

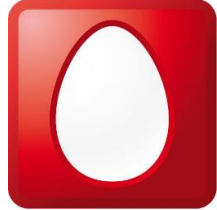
Управление данными для аналитики с использованием технологий Intel

Коваль Григорий, Big Data Architect, МТС
Шостко Дмитрий, СТО BigData/DWH, МТС



МТС

Ты знаешь, что можешь!

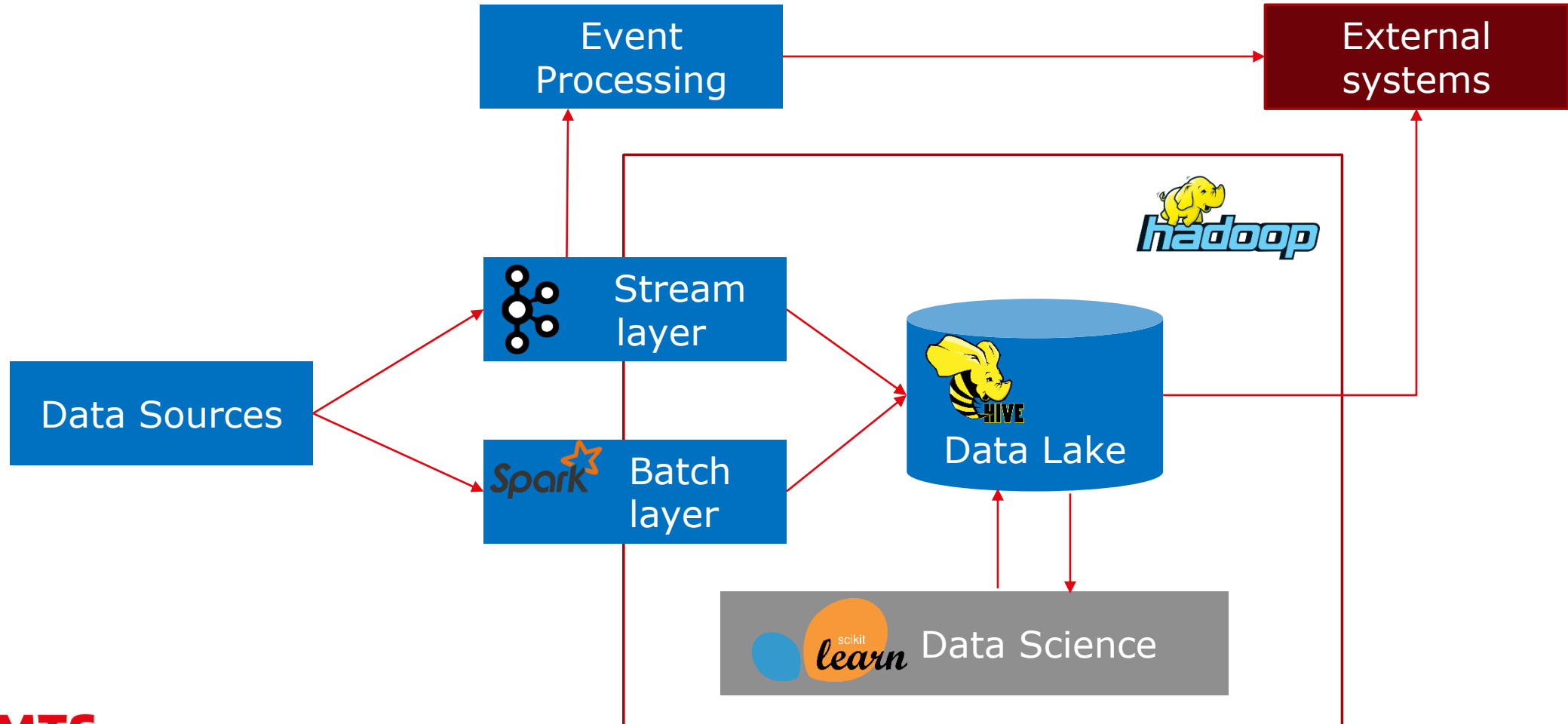


Источники данных МТС

- Мобильный оператор > **105** млн клиентов
- География: 4 страны
- Retail: 5500 точек продаж
- Финансовые услуги
- Каналы гибридного и спутникового телевидения
- Оператор №1 для бизнеса: более 40% корпоративных абонентов в сегменте мобильной связи в России
- Лидер рынка интернета вещей в России: 45% рынка IoT по числу телематических SIM-карт
- Облачные сервисы
- Гео сервисы
- Контекстная и мобильная реклама
- Мобильные приложения



Верхнеуровневая архитектура



Условия тестирования Intel NVMe SSD (как блочное устройство)

Среда тестирования

Кластер HDP из 12-и datanode

Запускаемые тесты

Тестирования осуществлялось с помощью стандартных тестов TeraGen, TeraSort и TeraValidate с использованием MapReduce. Тесты запускались в режиме кластера. Тестирование производилось на объёмах 1 Tb, 1Gb

Конфигурация серверов

Intel(R) Xeon(R) CPU E5-2660 v3 @ 2.60GHz (40 ядер)

128 Gb system memory (DRAM)

Oracle Linux Server 7.3 (3.10.0-514.21.1)

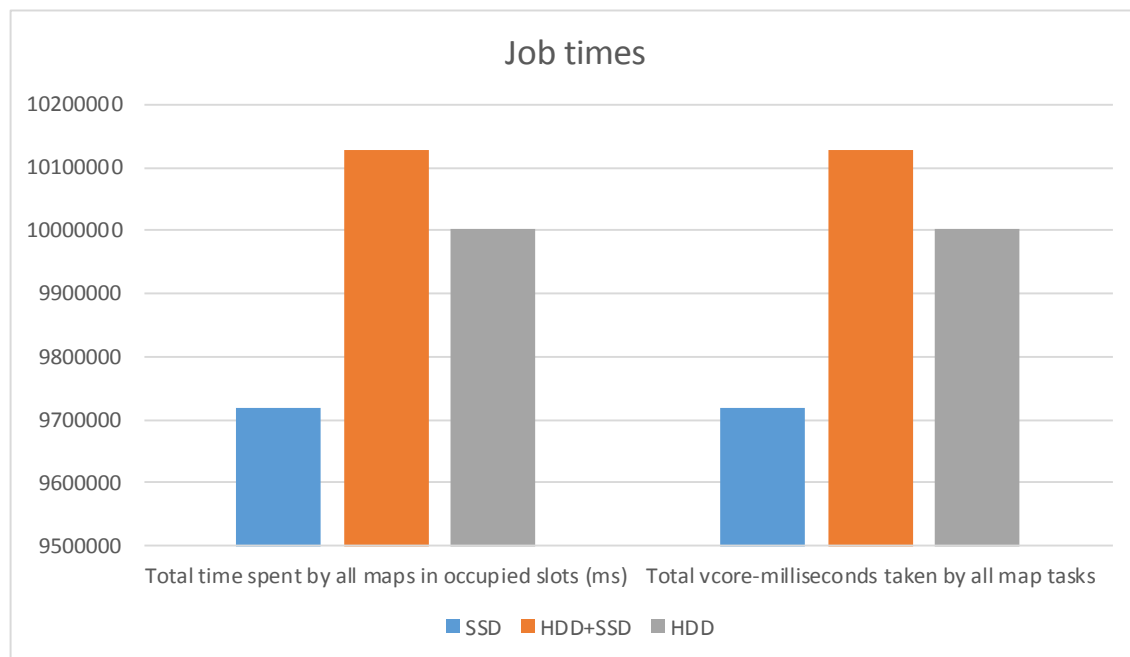
Java 1.8

HDP 2.6.1

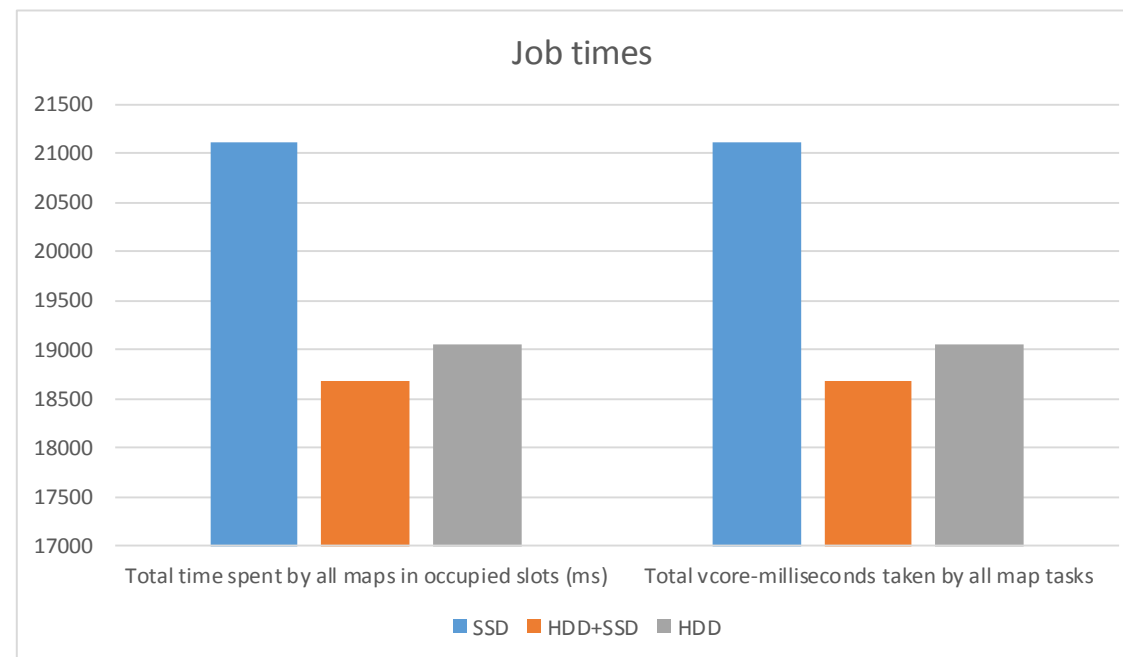
Результаты тестирования Intel NVMe SSD как блочное устройство



1 TB исходных данных



1 GB исходных данных



Условия тестирования Intel Optane SSD с технологией Intel Memory Drive Technology (Optane SSD DC P4800X)



Среда тестирования

Кластер HDP из 3-х нод, каждая из которой являлась datanode.

Запускаемые тесты

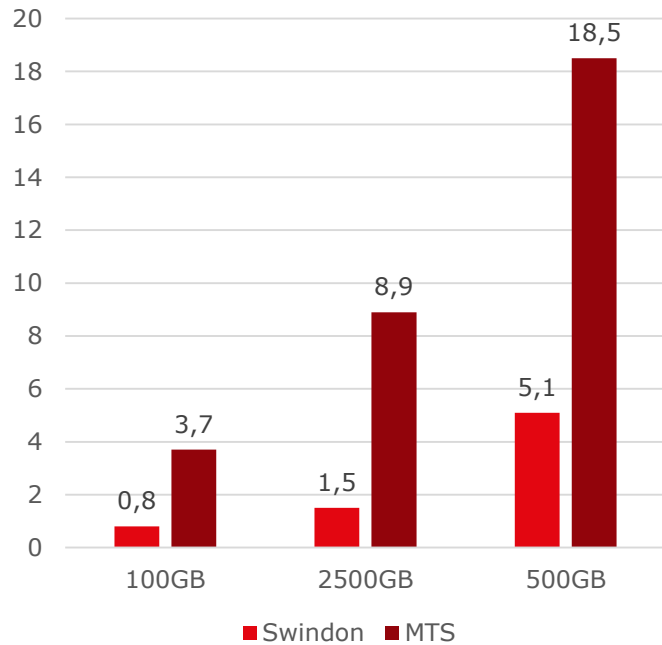
Тестирования осуществлялось с помощью тестов TeraGen, TeraSort и TeraValidate с использованием и Spark2 (<https://github.com/ehiggs/spark-terasort>). Тесты запускались в режиме кластера. Тестирование производилось на объёмах 100Gb, 250Gb и 500Gb.

Лаборатория Swindon	Тестовый контур МТС
<ul style="list-style-type: none">• Intel(R) Xeon(R) Gold 6148 CPU @ 2.40GHz (80 ядер)• 1 Tb system memory (DRAM + Optane)• Red Hat Enterprise 7.3 (kernel 3.10.0-693.5.2)• Java 1.8• HDP 2.6.1 (Spark 2)	<ul style="list-style-type: none">• Intel(R) Xeon(R) CPU E5-2660 v3 @ 2.60GHz (40 ядер)• 377 Gb system memory (DRAM)• Oracle Linux Server 7.3 (3.10.0-514.21.1)• Java 1.8• HDP 2.6.1 (Spark 2)

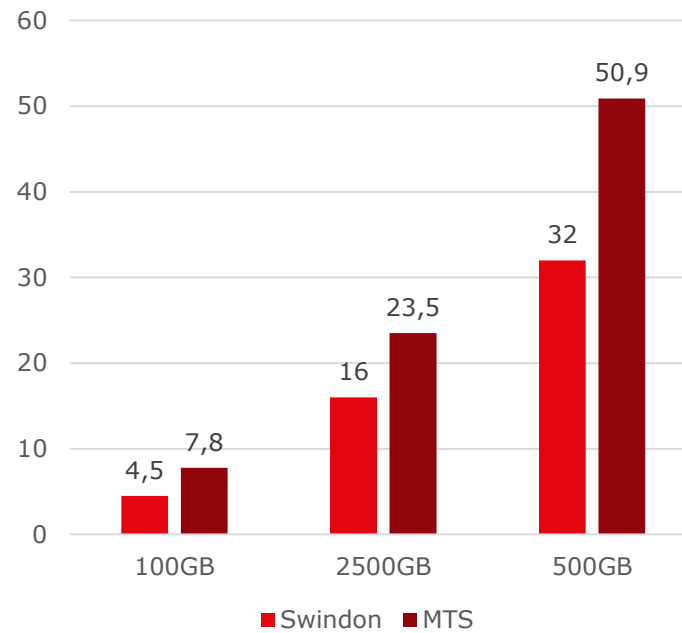
Результаты тестирования Intel Optane SSD с технологией Intel Memory Drive Technology (Optane SSD DC P4800X)



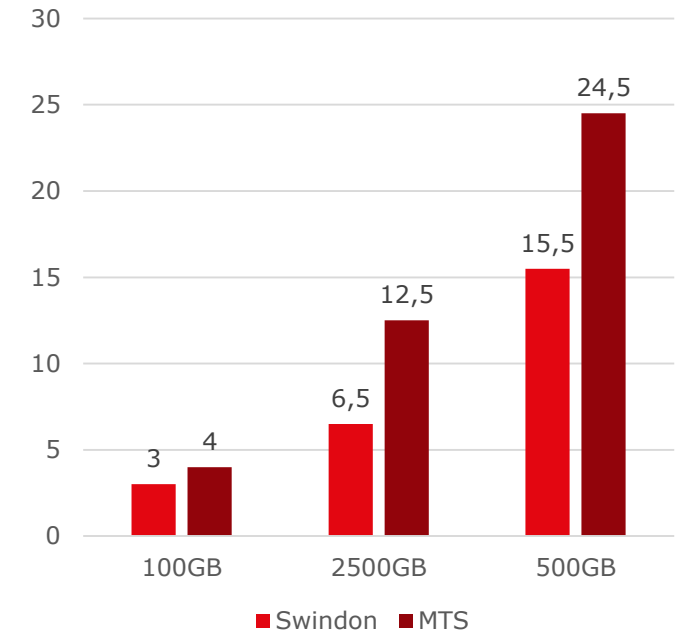
TeraValidate



TeraSort



TeraGen



Ты знаешь, что можешь!

Спасибо!



МТС

Ты знаешь, что можешь!