

ICDE2016 Industrial and Applications 4-1:

A Column Store Engine for Real-Time Streaming Analytics

Alexander Skidanov (MemSQL), Anders Papito (MemSQL), Adam Prout (MemSQL)

塩井 隆円 (同志社大, 文化情報学研究科)

* Some figures are copied from ICDE 2016 proceedings.

目的

目的：リアルタイムOLTP/OLAP

- MemSQL：列指向型DBMS
 - セグメント：列データの集合（1セグメントに数千行等）
 - ▶ セグメント単位で列データをソート，圧縮してディスクに格納
 - ▶ ソートを工夫することでOLTP/OLAP高速化を図る

この論文の解く問題 (1/4)

- Internal Sorting

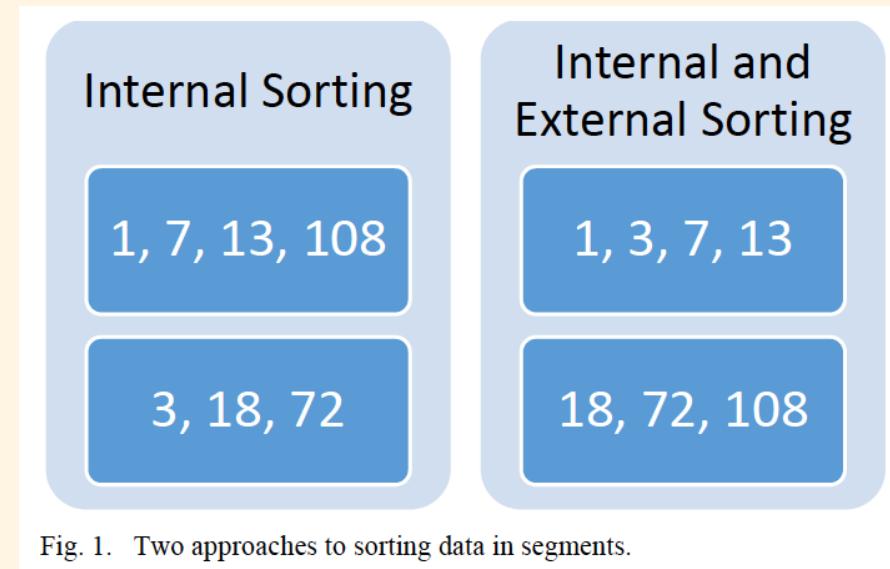
- セグメント内のデータのソート

- External Sorting

- セグメント間のデータのソート

- Internal Sortingされていることが前提

- ▶ [1,13]→[18,108]

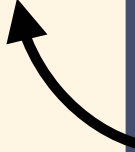


この論文の解く問題 (2/4)

- Segment Elimination

- ▶ diskに格納されているデータ (ソート済み) :

[0, 150], [150, 250], [250, 400], [400, 500]



```
SELECT COUNT(col)
FROM table
WHERE col > 100 AND col < 300 ;
```

◎[400, 500] の問合せに不要なデータは
あらかじめ省いて処理できる

この論文の解く問題 (3/4)

- sorted run algorithm (segment elimination ↗)
 - ▶ あるセグメント内の最大値が、他のセグメントの最小値よりも小さければ集合 S (segment group) に含む ($n \log n$)

- internal sortが出来ている理由は列指向型DBMSがメモリ上で1segmentをソートし満杯のsegmentをディスクに格納するから



本論文の解く問題 (4/4)

- sorted run algorithm in Insert and Update

- ▶ 新しいセグメントを作ってsegment groupの
小さいセグメントから順にマージソート

- 途中でデータのreadがあってもソートはバックグラウンド
で行われているのでリアルタイムにOLAPに対応できる
(ディスク使用率は二倍)

- ○ディスク使用効率が悪いが、速度は上がる

- ×ディスク使用効率が良いが、速度は下がる

評価(1/2)

- TPC-H (100GB, SF=100)

- OLAPのみを評価

- DB B :
格納する全ての列
データをソート
してしまうため,
MemSQLよりも
速い

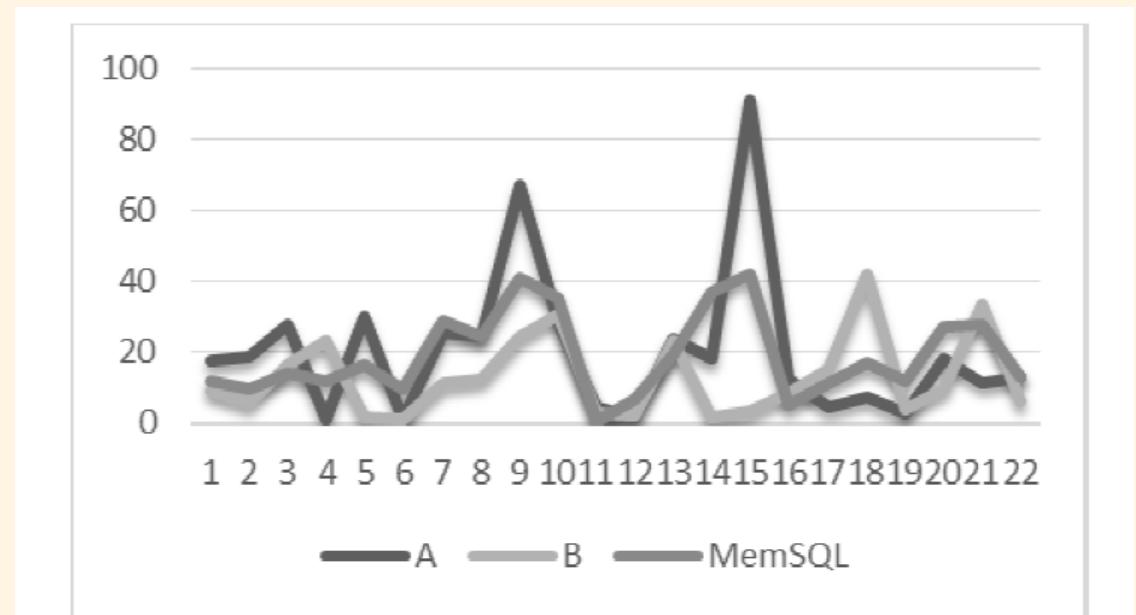


Fig. 4. TPC-H comparison without concurrent write queries.

評価(2/2)

- OLTPを繰り返しながらOLAPを評価

DB A

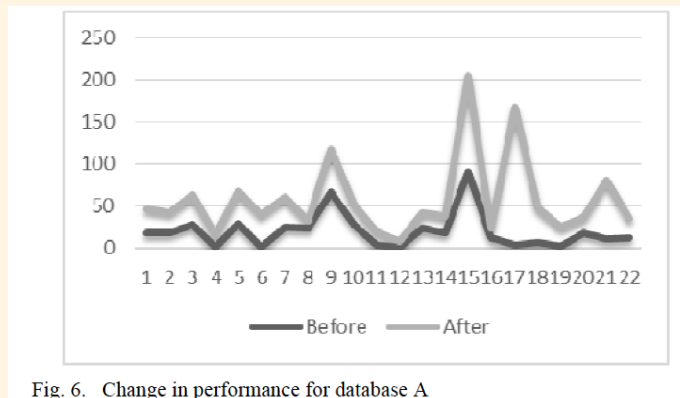


Fig. 6. Change in performance for database A

MemSQL

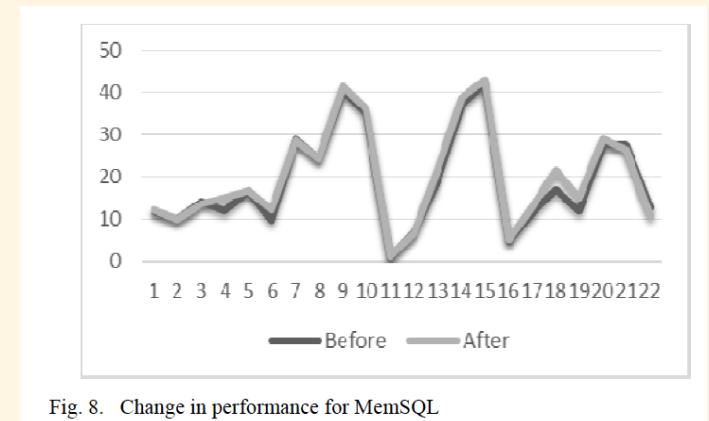


Fig. 8. Change in performance for MemSQL

DB B

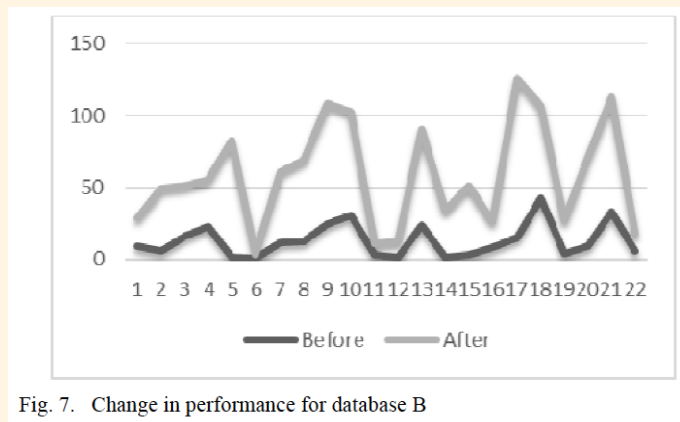


Fig. 7. Change in performance for database B

