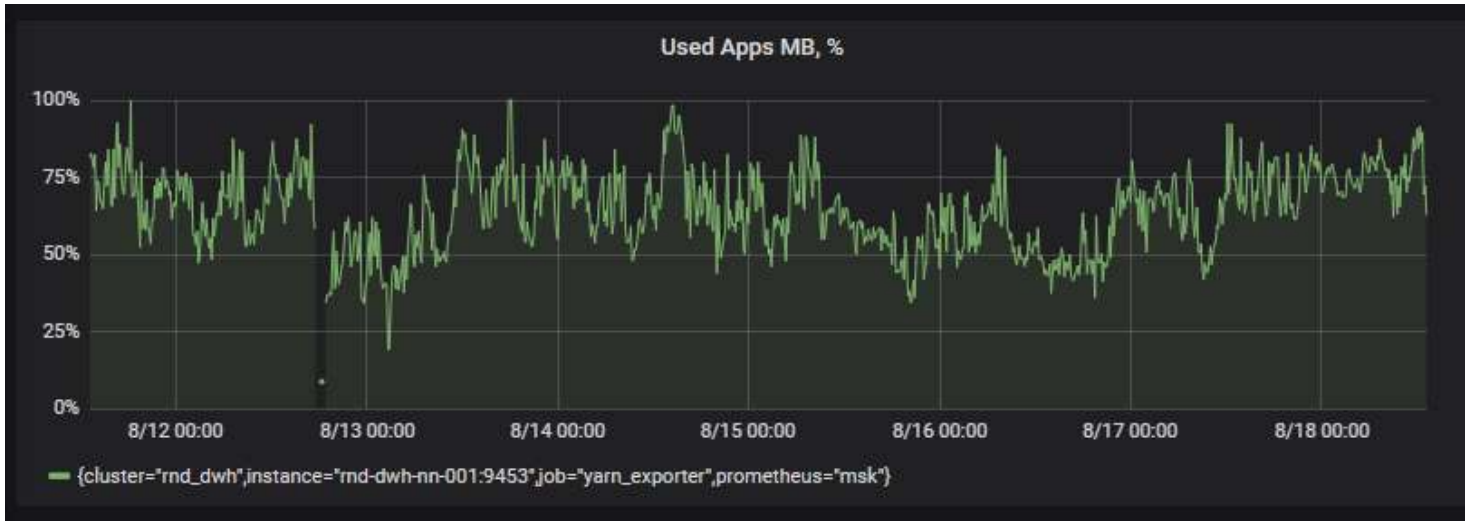
Hello, Conference!_

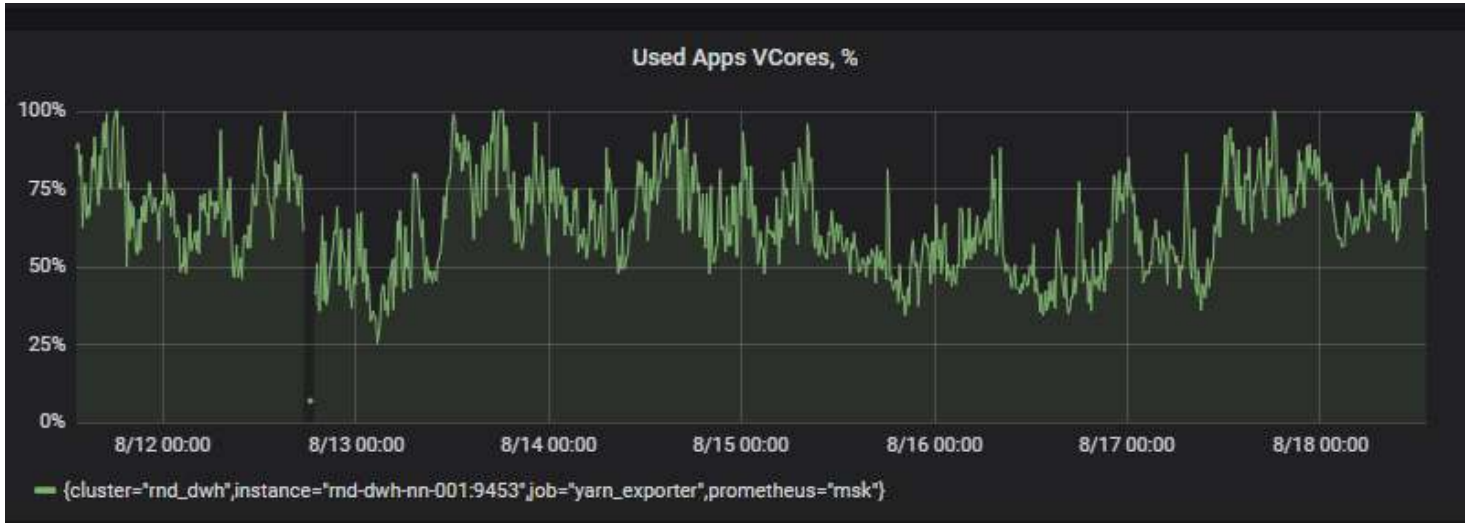
# Построение оптимальной инфраструктуры для больших данных

Григорий Коваль
Менеджер по продукту
Центр управления данными
Центр Big Data

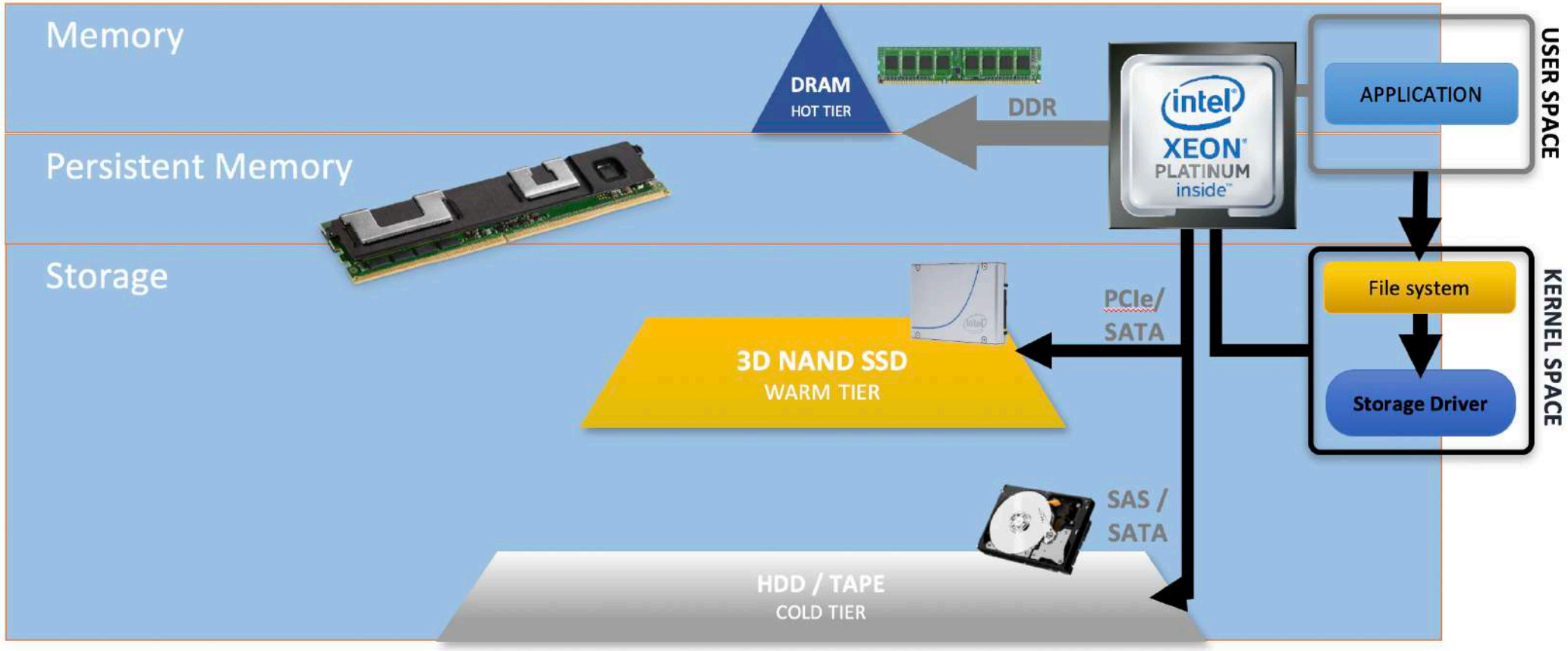
22 октября 2020

# ПРИМЕР ТЕКУЩЕЙ НАГРУЗКИ НА HADOOP КЛАСТЕР



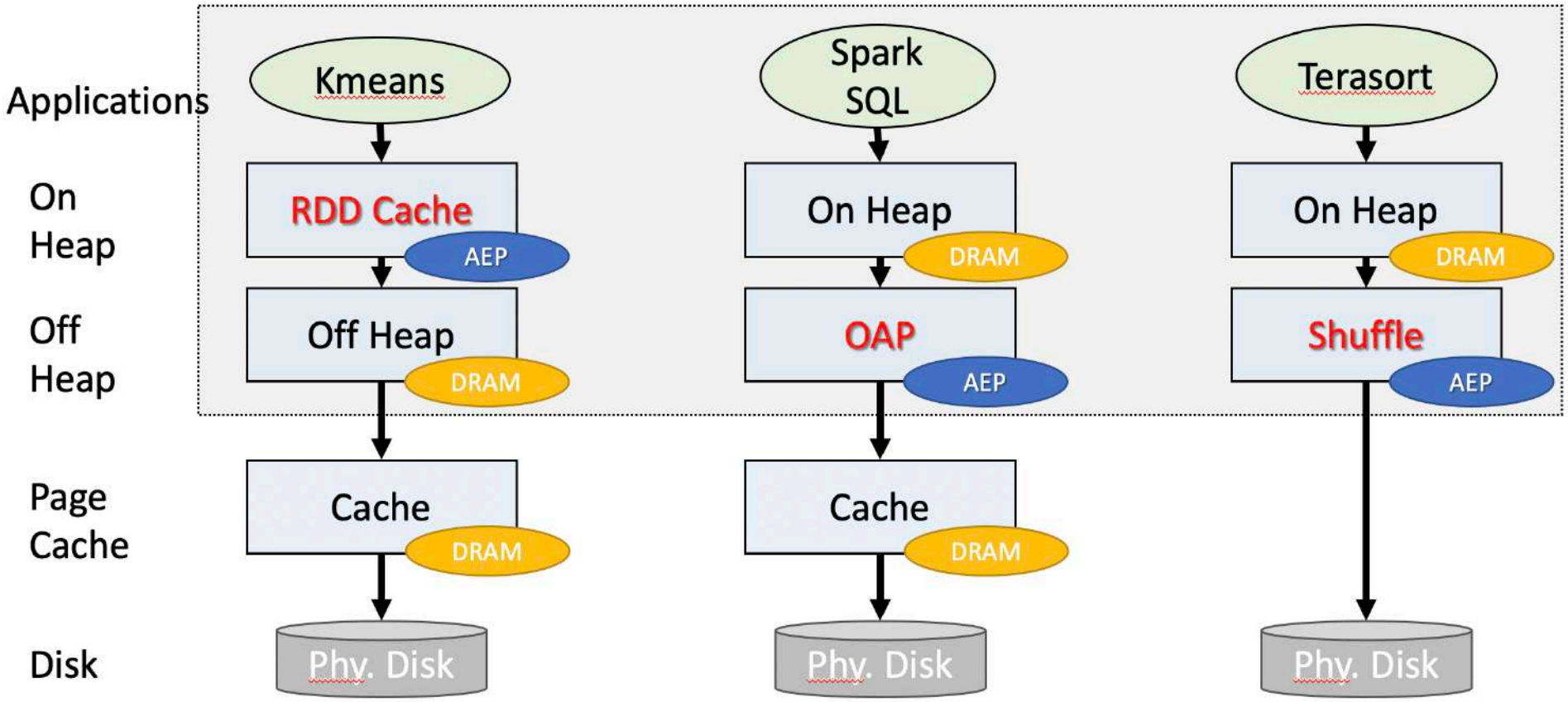



# НОВЫЙ ТИП ПОСТОЯННОЙ ПАМЯТИ




Copyright ©, Intel Corporation. All rights reserved.

# АРХИТЕКТУРА SPARK + AEP




Copyright ©, Intel Corporation. All rights reserved.

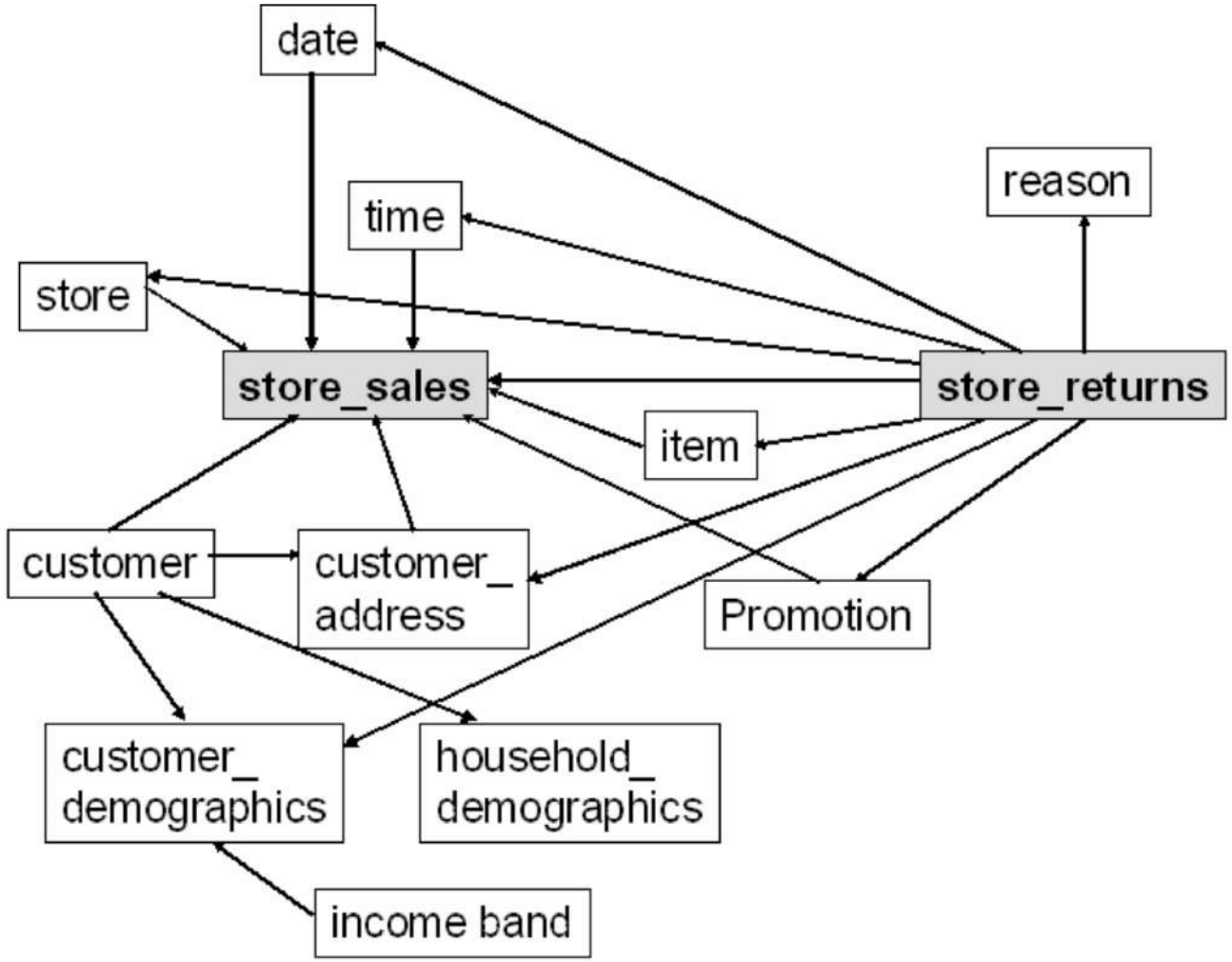
# ЗАДАЧИ ТЕСТИРОВАНИЯ

- Проанализировать как OAP кэш может повлиять на разные типы Spark SQL запросов (как простые так и сложные - аналитические)
- Проверить на сколько эффективно работает OAP кэш, дает ли он такой выигрыш, который оправдывает затраты на его внедрение

# КОНФИГУРАЦИЯ ОБОРУДОВАНИЯ И ПО НА ТЕСТОВЫХ СТЕНДАХ

| Hardware/Software | Model/Version Intel Test Lab in Swindon (a) | Model/Version Intel Test Lab in Swindon (b) | Model/Version MTS Test Lab in Moscow (c) | Model/Version MTS Test Lab in Moscow (d) | Model/Version MTS baseline configuration ( e) |
|---|---|---|---|---|---|
| CPU | 2 * Intel(R) Xeon(R) Platinum 8280 CPU @ 2.70GHz | 2 * Intel(R) Xeon(R) Platinum 8280 CPU @ 2.70GHz | Intel(R) Xeon(R) Gold 6140 CPU @ 2.30GHz | Intel(R) Xeon(R) Gold 6140 CPU @ 2.30GHz | Intel(R) Xeon(R) Gold 6140 CPU @ 2.30GHz |
| RAM | 192 GB | 192 GB | 384 GB | 384 GB | 384 GB |
| PMEM | 12*128GB = 1,5 TB | | | | |
| Optane SSD + IMDT | | | 640 GB Optane SSD + 384 GB RAM | | |
| Network | 10 Gb/s | 10 Gb/s | 10 Gb/s | 10 Gb/s | 10 Gb/s |
| Spark Version | 2.3.4 | 2.3.4 | 2.4.4 | 2.4.4 | 2.4.4 |
| Hadoop | Apache Hadoop 2.7 | Apache Hadoop 2.7 | HDP 2.6.5 | HDP 2.6.5 | HDP 2.6.5 |
| Cluster nodes | 3 hardware data nodes and 1 master node | 3 hardware data nodes and 1 master node | 3 hardware data nodes and 1 master node | 3 hardware data nodes and 1 master node | |

# МОДЕЛЬ ДАННЫХ И ПАРАМЕТРЫ ДЛЯ TPC-DS ТЕСТА




| | | |
|---|---|---|
| Number of fact tables | 7 | |
| Number of dimension tables | 17 | |
| Number of columns | min | 3 |
| | max | 34 |
| | avg | 18 |
| Number of foreign keys | 104 | |
| Row length [bytes][4] | min | 16 |
| | max | 317 |
| | avg | 136 |

| Table | Rowcounts | | | |
|---|---|---|---|---|
| | 100GB | 1TB | 10TB | 100TB |
| store_ sales | 288M | 2.9B | 30B | 297B |
| store_ returns | 14M | 147M | 1.5B | 15B |
| store | 200 | 500 | 750 | 1,500 |
| customer | 2M | 8M | 20M | 100M |
| Items | 200K | 300K | 400K | 500K |

# ЗАПРОСЫ TPC-DS ИСПОЛЬЗУЕМЫЕ В ТЕСТАХ

| Тип | Номер запроса TPC-DS |
|---|---|
| **Интерактивные** (данные за 1-3 месяца) — Простые запросы | 32, 42, 55 |
| **Отчеты** (данные за 1 год) — относительно простые запросы | 26, 27, 53 |
| **Аналитические** (Несколько лет, регулярные выражения в фильтрах) — | 68,48 |
| **Сложные** — Объединение несколько таблиц с фактами за несколько лет, оконные функции | 76 |

**Условия тестирования:**

1. Использовались дата сеты 1 и 6 ТВ
2. Скрипты запускались в двух конфигурациях – с OAP and и без включения OAP
3. Каждый тест был запущен дважды (с “холодным” и “горячим” OAP кэшом)
4. Tests were performed in Intel Swindon lab (see configuration on slide #2)

# ПРИМЕРЫ ЗАПРОСОВ:

## Reporting query (#26)

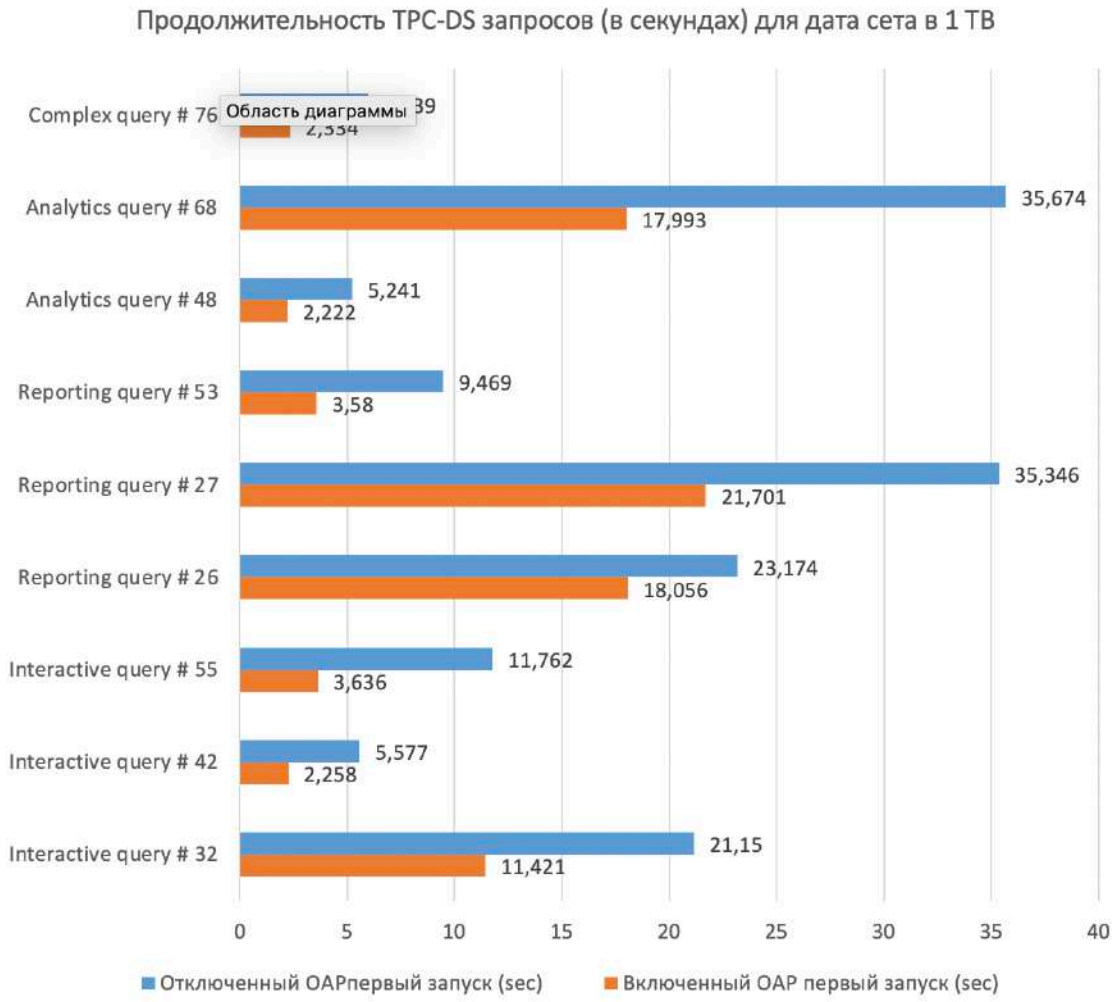
```
select i_item_id,
     avg(cs_quantity) agg1,
     avg(cs_list_price) agg2,
     avg(cs_coupon_amt) agg3,
     avg(cs_sales_price) agg4
from catalog_sales, customer_demographics, date_dim, item, promotion
where cs_sold_date_sk = d_date_sk and
    cs_item_sk = i_item_sk and
    cs_bill_cdemo_sk = cd_demo_sk and
    cs_promo_sk = p_promo_sk and
    cd_gender = 'M' and
    cd_marital_status = 'S' and
    cd_education_status = 'College' and
    (p_channel_email = 'N' or p_channel_event = 'N') and
    d_year = 2000
group by i_item_id
order by i_item_id
limit 100
```

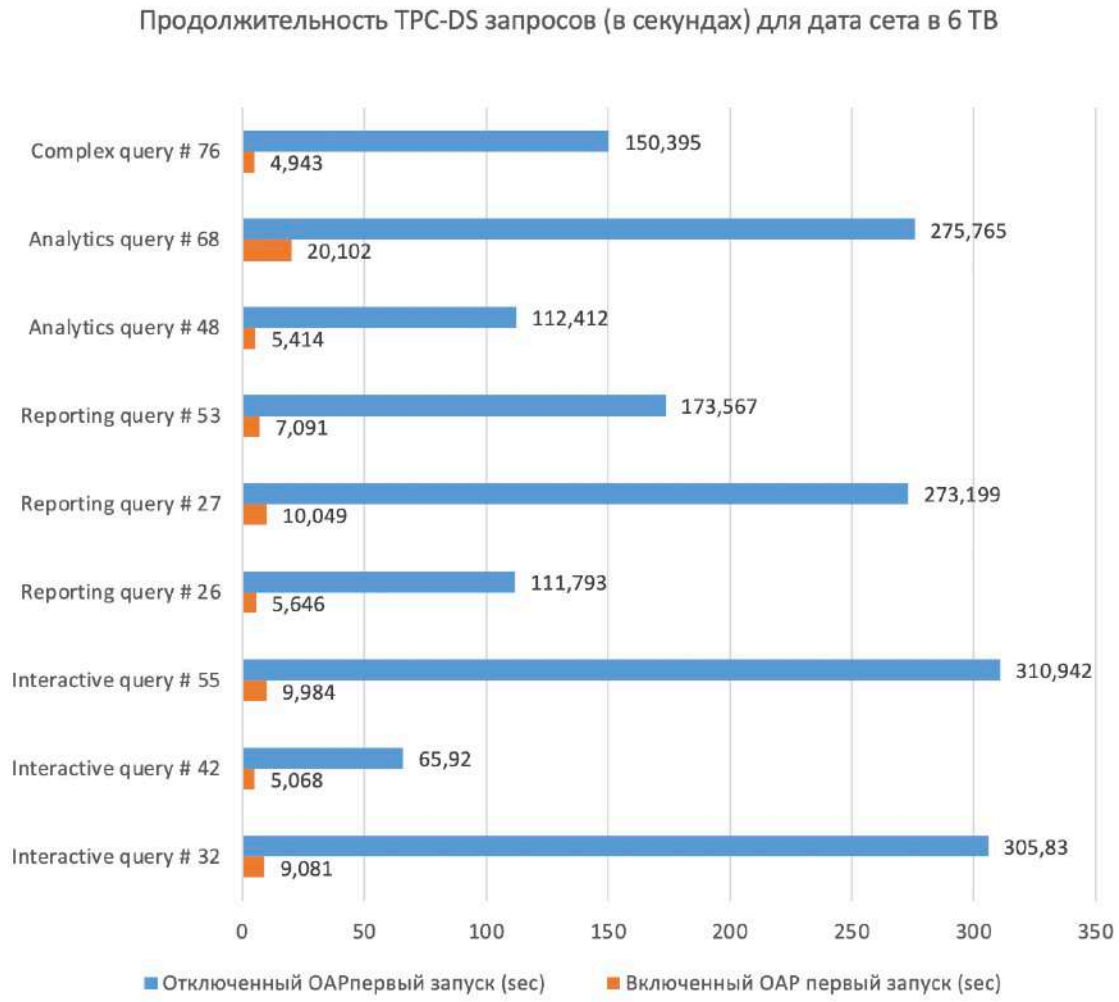
## Complex query (#76)

```
SELECT
  channel, col_name, d_year, d_qoy, i_category, COUNT(*) sales_cnt,
  SUM(ext_sales_price) sales_amt
FROM(
  SELECT
    'store' as channel, ss_store_sk col_name, d_year, d_qoy, i_category,
    ss_ext_sales_price ext_sales_price
  FROM store_sales, item, date_dim
  WHERE ss_store_sk IS NULL
   AND ss_sold_date_sk=d_date_sk
   AND ss_item_sk=i_item_sk
  UNION ALL
  SELECT
    'web' as channel, ws_ship_customer_sk col_name, d_year, d_qoy, i_category,
    ws_ext_sales_price ext_sales_price
  FROM web_sales, item, date_dim
  WHERE ws_ship_customer_sk IS NULL
   AND ws_sold_date_sk=d_date_sk
   AND ws_item_sk=i_item_sk
  UNION ALL
  SELECT
    'catalog' as channel, cs_ship_addr_sk col_name, d_year, d_qoy, i_category,
    cs_ext_sales_price ext_sales_price
  FROM catalog_sales, item, date_dim
  WHERE cs_ship_addr_sk IS NULL
   AND cs_sold_date_sk=d_date_sk
   AND cs_item_sk=i_item_sk) foo
GROUP BY channel, col_name, d_year, d_qoy, i_category
ORDER BY channel, col_name, d_year, d_qoy, i_category
limit 100
```

# РЕЗУЛЬТАТЫ

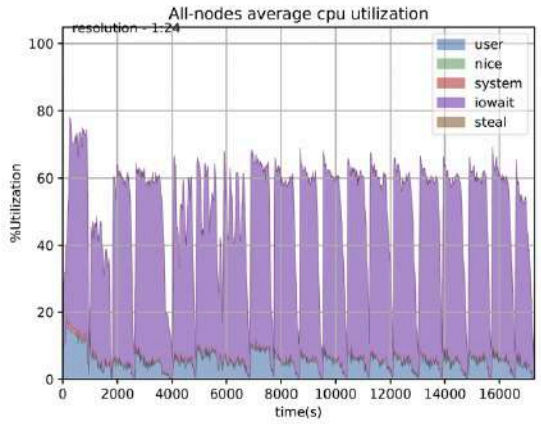
Продолжительность разных типов запросов (с включенным oap и отключенным oap)

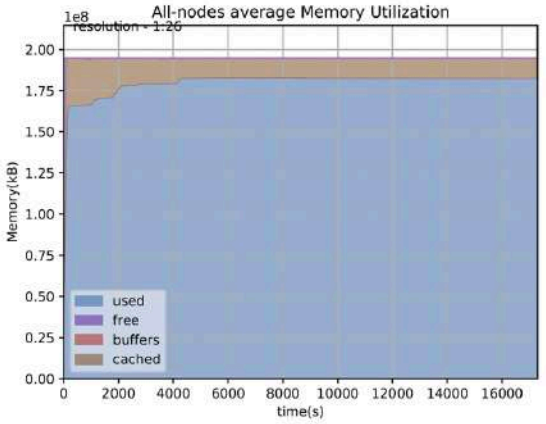


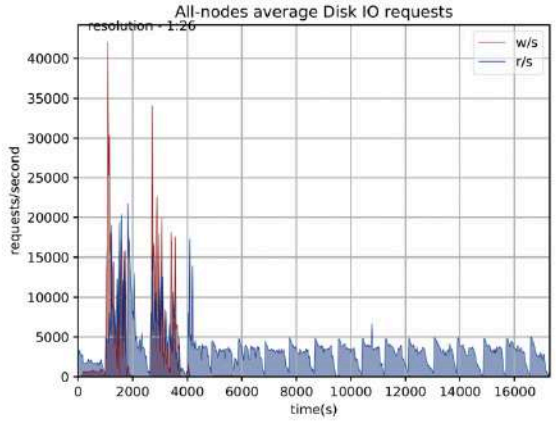



# РЕЗУЛЬТАТЫ

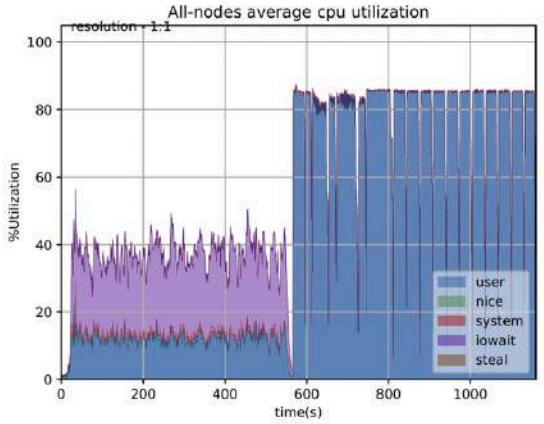
Отличие в нагрузке на кластер при включенном и отключенном OAP

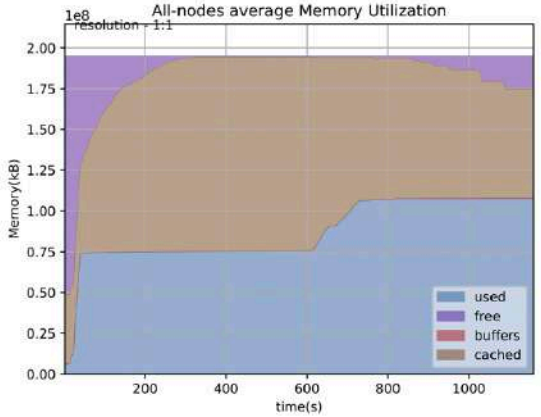


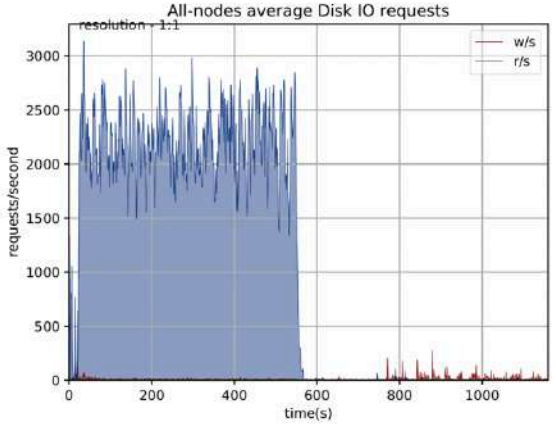




Системные метрики во время тестов при отключенном OAP




Системные метрики во время тестов при включенном OAP

# ВЫВОДЫ

- Тесты TPC-DS для Spark включают в себя разные типы запросов, все типы запросов ускоряются при влечении OAP кэша
- OAP кэш может ускорить  тяжелые запросы к данным до 30 раз

# СПАСИБО!