

# TPC-H performance measure

Keisuke Suzuki

2013 年 1 月 9 日

## 1 実験環境

- CPU : Xeon X7560 @ 2.27GHz x4
- Memory : 64GB
- DBMS : PostgreSQL 9.2
- RAID0 : iohdrive x8 (chunk size = 64KB)
- 各テーブルの primary key 上に B-tree index を構築
- Scale Factor = 100
- shared buffer = 8GB
- 各クエリの実行時の状況を iostat と mpstat で 1 秒おきに監視

## 2 Query 1 by index scan on l\_shipdate

### 2.1 microbenchmark

microbench で計測した IO の性能が read-ahead の有無や O\_DIRECT の有無によって、乱れる場合があった。乱れの原因となっている部分を確認するため、以下のパラメタを変化させて挙動を観察する。(iosize = 8KB)

パラメタ

- read-ahead: 0 or 2048
- O\_DIRECT: 有 or 無
- iohdrive 単体 or raid0

### 2.1.1 random read microbenchmark 結果

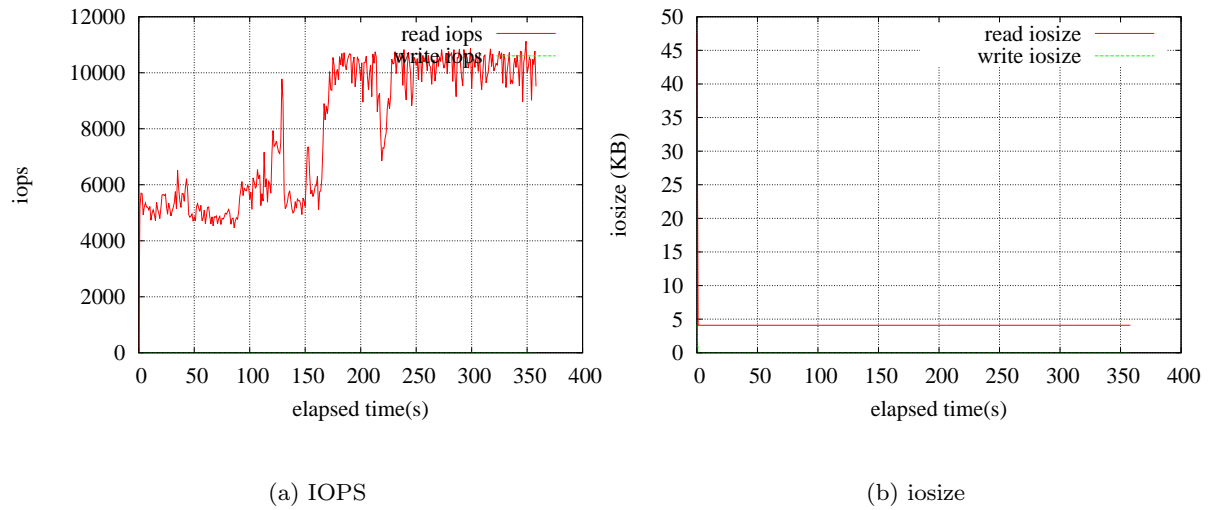


図 1 read-ahead = 0, w/o O\_DIRECT, raid0

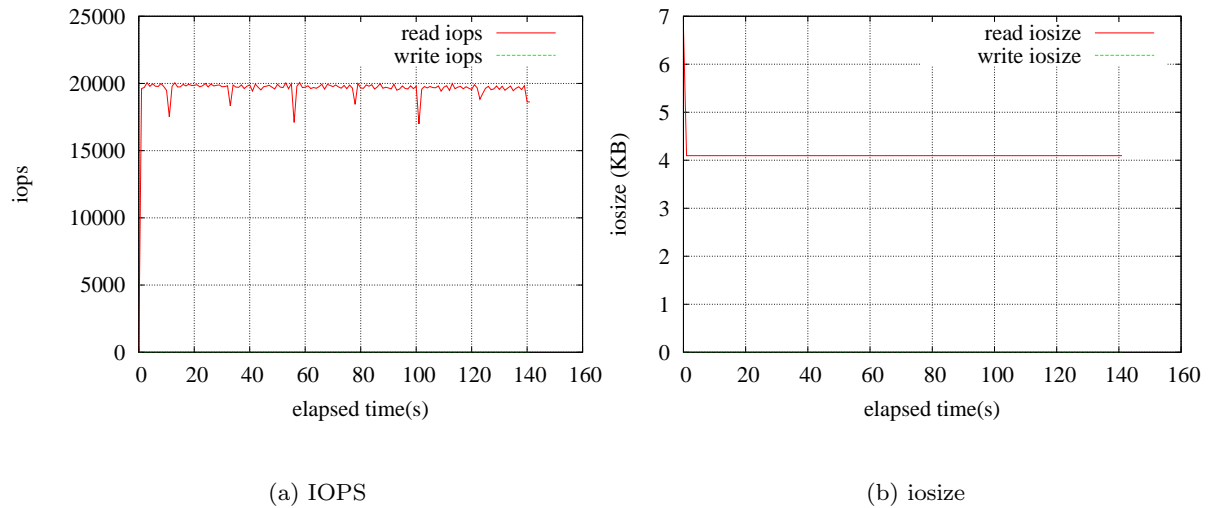
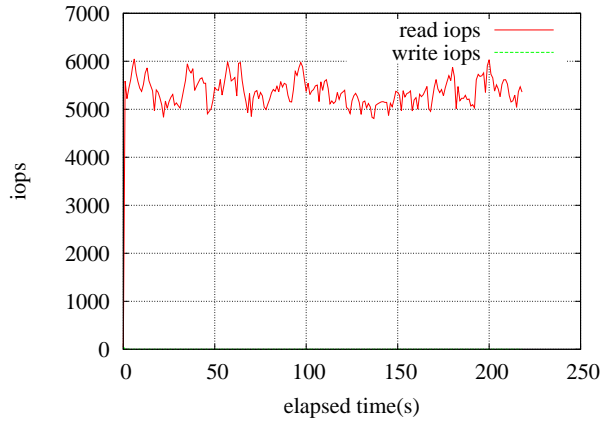
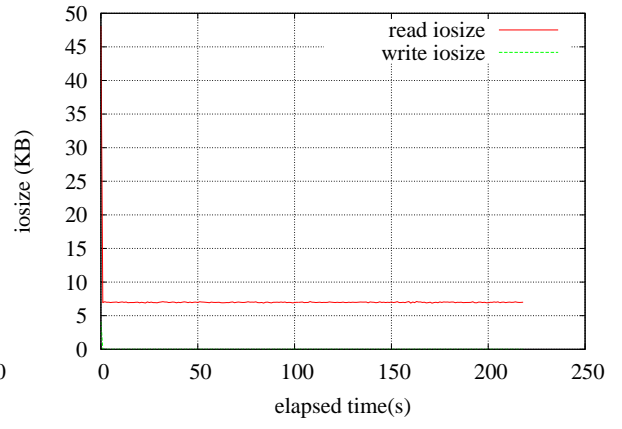


図 2 read-ahead = 0, w/o O\_DIRECT, standalone

図 1(a) では、性能が大きく乱れており、同じ計測を繰り返すと毎回異なるグラフになっていて、実行時間も大きく変化していた。iodrive 単体の上での計測ではこのような結果にはならなかったので、software raid の部分に原因があると想定される。

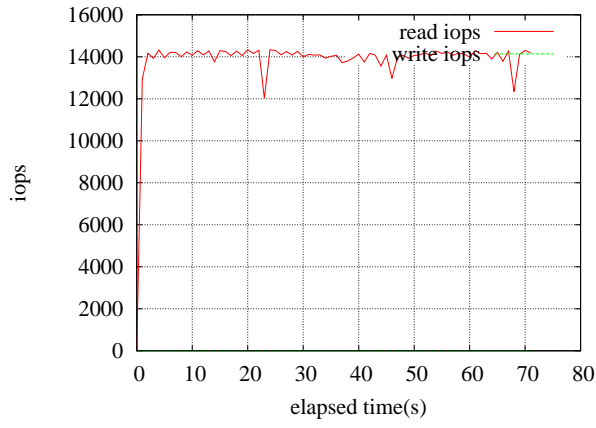


(a) IOPS

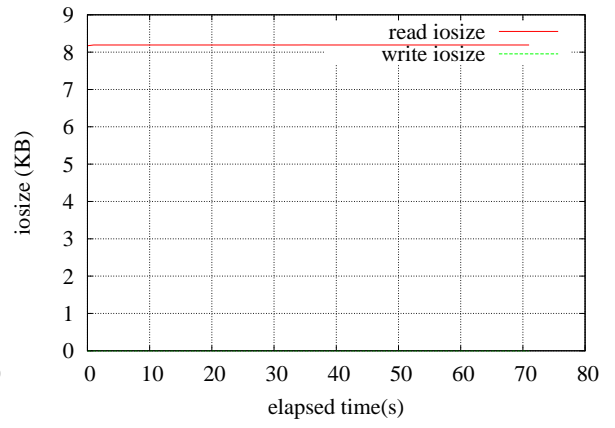


(b) iosize

Figure 3 read-ahead = 0, w/ O\_DIRECT, raid0

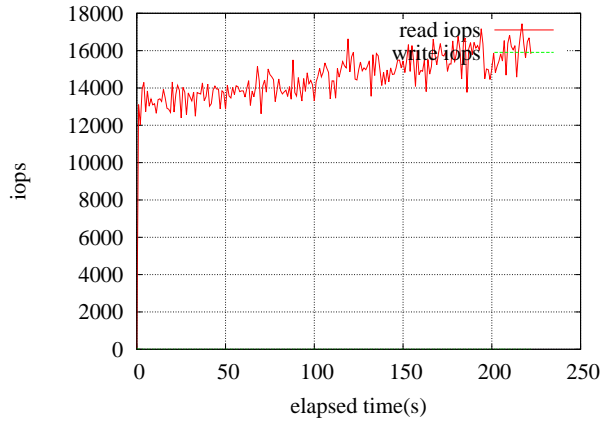


(a) IOPS

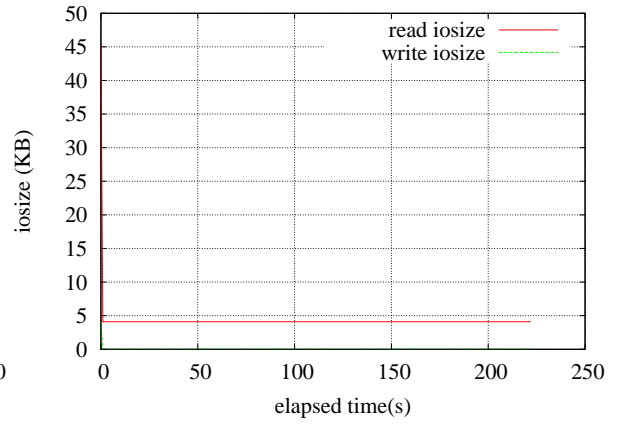


(b) iosize

Figure 4 read-ahead = 0, w/ O\_DIRECT, standalone

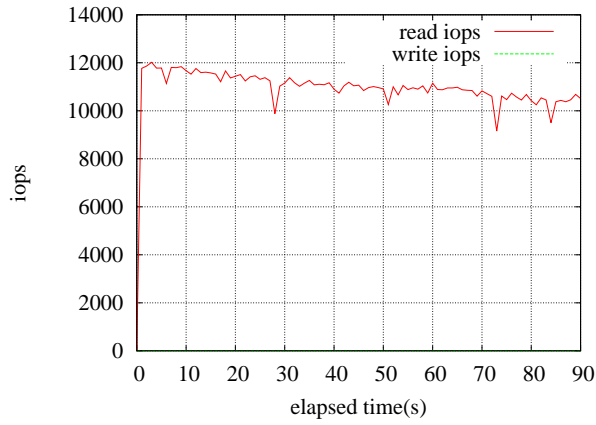


(a) IOPS

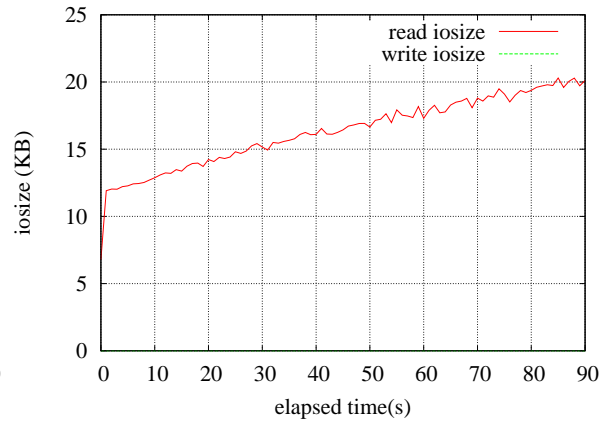


(b) iosize

图 5 read-ahead = 2048, w/o O\_DIRECT, raid0

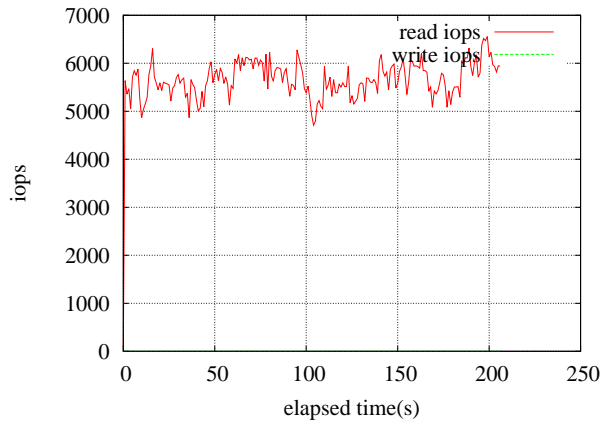


(a) IOPS

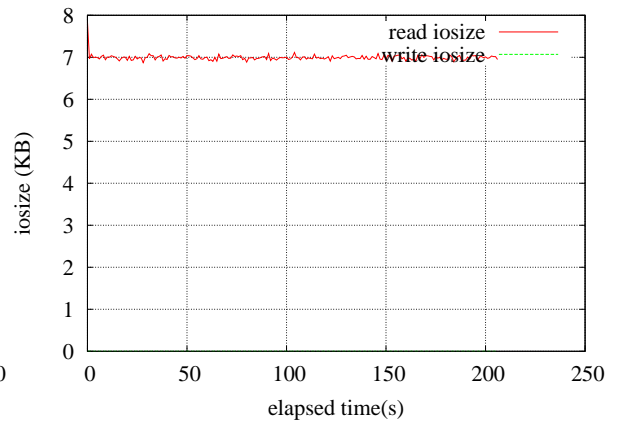


(b) iosize

图 6 read-ahead = 256, w/o O\_DIRECT, standalone

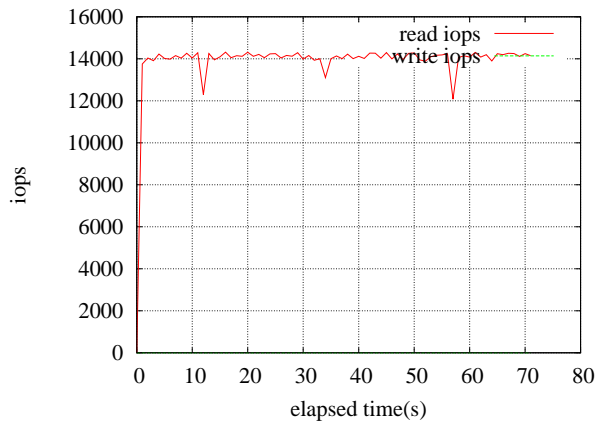


(a) IOPS

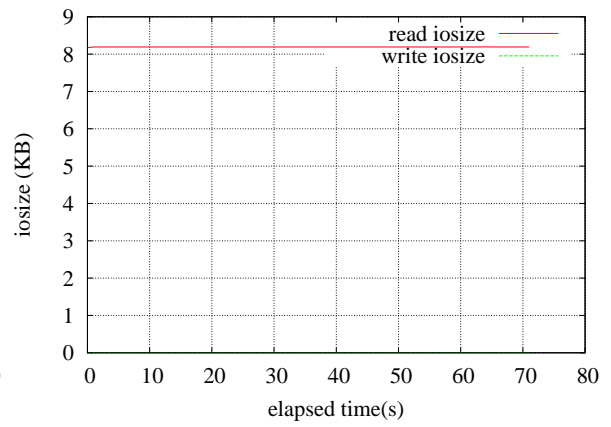


(b) iosize

図 7 read-ahead = 2048, w/ O\_DIRECT, raid0



(a) IOPS



(b) iosize

図 8 read-ahead = 256, w/ O\_DIRECT, standalone

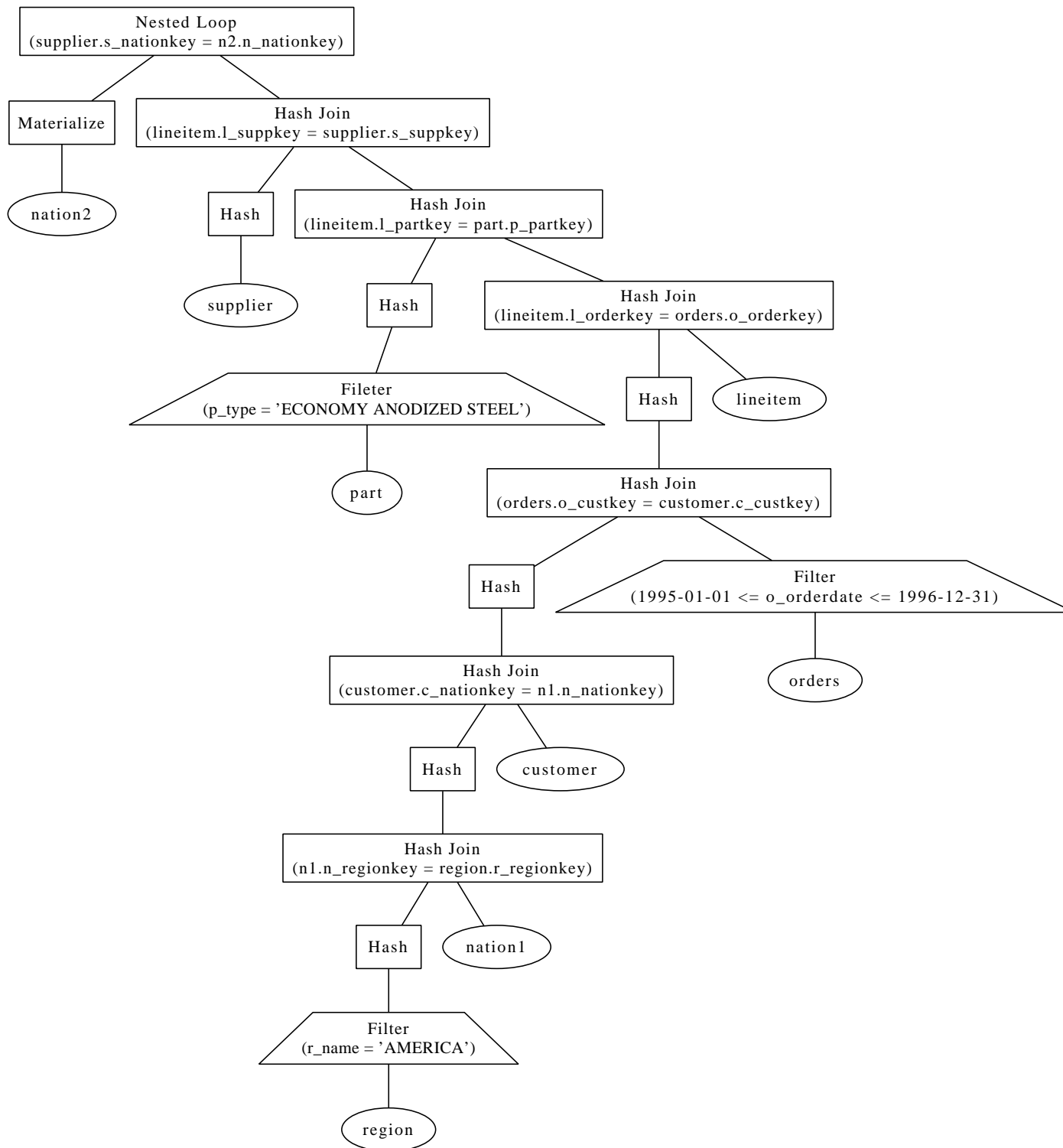
raid の方では複数の LU にまたがるアドレスへのアクセスが含まれる場合があるため、iosize が 8k よりは小さくなっている。

### 3 Query 8

簡単の為、Query8 のうちの IO がメインとなる部分のみを実行する。

#### 3.1 Query and Execution Plan

```
select
extract(year from o_orderdate) as o_year,
l_extendedprice * (1 - l_discount) as volume,
n2.n_name as nation
from
part, supplier, lineitem, orders,
customer, nation n1, nation n2, region
where
p_partkey = l_partkey
and s_suppkey = l_suppkey
and l_orderkey = o_orderkey
and o_custkey = c_custkey
and c_nationkey = n1.n_nationkey
and n1.n_regionkey = r_regionkey
and r_name = 'AMERICA'
and s_nationkey = n2.n_nationkey
and o_orderdate between date '1995-01-01' and date '1996-12-31'
and p_type = 'ECONOMY ANODIZED STEEL'
```



☒ 9 Query 8 execution plan

### 3.2 結果

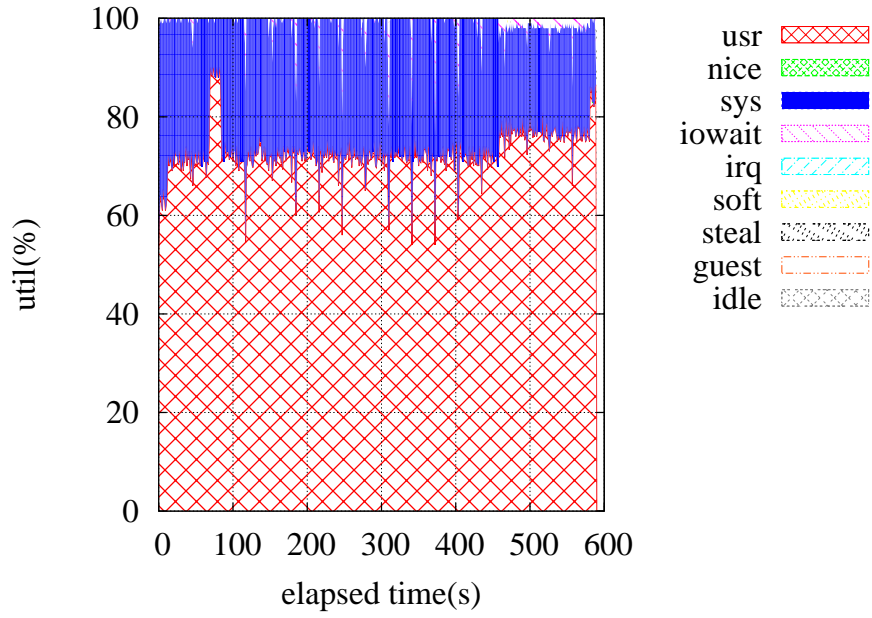


図 10 cpu utilization

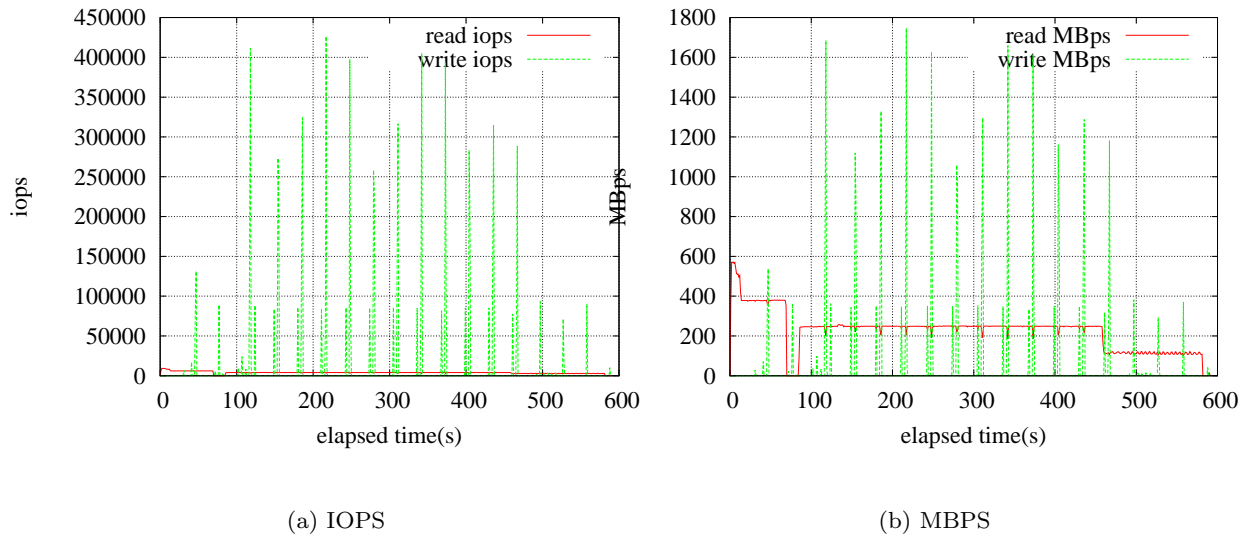


図 11 IO spec