le4 DB 課題6

1029305381 佐藤新太

索引に関する考察

発行する頻度が高いクエリを考える。

部屋の検索

```
select * from rooms
where host_id = 'bu7s8inb0m7rpebtrp7g' and id > 'bu7s8hnb0m7rpebtqpkg' and
price between 5000 and 10000
limit 20;
```

```
+-----------------------------------------------------------------------------
-----------------------------------------------------------------------------
--------------+
|QUERY PLAN
|
+-----------------------------------------------------------------------------
-----------------------------------------------------------------------------
--------------+
|Limit  (cost=1000.00..22905.31 rows=1 width=56) (actual
time=6.404..160.490 rows=4 loops=1)
|
|  ->  Gather  (cost=1000.00..22905.31 rows=1 width=56) (actual
time=6.375..160.091 rows=4 loops=1)
|
|        Workers Planned: 2
|
|        Workers Launched: 2
|
|        ->  Parallel Seq Scan on rooms  (cost=0.00..21905.21 rows=1
width=56) (actual time=90.122..140.633 rows=1 loops=3)
|
|              Filter: (((id)::text > 'bu7s8hnb0m7rpebtqpkg'::text) AND
(price >= 5000) AND (price <= 10000) AND ((host_id)::text =
'bu7s8inb0m7rpebtrp7g'::text))|
|              Rows Removed by Filter: 372367
|
|Planning Time: 0.434 ms
|
|Execution Time: 160.954 ms
|
+-----------------------------------------------------------------------------
-----------------------------------------------------------------------------
```

```
--------------+
```

WHERE句で160msかかっているので、以下のインデックスを貼る。idはprimary keyなので自動でindexが貼られていると判断した。

```
create index on rooms (price);
create index on rooms (host_id);
```

```
+-----------------------------------------------------------------------------
-----------------------------------------------+
|QUERY PLAN
|
+-----------------------------------------------------------------------------
-----------------------------------------------+
|Limit  (cost=0.43..8.56 rows=1 width=56) (actual time=0.289..0.521 rows=4
loops=1)                                       |
|  ->  Index Scan using rooms_host_id_idx on rooms  (cost=0.43..8.56
rows=1 width=56) (actual time=0.259..0.385 rows=4 loops=1)|
|        Index Cond: ((host_id)::text = 'bu7s8inb0m7rpebtrp7g'::text)
|
|        Filter: (((id)::text > 'bu7s8hnb0m7rpebtqpkg'::text) AND (price
>= 5000) AND (price <= 10000))                  |
|        Rows Removed by Filter: 18
|
|Planning Time: 0.685 ms
|
|Execution Time: 0.776 ms
|
+-----------------------------------------------------------------------------
-----------------------------------------------+
```

以上のように、Index Scanが実行され、実行時間が160.954 msから0.776msと大幅に性能が向上した。

特定の部屋の情報を取得

```
select * from rooms where id = 'bu7s8hnb0m7rpebtqpkg';
```

```
+-----------------------------------------------------------------------------
---------------------------------------+
|QUERY PLAN
|
+-----------------------------------------------------------------------------
---------------------------------------+
|Index Scan using rooms_pkey on rooms  (cost=0.43..8.45 rows=1 width=56)
```

```
(actual time=0.094..0.117 rows=1 loops=1)|
|  Index Cond: ((id)::text = 'bu7s8hnb0m7rpebtqpkg'::text)
|
|Planning Time: 0.218 ms
|
|Execution Time: 0.211 ms
|
+-------------------------------------------------------------------------------
-------------------------------------+
```

index scanが行われているので、新たなindexは貼らない。

あるゲストの予約を確認

```
select res.id, check_in, check_out, guest_id, room_id, r.name from
reservations as res
inner join rooms r on r.id = res.room_id
where guest_id = 'bu7s6l7b0m7rpm50qv4g' and check_in > '2020-10-21
00:00:00' and check_out < '2020-10-23 00:00:00'
limit 20;
```

```
+-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
-----------------------------------------------------------------+
|QUERY PLAN
|
+-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
-----------------------------------------------------------------+
|Limit  (cost=1000.43..7638.17 rows=5 width=89) (actual time=0.813..63.460
rows=2 loops=1)
|
|  ->  Gather  (cost=1000.43..7638.17 rows=5 width=89) (actual
time=0.790..63.366 rows=2 loops=1)
|
|        Workers Planned: 2
|
|        Workers Launched: 2
|
|        ->  Nested Loop  (cost=0.43..6637.67 rows=2 width=89) (actual
time=26.455..44.986 rows=1 loops=3)
|
|              ->  Parallel Seq Scan on reservations res
(cost=0.00..6620.78 rows=2 width=79) (actual time=26.396..44.793 rows=1
loops=3)
|
|                    Filter: ((check_in > '2020-10-21
00:00:00+00'::timestamp with time zone) AND (check_out < '2020-10-23
00:00:00+00'::timestamp with time zone) AND ((guest_id)::text =
```

```
'bu7s6l7b0m7rpm50qv4g'::text))|
|                    Rows Removed by Filter: 102252
|
|              ->  Index Scan using rooms_pkey on rooms r
(cost=0.43..8.45 rows=1 width=31) (actual time=0.045..0.050 rows=1
loops=2)
|
|                    Index Cond: ((id)::text = (res.room_id)::text)
|
|Planning Time: 0.732 ms
|
|Execution Time: 63.596 ms
|
+-------------------------------------------------------------------------
-------------------------------------------------------------------------
-----------------------------------------------------------------+
```

where句で63msかかっているので、以下のindexを作成した。

```
create index on reservations (guest_id);
create index on reservations (check_in);
create index on reservations (check_out);
```

```
+-------------------------------------------------------------------------
-------------------------------------------------------------------------
---------+
|QUERY PLAN
|
+-------------------------------------------------------------------------
-------------------------------------------------------------------------
---------+
|Limit  (cost=0.85..50.81 rows=5 width=89) (actual time=0.299..0.645
rows=2 loops=1)
|
|  ->  Nested Loop  (cost=0.85..50.81 rows=5 width=89) (actual
time=0.220..0.381 rows=2 loops=1)
|
|        ->  Index Scan using reservations_guest_id_idx on reservations
res  (cost=0.42..8.59 rows=5 width=79) (actual time=0.154..0.201 rows=2
loops=1)     |
|              Index Cond: ((guest_id)::text =
'bu7s6l7b0m7rpm50qv4g'::text)
|
|              Filter: ((check_in > '2020-10-21 00:00:00+00'::timestamp
with time zone) AND (check_out < '2020-10-23 00:00:00+00'::timestamp with
time zone))|
|        ->  Index Scan using rooms_pkey on rooms r  (cost=0.43..8.45
rows=1 width=31) (actual time=0.026..0.031 rows=1 loops=2)
|
```

```
|              Index Cond: ((id)::text = (res.room_id)::text)
|
|Planning Time: 1.628 ms
|
|Execution Time: 0.754 ms
|
+-------------------------------------------------------------------------
-------------------------------------------------------------------------
---------+
```

以上のように、Index Scanが実行され、実行時間が63.596 msから0.754 msと大幅に性能が向上した。

あるホストが登録している部屋を確認

```
select id, name, price, host_id from rooms
where host_id = 'bu7s8jnb0m7rpebtt0p0' and id > 'bu7s8jnb0m7rpebtt120'
limit 20;
```

```
+-------------------------------------------------------------------------
-------------------------------------------------------+
|QUERY PLAN
|
+-------------------------------------------------------------------------
-------------------------------------------------------+
|Limit  (cost=0.43..8.53 rows=5 width=56) (actual time=0.082..0.445
rows=11 loops=1)                                                       |
|  ->  Index Scan using rooms_host_id_idx on rooms  (cost=0.43..8.53
rows=5 width=56) (actual time=0.060..0.194 rows=11 loops=1)|
|        Index Cond: ((host_id)::text = 'bu7s8jnb0m7rpebtt0p0'::text)
|
|        Filter: ((id)::text > 'bu7s8jnb0m7rpebtt120'::text)
|
|        Rows Removed by Filter: 18
|
|Planning Time: 0.301 ms
|
|Execution Time: 0.695 ms
|
+-------------------------------------------------------------------------
-------------------------------------------------------+
```

すでにhost_idにはindexを貼っているのでそれほど実行時間はかかっていない。

あるホストの予約一覧を取得

```
select res.id, check_in, check_out, guest_id, room_id, r.name from rooms r
    inner join reservations res on r.id = res.room_id
```

```
where host_id = 'bu7t3qvb0m7rpf64tc0g' and check_in > '2020-10-21
00:00:00' and check_out < '2020-10-23 00:00:00'
limit 20;
```

```
+-------------------------------------------------------------------------
-------------------------------------------------------------------------
---------------+
|QUERY PLAN
|
+-------------------------------------------------------------------------
-------------------------------------------------------------------------
---------------+
|Limit  (cost=1008.58..7642.13 rows=1 width=89) (actual
time=1030.963..1416.056 rows=13 loops=1)
|
|  ->  Gather  (cost=1008.58..7642.13 rows=1 width=89) (actual
time=1030.930..1415.823 rows=13 loops=1)
|
|        Workers Planned: 2
|
|        Workers Launched: 2
|
|        ->  Hash Join  (cost=8.58..6642.03 rows=1 width=89) (actual
time=1218.829..1345.927 rows=4 loops=3)
|
|              Hash Cond: ((res.room_id)::text = (r.id)::text)
|
|              ->  Parallel Seq Scan on reservations res
(cost=0.00..6301.24 rows=126553 width=79) (actual time=0.095..677.838
rows=102186 loops=3)               |
|                    Filter: ((check_in > '2020-10-21
00:00:00+00'::timestamp with time zone) AND (check_out < '2020-10-23
00:00:00+00'::timestamp with time zone))|
|                    Rows Removed by Filter: 67
|
|              ->  Hash  (cost=8.52..8.52 rows=5 width=31) (actual
time=0.201..0.219 rows=3 loops=3)
|
|                    Buckets: 1024  Batches: 1  Memory Usage: 9kB
|
|                    ->  Index Scan using rooms_host_id_idx on rooms r
(cost=0.43..8.52 rows=5 width=31) (actual time=0.104..0.143 rows=3
loops=3)               |
|                          Index Cond: ((host_id)::text =
'bu7t3qvb0m7rpf64tc0g'::text)
|
|Planning Time: 1.612 ms
|
|Execution Time: 1416.949 ms
|
+-------------------------------------------------------------------------
```

```
-------------------------------------------------------------------------------
--------------+
```

room_idで結合する際に時間がかかっているため、room_idにindexを貼った。

```
create index on reservations (room_id);
```

```
+-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
---------+
|QUERY PLAN
|
+-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
---------+
|Limit  (cost=0.85..51.02 rows=1 width=89) (actual time=0.203..1.333
rows=13 loops=1)
|
|  ->  Nested Loop  (cost=0.85..51.02 rows=1 width=89) (actual
time=0.183..0.970 rows=13 loops=1)
|
|        ->  Index Scan using rooms_host_id_idx on rooms r
(cost=0.43..8.52 rows=5 width=31) (actual time=0.035..0.072 rows=3
loops=1)                        |
|              Index Cond: ((host_id)::text =
'bu7t3qvb0m7rpf64tc0g'::text)
|
|        ->  Index Scan using reservations_room_id_idx on reservations res
(cost=0.42..8.48 rows=2 width=79) (actual time=0.040..0.150 rows=4
loops=3)      |
|              Index Cond: ((room_id)::text = (r.id)::text)
|
|              Filter: ((check_in > '2020-10-21 00:00:00+00'::timestamp
with time zone) AND (check_out < '2020-10-23 00:00:00+00'::timestamp with
time zone))|
|Planning Time: 1.162 ms
|
|Execution Time: 1.596 ms
|
+-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
---------+
```

以上のように、Hash Joinが実行されなくなり、実行時間が1416.949 msから1.596 msと大幅に性能が向上した。

tags: le4