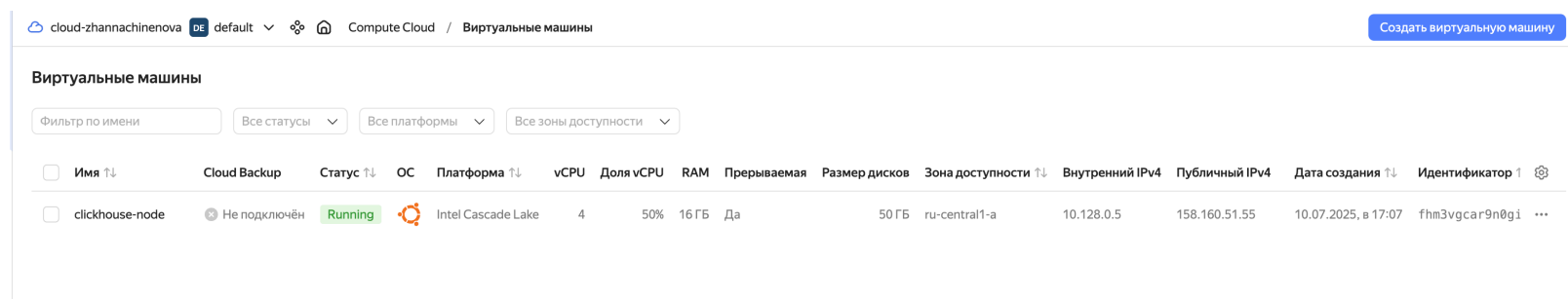


ClickHouse config

1. Создали машину с помощью команды :

```
yc compute instance create --name clickhouse-node --ssh-key /Users/zhannamakrusina/yandex-cloud/key_ssh_id.pub --create-boot-disk image-folder-id=standard-images,image-family=ubuntu-2404-lts,size=50,auto-delete=true --network-interface subnet-name=default-ru-central1-a,nat-ip-version=ipv4 --memory 16G --cores 4 --core-fraction 50 --zone ru-central1-a --preemptible --hostname clickhouse-node
```



The screenshot shows the Yandex Cloud console interface. At the top, there's a header with the user 'cloud-zhannachinenova', a 'default' dropdown, and a 'Compute Cloud / Виртуальные машины' breadcrumb. A blue button 'Создать виртуальную машину' is in the top right. Below the header, there's a section 'Виртуальные машины' with filters: 'Фильтр по имени', 'Все статусы', 'Все платформы', and 'Все зоны доступности'. A table lists the virtual machines. The first row is for 'clickhouse-node', which is 'Running' (green status), has 4 vCPU, 16 GB RAM, and is located in the 'ru-central1-a' zone. The table columns include: Имя, Cloud Backup, Статус, ОС, Платформа, vCPU, Доля vCPU, RAM, Прерываемая, Размер дисков, Зона доступности, Внутренний IPv4, Публичный IPv4, Дата создания, and Идентификатор.

Имя	Cloud Backup	Статус	ОС	Платформа	vCPU	Доля vCPU	RAM	Прерываемая	Размер дисков	Зона доступности	Внутренний IPv4	Публичный IPv4	Дата создания	Идентификатор
clickhouse-node	Не подключён	Running	Ubuntu 24.04 LTS	Intel Cascade Lake	4	50%	16 Гб	Да	50 Гб	ru-central1-a	10.128.0.5	158.160.51.55	10.07.2025, в 17:07	fhm3vgcar9n0gi

2. Выполнена инсталяция clickhouse

```
sudo apt -y install apt-transport-https ca-certificates dirmngr
```

```
sudo apt-key adv --keyserver hkp://keyserver.ubuntu.com:80 --recv  
8919F6BD2B48D754
```

```
sudo echo "deb https://packages.clickhouse.com/deb stable main" | sudo  
tee /etc/apt/sources.list.d/clickhouse.list
```

```
sudo apt update && sudo apt -y install clickhouse-server=25.3.5.42  
clickhouse-client=25.3.5.42 clickhouse-common-static=25.3.5.42
```

```
sudo service clickhouse-server start
```

```
apt install clickhouse-tools
```

3. Загружен датасет trips

CREATE TABLE trips

```
(  
    `trip_id` UInt32,  
    `vendor_id` Enum8('1' = 1, '2' = 2, '3' = 3, '4' = 4, 'CMT' = 5, 'VTS' = 6, 'DDS' = 7,  
    'B02512' = 10, 'B02598' = 11, 'B02617' = 12, 'B02682' = 13, 'B02764' = 14, '' = 15),  
    `pickup_date` Date,  
    `pickup_datetime` DateTime,  
    `dropoff_date` Date,  
    `dropoff_datetime` DateTime,  
    `store_and_fwd_flag` UInt8,  
    `rate_code_id` UInt8,  
    `pickup_longitude` Float64,  
    `pickup_latitude` Float64,  
    `dropoff_longitude` Float64,  
    `dropoff_latitude` Float64,  
    `passenger_count` UInt8,  
    `trip_distance` Float64,  
    `fare_amount` Float32,  
    `extra` Float32,  
    `mta_tax` Float32,  
    `tip_amount` Float32,  
    `tolls_amount` Float32,  
    `ehail_fee` Float32,  
    `improvement_surcharge` Float32,  
    `total_amount` Float32,  
    `payment_type` Enum8('UNK' = 0, 'CSH' = 1, 'CRE' = 2, 'NOC' = 3, 'DIS' = 4),  
    `trip_type` UInt8,
```

```
`pickup` FixedString(25),
`dropoff` FixedString(25),
`cab_type` Enum8('yellow' = 1, 'green' = 2, 'uber' = 3),
`pickup_nyct2010_gid` Int8,
`pickup_ctlabel` Float32,
`pickup_borocode` Int8,
`pickup_ct2010` String,
`pickup_boroct2010` String,
`pickup_cdeligibil` String,
`pickup_ntacode` FixedString(4),
`pickup_ntaname` String,
`pickup_puma` UInt16,
`dropoff_nyct2010_gid` UInt8,
`dropoff_ctlabel` Float32,
`dropoff_borocode` UInt8,
`dropoff_ct2010` String,
`dropoff_boroct2010` String,
`dropoff_cdeligibil` String,
`dropoff_ntacode` FixedString(4),
`dropoff_ntaname` String,
`dropoff_puma` UInt16
)
```

ENGINE = MergeTree

PARTITION BY toYYYYMM(pickup_date)

ORDER BY pickup_datetime

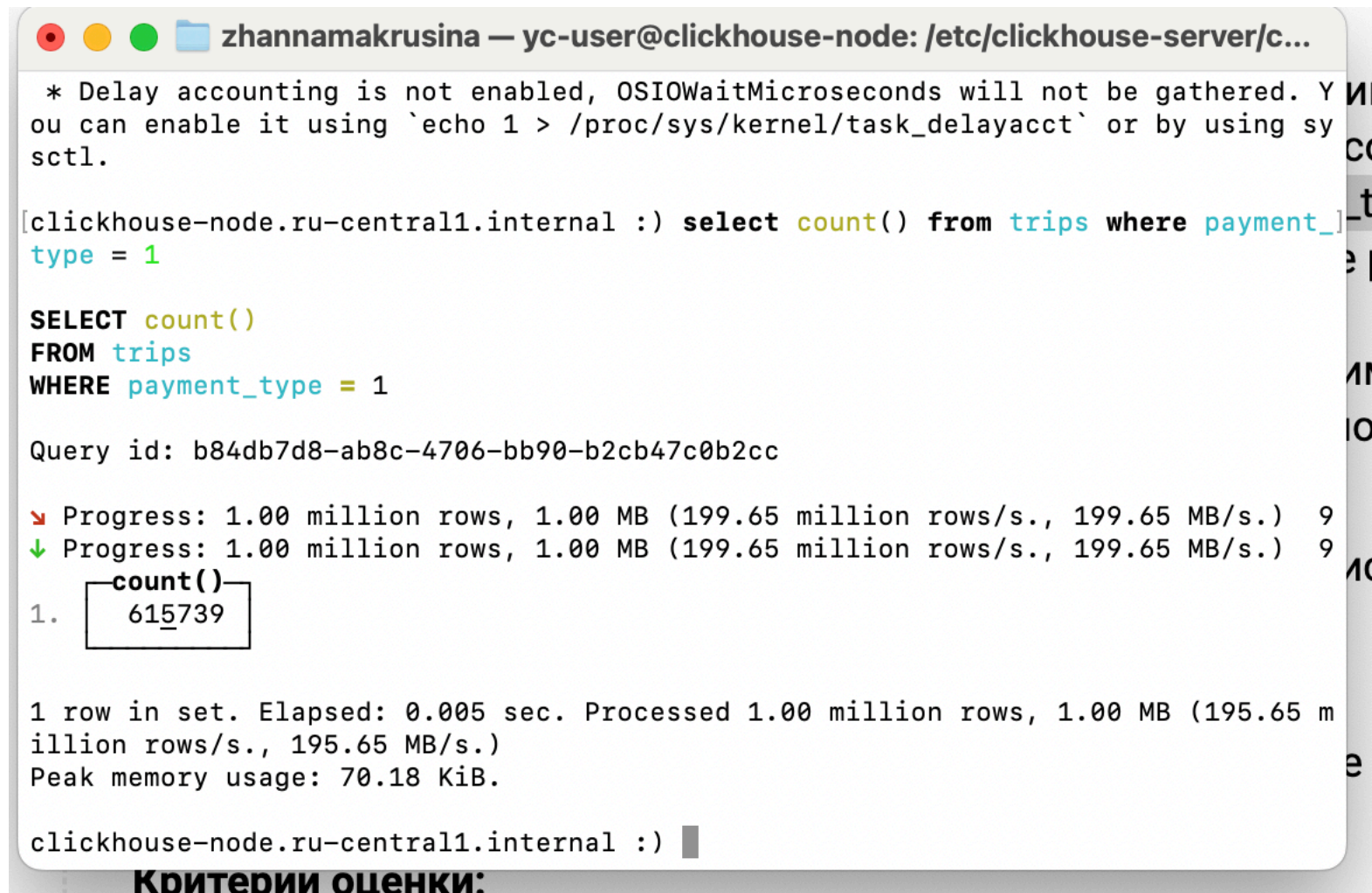
INSERT INTO trips

SELECT *

```
FROM s3('https://datasets-documentation.s3.eu-west-3.amazonaws.com/nyc-taxi/
trips_*.gz', 'TabSeparatedWithNames')
```

```
LIMIT 1000000;
```

4. Результат запроса: `select count() from trips where payment_type = 1.`



The screenshot shows a terminal window titled "zhannamakrusina — yc-user@clickhouse-node: /etc/clickhouse-server/c...". It displays a ClickHouse query and its execution details. The query is `select count() from trips where payment_type = 1`. The results show a single row with the value 615739. The terminal also shows progress information and system metrics.

```
* Delay accounting is not enabled, OSIOWaitMicroseconds will not be gathered. You can enable it using `echo 1 > /proc/sys/kernel/task_delayacct` or by using sysctl.

[clickhouse-node.ru-central1.internal :) select count() from trips where payment_type = 1

SELECT count()
FROM trips
WHERE payment_type = 1

Query id: b84db7d8-ab8c-4706-bb90-b2cb47c0b2cc

Progress: 1.00 million rows, 1.00 MB (199.65 million rows/s., 199.65 MB/s.)
Progress: 1.00 million rows, 1.00 MB (199.65 million rows/s., 199.65 MB/s.)

count()
1. 615739

1 row in set. Elapsed: 0.005 sec. Processed 1.00 million rows, 1.00 MB (195.65 million rows/s., 195.65 MB/s.)
Peak memory usage: 70.18 KiB.

clickhouse-node.ru-central1.internal :) 
```

Критерии оценки:

5. Тестируем производительность командой `echo "SELECT * FROM default.trips LIMIT 10000000 OFFSET 10000000" | clickhouse-benchmark -i 10`

6. Создаем файл `max_query_size.xml`

```
<clickhouse>
  <profiles>
    <default>
      <max_query_size replace="replace">150000</max_query_size>
    </default>
  </profiles>
  <users>
    <default>
      <profile>default</profile>
    </default>
  </users>
</clickhouse>
```

7. Применяем параметр и проверяем что параметр изменился в системной таблице и скорость выполнения чуть снизилась, так как показатель мы снизили по сравнению со значением по умолчанию

8. Увеличили значение `max_query_size` до 15000000 и убрали `limit` из запроса увидели эффект от применения , а потом выключения параметра

```
zhanamakrusina — yc-user@clickhouse-node: ~ — ssh -i ~/yandex-cloud/key_ssh_id yc-user@158.1...

14 of these updates are standard security updates.
To see these additional updates run: apt list --upgradable

Enable ESM Apps to receive additional future security updates.
See https://ubuntu.com/esm or run: sudo pro status

Last login: Thu Jul 10 15:52:57 2025 from 46.22.56.196
yc-user@clickhouse-node:~$ echo "SELECT * FROM default.trips LIMIT 10000000 OFFSET 10000000" | clickhouse-
benchmark -i 10
Loaded 1 queries.

Queries executed: 10.
[localhost:9000, queries: 10, QPS: 7.742, RPS: 7741869.461, MiB/s: 2270.722, result RPS: 0.000, result MiB/
s: 0.000.
[
0%          0.097 sec.
10%         0.111 sec.
20%         0.114 sec.
30%         0.114 sec.
40%         0.115 sec.
50%         0.117 sec.
60%         0.117 sec.
70%         0.122 sec.
80%         0.123 sec.
90%         0.123 sec.
95%         0.133 sec.
99%         0.133 sec.
99.9%       0.133 sec.
99.99%      0.133 sec.
yc-user@clickhouse-node:~$
```

Пункт 5

```
root@clickhouse-node:/etc/clickhouse-server/users.d# ls -l
total 4
-rw-rw-r-- 1 yc-user yc-user 266 Jul 15 17:58 max_query_size.xml
root@clickhouse-node:/etc/clickhouse-server/users.d#
```

Пункт 6

```
root@clickhouse-node:/etc/clickhouse-server/users.d# echo "SELECT * FROM default.trips LIMIT 10000000 OFFSET 10000000" | clickhouse-benchmark -i 10
Loaded 1 queries.

Queries executed: 10.

localhost:9000, queries: 10, QPS: 7.755, RPS: 7754927.695, MiB/s: 2274.552, result RPS: 0.000, result MiB/s: 0.000.
0%          0.105 sec.
10%         0.107 sec.
20%         0.107 sec.
30%         0.109 sec.
40%         0.111 sec.
50%         0.119 sec.
60%         0.119 sec.
70%         0.119 sec.
80%         0.120 sec.
90%         0.123 sec.
95%         0.151 sec.
99%         0.151 sec.
99.9%       0.151 sec.
99.99%      0.151 sec.
```

Пункт 7


```
zhannamakrusina — root@clickhouse-node: /etc/clickhouse-server/users.d — ssh -i ~/yandex-cloud/key_ssh_id yc-user@158.160.1'
localhost:9000, queries: 10, QPS: 1.851, RPS: 1850638.253, MiB/s: 542.800, result RPS: 1850638.253, result MiB/s: 542.800.

0%          0.470 sec.
10%         0.471 sec.
20%         0.472 sec.
30%         0.488 sec.
40%         0.503 sec.
50%         0.510 sec.
60%         0.510 sec.
70%         0.538 sec.
80%         0.600 sec.
90%         0.605 sec.
95%         0.619 sec.
99%         0.619 sec.
99.9%       0.619 sec.
99.99%      0.619 sec.

root@clickhouse-node:/etc/clickhouse-server/users.d# mv max_query_size.xml max_query_size.xml_old
root@clickhouse-node:/etc/clickhouse-server/users.d# echo "SELECT * FROM default.trips " | clickhouse-benchmark -i 10
Loaded 1 queries.

Queries executed: 2.

localhost:9000, queries: 2, QPS: 1.505, RPS: 1504971.262, MiB/s: 441.414, result RPS: 1504971.262, result MiB/s: 441.414.

0%          0.556 sec.
10%         0.556 sec.
20%         0.556 sec.
30%         0.556 sec.
40%         0.556 sec.
50%         0.768 sec.
60%         0.768 sec.
70%         0.768 sec.
80%         0.768 sec.
90%         0.768 sec.
95%         0.768 sec.
99%         0.768 sec.
99.9%       0.768 sec.
99.99%      0.768 sec.
```

Пункт 8