- Создание запроса без первичного ключа

```
CREATE TABLE wo_pk (id UInt32, name String) ENGINE = MergeTree() ORDER BY tuple();
```

```
Main. :) CREATE TABLE wo_pk (id UInt32, name String) ENGINE = MergeTree() ORDER BY tuple();

CREATE TABLE wo_pk
(
    `id` UInt32,
    `name` String
)
ENGINE = MergeTree
ORDER BY tuple()

Query id: 813e0276-9e52-47aa-bac2-61175b09db73

Ok.

0 rows in set. Elapsed: 0.018 sec.
```

- Создание запроса с первичным ключом

```
CREATE TABLE w_pk (id UInt32, name String) ENGINE = MergeTree() ORDER BY id;
```

```
Main. :) CREATE TABLE w_pk (id UInt32, name String) ENGINE = MergeTree() ORDER BY id;

CREATE TABLE w_pk
(
    `id` UInt32,
    `name` String
)
ENGINE = MergeTree
ORDER BY id

Query id: 0b18d5ce-b0e2-47ff-b007-eee5471d8e74

Ok.

0 rows in set. Elapsed: 0.047 sec.
```

- Заполнение данными таблицы без первичного ключа

```
INSERT INTO wo_pk SELECT number, concat('Name', toString(number)) FROM numbers(1000000);
```

```
Main. :) INSERT INTO wo_pk SELECT number, concat('Name', toString(number)) FROM numbers(1000000);

INSERT INTO wo_pk SELECT
    number,
    concat('Name', toString(number))
FROM numbers(1000000)

Query id: f44ae39e-bb92-436d-8e92-4b3ae6f6932f

Ok.

0 rows in set. Elapsed: 0.115 sec. Processed 1.00 million rows, 8.00 MB (8.69 million rows/s., 69.55 MB/s.)
Peak memory usage: 55.64 MiB.
```

- Заполнение данными таблицы с первичным ключом

INSERT INTO w_pk SELECT number, concat('Name', toString(number)) FROM numbers(1000000);

```
Main. :) INSERT INTO w_pk SELECT number, concat('Name', toString(number)) FROM numbers(1000000);

INSERT INTO w_pk SELECT
    number,
    concat('Name', toString(number))
FROM numbers(1000000)

Query id: 37019ab5-f07d-40fb-b254-63b2ab8612ff

Ok.

0 rows in set. Elapsed: 0.095 sec. Processed 1.00 million rows, 8.00 MB (10.57 million rows/s., 84.55 MB/s.)
Peak memory usage: 55.64 MiB.
```

- Выполнение запросов с условием по полю «id»

SELECT * FROM wo_pk WHERE id = 500000;

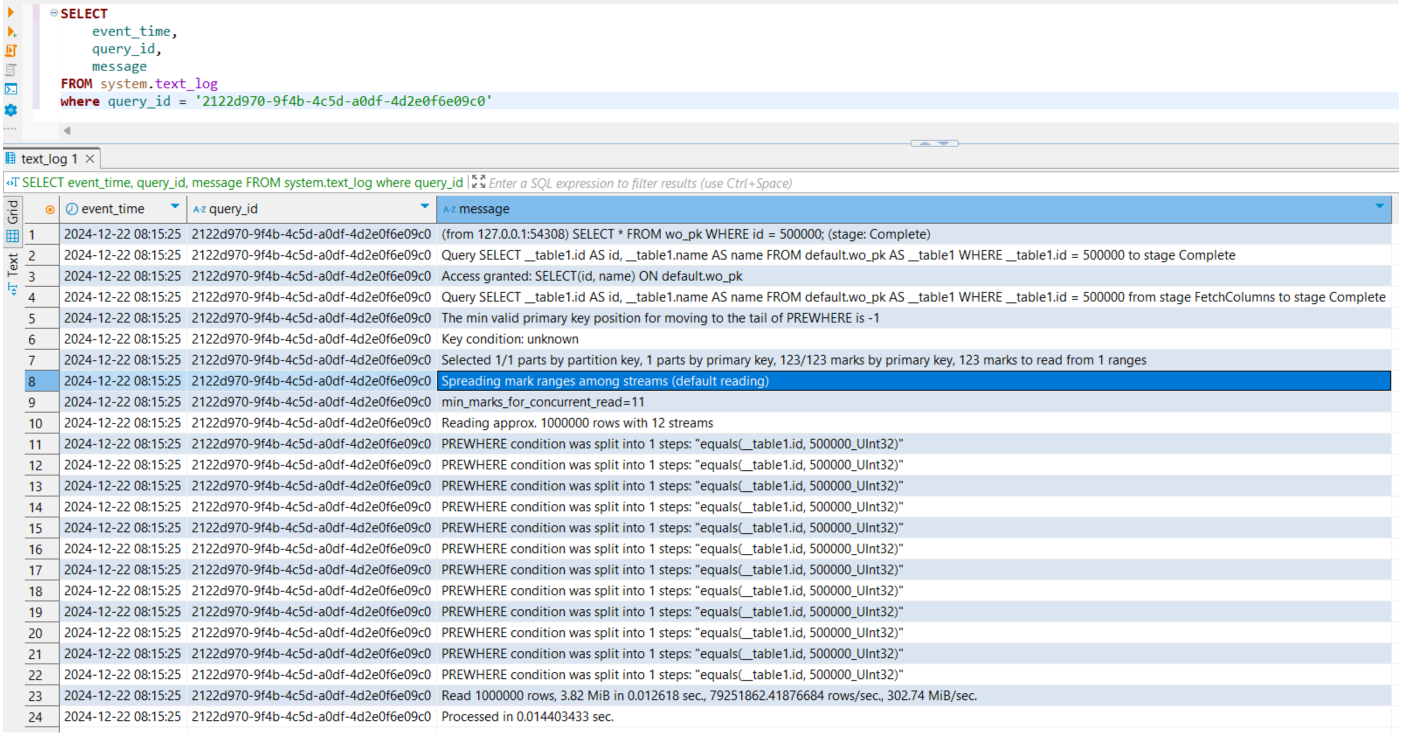
```
Main. :) SELECT * FROM wo_pk WHERE id = 500000;

SELECT *
FROM wo_pk
WHERE id = 500000

Query id: 9044d178-4eee-4fb2-99cb-bf039903a8a7

    ┌──────id─┬─name──────┐
1.  │  500000 │ Name500000 │
    └─────────┴────────────┘

1 row in set. Elapsed: 0.034 sec. Processed 1.00 million rows, 4.01 MB (29.83 million rows/s., 119.47 MB/s.)
Peak memory usage: 64.01 KiB.
```

SELECT * FROM w_pk WHERE id = 500000;

```
Main. :) SELECT * FROM w_pk WHERE id = 500000;

SELECT *
FROM w_pk
WHERE id = 500000

Query id: f0e227cf-4af0-46a2-8e5b-67e8578cfa55

    ┌──────id─┬─name──────┐
1.  │  500000 │ Name500000 │
    └─────────┴────────────┘

1 row in set. Elapsed: 0.009 sec. Processed 8.19 thousand rows, 38.26 KB (889.30 thousand rows/s., 4.15 MB/s.)
Peak memory usage: 7.15 KiB.
```

- Анализ информации из логов для таблицы без первичного ключа

```
SELECT
    event_time,
    query_id,
    message
FROM system.text_log
where query_id = '2122d970-9f4b-4c5d-a0df-4d2e0f6e09c0'
```



```
SELECT
    event_time,
    query_id,
    message
FROM system.text_log
where query_id = '2122d970-9f4b-4c5d-a0df-4d2e0f6e09c0'
```

text_log 1

SELECT event_time, query_id, message FROM system.text_log where query_id

Enter a SQL expression to filter results (use Ctrl+Space)

	event_time	query_id	message
1	2024-12-22 08:15:25	2122d970-9f4b-4c5d-a0df-4d2e0f6e09c0	(from 127.0.0.1:54308) SELECT * FROM wo_pk WHERE id = 500000; (stage: Complete)
2	2024-12-22 08:15:25	2122d970-9f4b-4c5d-a0df-4d2e0f6e09c0	Query SELECT __table1.id AS id, __table1.name AS name FROM default.wo_pk AS __table1 WHERE __table1.id = 500000 to stage Complete
3	2024-12-22 08:15:25	2122d970-9f4b-4c5d-a0df-4d2e0f6e09c0	Access granted: SELECT(id, name) ON default.wo_pk
4	2024-12-22 08:15:25	2122d970-9f4b-4c5d-a0df-4d2e0f6e09c0	Query SELECT __table1.id AS id, __table1.name AS name FROM default.wo_pk AS __table1 WHERE __table1.id = 500000 from stage FetchColumns to stage Complete
5	2024-12-22 08:15:25	2122d970-9f4b-4c5d-a0df-4d2e0f6e09c0	The min valid primary key position for moving to the tail of PREWHERE is -1
6	2024-12-22 08:15:25	2122d970-9f4b-4c5d-a0df-4d2e0f6e09c0	Key condition: unknown
7	2024-12-22 08:15:25	2122d970-9f4b-4c5d-a0df-4d2e0f6e09c0	Selected 1/1 parts by partition key, 1 parts by primary key, 123/123 marks by primary key, 123 marks to read from 1 ranges
8	2024-12-22 08:15:25	2122d970-9f4b-4c5d-a0df-4d2e0f6e09c0	Spreading mark ranges among streams (default reading)
9	2024-12-22 08:15:25	2122d970-9f4b-4c5d-a0df-4d2e0f6e09c0	min_marks_for_concurrent_read=11
10	2024-12-22 08:15:25	2122d970-9f4b-4c5d-a0df-4d2e0f6e09c0	Reading approx. 1000000 rows with 12 streams
11	2024-12-22 08:15:25	2122d970-9f4b-4c5d-a0df-4d2e0f6e09c0	PREWHERE condition was split into 1 steps: "equals(__table1.id, 500000_UInt32)"
12	2024-12-22 08:15:25	2122d970-9f4b-4c5d-a0df-4d2e0f6e09c0	PREWHERE condition was split into 1 steps: "equals(__table1.id, 500000_UInt32)"
13	2024-12-22 08:15:25	2122d970-9f4b-4c5d-a0df-4d2e0f6e09c0	PREWHERE condition was split into 1 steps: "equals(__table1.id, 500000_UInt32)"
14	2024-12-22 08:15:25	2122d970-9f4b-4c5d-a0df-4d2e0f6e09c0	PREWHERE condition was split into 1 steps: "equals(__table1.id, 500000_UInt32)"
15	2024-12-22 08:15:25	2122d970-9f4b-4c5d-a0df-4d2e0f6e09c0	PREWHERE condition was split into 1 steps: "equals(__table1.id, 500000_UInt32)"
16	2024-12-22 08:15:25	2122d970-9f4b-4c5d-a0df-4d2e0f6e09c0	PREWHERE condition was split into 1 steps: "equals(__table1.id, 500000_UInt32)"
17	2024-12-22 08:15:25	2122d970-9f4b-4c5d-a0df-4d2e0f6e09c0	PREWHERE condition was split into 1 steps: "equals(__table1.id, 500000_UInt32)"
18	2024-12-22 08:15:25	2122d970-9f4b-4c5d-a0df-4d2e0f6e09c0	PREWHERE condition was split into 1 steps: "equals(__table1.id, 500000_UInt32)"
19	2024-12-22 08:15:25	2122d970-9f4b-4c5d-a0df-4d2e0f6e09c0	PREWHERE condition was split into 1 steps: "equals(__table1.id, 500000_UInt32)"
20	2024-12-22 08:15:25	2122d970-9f4b-4c5d-a0df-4d2e0f6e09c0	PREWHERE condition was split into 1 steps: "equals(__table1.id, 500000_UInt32)"
21	2024-12-22 08:15:25	2122d970-9f4b-4c5d-a0df-4d2e0f6e09c0	PREWHERE condition was split into 1 steps: "equals(__table1.id, 500000_UInt32)"
22	2024-12-22 08:15:25	2122d970-9f4b-4c5d-a0df-4d2e0f6e09c0	PREWHERE condition was split into 1 steps: "equals(__table1.id, 500000_UInt32)"
23	2024-12-22 08:15:25	2122d970-9f4b-4c5d-a0df-4d2e0f6e09c0	Read 1000000 rows, 3.82 MiB in 0.012618 sec., 79251862.41876684 rows/sec., 302.74 MiB/sec.
24	2024-12-22 08:15:25	2122d970-9f4b-4c5d-a0df-4d2e0f6e09c0	Processed in 0.014403433 sec.

От сюда видно, что идет сканирование всей таблицы:

```
Selected 1/1 parts by partition key, 1 parts by primary key, 123/123 marks by primary
key, 123 marks to read from 1 ranges
```

В 12 потоков:

```
Reading approx. 1000000 rows with 12 streams
PREWHERE condition was split into 1 steps: "equals(__table1.id, 500000_UInt32)"
PREWHERE condition was split into 1 steps: "equals(__table1.id, 500000_UInt32)"
PREWHERE condition was split into 1 steps: "equals(__table1.id, 500000_UInt32)"
PREWHERE condition was split into 1 steps: "equals(__table1.id, 500000_UInt32)"
PREWHERE condition was split into 1 steps: "equals(__table1.id, 500000_UInt32)"
PREWHERE condition was split into 1 steps: "equals(__table1.id, 500000_UInt32)"
PREWHERE condition was split into 1 steps: "equals(__table1.id, 500000_UInt32)"
PREWHERE condition was split into 1 steps: "equals(__table1.id, 500000_UInt32)"
PREWHERE condition was split into 1 steps: "equals(__table1.id, 500000_UInt32)"
PREWHERE condition was split into 1 steps: "equals(__table1.id, 500000_UInt32)"
PREWHERE condition was split into 1 steps: "equals(__table1.id, 500000_UInt32)"
PREWHERE condition was split into 1 steps: "equals(__table1.id, 500000_UInt32)"

Read 1000000 rows, 3.82 MiB in 0.012618 sec., 79251862.41876684 rows/sec., 302.74
MiB/sec.
```

- Анализ информации из логов для таблицы с первичным ключом

```
SELECT
    event_time,
    query_id,
    message
FROM system.text_log
where query_id = '3e92b1f9-3d3f-472d-9243-9d72b11e0443'
```

```
SELECT
    event_time,
    query_id,
    message
FROM system.text_log
where query_id = '3e92b1f9-3d3f-472d-9243-9d72b11e0443'
```

text_log 1

SELECT event_time, query_id, message FROM system.text_log where query_id | Enter a SQL expression to filter results (use Ctrl+Space)

	event_time	query_id	message
1	2024-12-22 08:15:34	3e92b1f9-3d3f-472d-9243-9d72b11e0443	(from 127.0.0.1:54308) SELECT * FROM w_pk WHERE id = 500000; (stage: Complete)
2	2024-12-22 08:15:34	3e92b1f9-3d3f-472d-9243-9d72b11e0443	Query SELECT __table1.id AS id, __table1.name AS name FROM default.w_pk AS __table1 WHERE __table1.id = 500000 to stage Complete
3	2024-12-22 08:15:34	3e92b1f9-3d3f-472d-9243-9d72b11e0443	Access granted: SELECT(id, name) ON default.w_pk
4	2024-12-22 08:15:34	3e92b1f9-3d3f-472d-9243-9d72b11e0443	Query SELECT __table1.id AS id, __table1.name AS name FROM default.w_pk AS __table1 WHERE __table1.id = 500000 from stage FetchColumns to stage Complete
5	2024-12-22 08:15:34	3e92b1f9-3d3f-472d-9243-9d72b11e0443	The min valid primary key position for moving to the tail of PREWHERE is 0
6	2024-12-22 08:15:34	3e92b1f9-3d3f-472d-9243-9d72b11e0443	Key condition: (column 0 in [500000, 500000])
7	2024-12-22 08:15:34	3e92b1f9-3d3f-472d-9243-9d72b11e0443	Running binary search on index range for part all_1_1_0 (124 marks)
8	2024-12-22 08:15:34	3e92b1f9-3d3f-472d-9243-9d72b11e0443	Found (LEFT) boundary mark: 61
9	2024-12-22 08:15:34	3e92b1f9-3d3f-472d-9243-9d72b11e0443	Found (RIGHT) boundary mark: 62
10	2024-12-22 08:15:34	3e92b1f9-3d3f-472d-9243-9d72b11e0443	Found continuous range in 12 steps
11	2024-12-22 08:15:34	3e92b1f9-3d3f-472d-9243-9d72b11e0443	Selected 1/1 parts by partition key, 1 parts by primary key, 1/123 marks by primary key, 1 marks to read from 1 ranges
12	2024-12-22 08:15:34	3e92b1f9-3d3f-472d-9243-9d72b11e0443	Spreading mark ranges among streams (default reading)
13	2024-12-22 08:15:34	3e92b1f9-3d3f-472d-9243-9d72b11e0443	Reading 1 ranges in order from part all_1_1_0, approx. 8192 rows starting from 499712
14	2024-12-22 08:15:34	3e92b1f9-3d3f-472d-9243-9d72b11e0443	PREWHERE condition was split into 1 steps: "equals(__table1.id, 500000_UInt32)"
15	2024-12-22 08:15:34	3e92b1f9-3d3f-472d-9243-9d72b11e0443	Read 8192 rows, 37.36 KiB in 0.007382 sec., 1109726.3614196696 rows/sec., 4.94 MiB/sec.
16	2024-12-22 08:15:34	3e92b1f9-3d3f-472d-9243-9d72b11e0443	Processed in 0.008953605 sec.

Вместо сканирования всей таблицы видно, что идет чтение только одной парты:

Selected 1/1 parts by partition key, 1 parts by primary key, 1/123 marks by primary key, 1 marks to read from 1 ranges

Reading 1 ranges in order from part all_1_1_0, approx. 8192 rows starting from 499712

Read 8192 rows, 37.36 KiB in 0.007382 sec., 1109726.3614196696 rows/sec., 4.94 MiB/sec.

- Анализ плана запроса к таблице без первичного ключа
  EXPLAIN indexes = 1 **SELECT** * **FROM** wo_pk **WHERE** id = 500000;

```
EXPLAIN indexes = 1 SELECT * FROM wo_pk WHERE id = 500000;
```

unknown 1

EXPLAIN indexes = 1 SELECT * FROM wo_pk WHERE id = 500000 | Enter a SQL expre

	explain
1	Expression ((Project names + Projection))
2	Expression
3	ReadFromMergeTree (default.wo_pk)

Видно, что в плане запроса отсутствует использование первичного ключа и происходит сканирование всей таблицы

- Анализ плана запроса к таблице с первичным ключом

```
EXPLAIN indexes = 1 SELECT * FROM w_pk WHERE id = 500000;
```

unknown 1

EXPLAIN indexes = 1 SELECT * FROM w_pk WHERE id = 500000

	explain
1	Expression ((Project names + Projection))
2	Expression
3	ReadFromMergeTree (default.w_pk)
4	Indexes:
5	PrimaryKey
6	Keys:
7	id
8	Condition: (id in [500000, 500000])
9	Parts: 1/1
10	Granules: 1/123

Видно, что в плане запроса используется первичный ключ для быстрого поиска партов по предикату с условием в запросе. При этом из всех гранул и партов читается только одна гранула и одна парта