL InnoDB	723	6,920	74,890
PostgreSQL	920	6,003	22,013
Cassandra 同期	4,176	12,407	94,683
Cassandra 非同期	2,136	6,270	30,703
HBase/Hadoop	<i>5,503</i>	<i>22,227</i>	<i>190,803</i>

表5 マシン1台での読み出し計測結果(ms)

データ件数	10,000	100,000	1,000,000
MySQL MyISAM	763	5,747	54,387
MySQL InnoDB	683	5,900	52,943
PostgreSQL	490	3,863	37,506
Cassandra 同期	<i>3,637</i>	<i>13,270</i>	<i>216,740</i>
Cassandra 非同期	1,423	3,530	60,377
HBase/Hadoop	1,900	8,903	89,260

Comparisons of Database Systems Suitable for Large-scale Sensor Networks

[†]Naoki Suematsu, ^{††}Hiroki Iwata, ^{††}Toshiaki Miyazaki

[†]School of Computer Science and Engineering, The University of Aizu

^{††}Graduate School of Computer Science and Engineering, The University of Aizu

書き込みにおいては、RDB である MySQL と PostgreSQL が速い結果となったが、データ数 100,000 件以上においては Cassandra の非同期書き込みと大差はない。一方、HBase/Hadoop は全ての件数において最も遅い結果となった。

3.2 複数マシンを用いた分散 DB 比較

3.1 章では 1 台のマシンでサーバとクライアントの両方を起動し計測した。しかし、分散 DB である Cassandra と HBase/Hadoop に関しては、本来想定する使用環境ではない。ここでは 3.1 章で扱った 2 種類の分散 DB を、データの保存されるマシンの台数を 4 台に統一し、読み書き速度を計測し比較する。Cassandra は 3.1 章と同様、同期読み書き、非同期読み書きの 2 種類の計測を行う。本計測では、クライアントとして表 2 のマシンを、サーバとして表 6 のマシンを用いた。書き込み計測結果を表 7 に、読み出し計測結果を表 8 に示す。

表 6 サーバ用ハードウェア環境

CPU	Intel® Core™ i7-3770 CPU @ 3.40GHz
メモリ	8 GB
ディスク	サイズ: 1 TB 回転速度: 7200 RPM

表 7 分散 DB の比較：書き込み計測結果 (ms)

データ件数	10,000	100,000	1,000,000
HBase/Hadoop	16,019	130,240	1,294,799
Cassandra 同期	21,626	112,699	873,779
Cassandra 非同期	749	2,614	9,649

表 8 分散 DB の比較：読み出し計測結果 (ms)

データ件数	10,000	100,000	1,000,000
HBase/Hadoop	5,605	50,498	513,129
Cassandra 同期	18,324	113,808	1,163,344
Cassandra 非同期	521	2,883	18,130

書き込みに関しては、Cassandra の非同期書き込みが他を凌駕し最速の結果となった。また HBase/Hadoop と Cassandra の同期書き込みでは、データ数 10,000 件では Cassandra の同期書き込みの方が遅いが、それより大きなデータ件数では HBase/Hadoop の方が遅いという結果になった。

読み出しに関しても、Cassandra の非同期読み出しが他を凌駕し最速の結果となった。しかし、読み出しは書き込みとは異なり、HBase/Hadoop の方が Cassandra の同期書き込みより速いという結果になった。

総じて、3.1 章での 1 台のマシン上で行った比較結果と比べて、優劣の点では変わらない結果となった。

4. 考察

3 章では DB の読み書きの速度を比較した。本稿で想定する大規模センサネットワークでは、統合するセンサネットワーク数が増え、センサ数が膨大になることが予測されるため、読み出し速度より書き込み速度がより重要となる。

CAP 定理 [7] によれば、分散 DB では、一貫性・可用性・分断耐性の 3 つの性質の内、多くても 2 つまでしか保証

出来ない。それによれば、今回比較した DB は、表 9 のように分類できる。

表 9 CAP 定理による分類

	一貫性	可用性	分断耐性
RDB	○	○	×
Cassandra	×	○	○
HBase/Hadoop	○	×	○

CAP 定理の観点から見ると今回想定する大規模センサネットワークでは、書き込まれたデータは”過去の情報”として早くとも数秒後に参照されれば充分であり、ミリ秒単位の厳密な一貫性は必要としない。また、各センサネットワークは同等な扱いになることが望ましいため、単一障害点 (Single Point of Failure) を構造的に持つ DB の使用は避け、可用性は確保したい。さらに、ネットワーク分断があった時にシステムがダウンすることは許されないため、分断耐性も求められる。Cassandra の一貫性を犠牲にした書き込み速度および読み込み速度性能は、前章の性能比較で示したとおり高速である。以上の考察より、本稿で想定する大規模センサネットワークには Cassandra が好適だと言える。

5. 結論

複数のセンサネットワークを統合して構成される大規模センサネットワークの DB には、Cassandra が好適であることを、求められる機能およびアクセス速度の計測結果から示した。

ネットワーク分断があった場合の読み取り可能なデータを考慮すると、各センサネットワークから取得した観測データは、当該センサネットワークをインターネットに接続するために設けるゲートウェイの近傍に位置する DB に保存されることが望ましい。しかし、今回用いた Cassandra の設定では、必ずしも、近傍の DB にデータが保存されるとは限らない。今後は、上記を考慮し、具体的なデータ構造を用いて Cassandra を用いた DB システムの構築と評価を行っていく。

謝辞 本研究の一部は、総務省戦略的情報通信研究開発推進制度 (SCOPE No. 121802001) および科研費 (No. 23500095) の支援を受けて実施したものである。

参考文献

- [1] Lakshman Avinash, Malik Prashant, "Cassandra: a decentralized structured storage system," SIGOPS Operating Systems Review, vol. 44, no. 2. pp. 35-40. Apr. 2010
- [2] Jan Sipke van der Veen, Bram van der Waaij, Robert J. Meijer, "Sensor Data Storage Performance: SQL or NoSQL, Physical or Virtual," IEEE 5th International Conference on Cloud Computing (CLOUD2012), pp.431-438, June 2012
- [3] MySQL, <http://www.mysql.com>
- [4] PostgreSQL, <http://www.postgresql.org>
- [5] T. White, "Hadoop: The Definitive Guide," O'Reilly Media, Inc., 2009
- [6] George Lars, "HBase: The Definitive Guide," 1st edition, O'Reilly Media, September 2011, ISBN 9781449396107
- [7] Nancy Lynch, Seth Gilbert, "Brewer's conjecture and the feasibility of consistent, available, partition-tolerant web services", ACM SIGACT News, vol.3, no.2, pp. 51-59, 2002