

Суперкомпьютеры для Big Data. Инфраструктура VAR

Васильев Н.П., Ровнягин М. М.

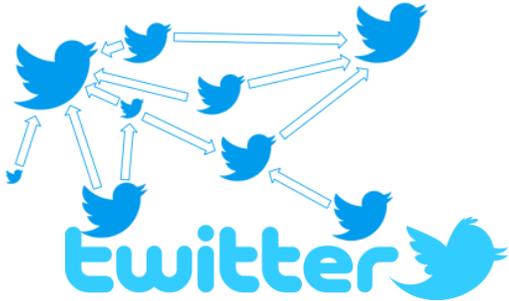


Компьютерные  
системы 
& *технологии*

Big Data сегодня



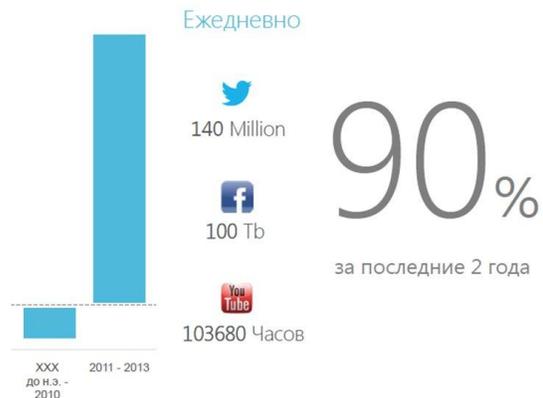
- К 2016 число пользователей IP-сетей превысит население земли в 3 раза



- В Twitter ежемесячно выполняется более 30 млрд операций поиска

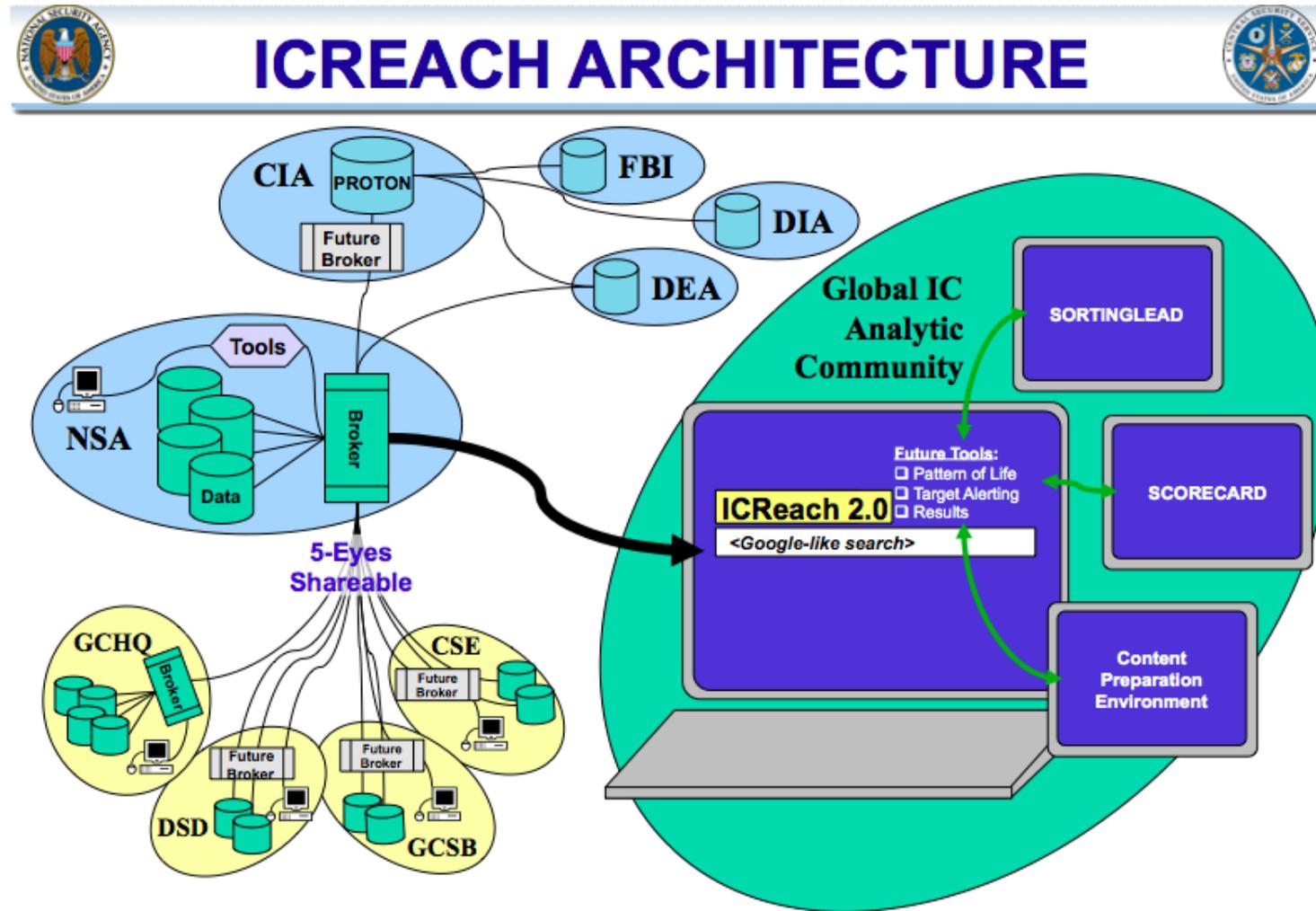


- 250 млн фотографий загружается на facebook ежедневно



- За последние 2 года человечество собрало больше данных чем за всю предыдущую историю

Big Data сегодня: Поисковая система АНБ США



- Система ICReach, используемая АНБ США рассчитана на добавление 5 млрд. записей ежедневно
- Позволяет осуществлять полнотекстовый поиск по более 850 млрд. записям с конфиденциальной информацией
- В общем случае система индексирует метаданные записей (объем которых в 2011 году превысил 1 трлн)
- Обладает удобным интерфейсом

ПРИМЕРЫ ЗАДАЧ ДЛЯ BIG DATA



Анализ данных о состоянии больного для выявления наиболее эффективных методов лечения и медикаментов.



Заблаговременное выявление опасных ситуаций и участков.



Разработка более эффективных маркетинговых программ, максимальная информированность при принятии ключевых бизнес-решений, прогнозирование рынка, конкуренции и стратегии развития.



Предсказание преступлений на основе анализа данных о предыдущих нарушениях.



Повышение качества работы и понимания нужд граждан, увеличение бюджета на другие программы за счет средств, сэкономленных с помощью Big Data.

ОТРАСЛИ, ВЛАДЕЮЩИЕ МАССИВОМ BIG DATA, ПБ



По данным с сайта: cnews.ru

Хранение и анализ транзакций в банковской сфере.

ТРАНЗАКЦИИ ПО КАРТАМ СБЕРБАНКА В РЕАЛЬНОМ ВРЕМЕНИ

АКТИВНОСТЬ ПО СТРАНАМ

млн руб. за последний час



МОСКОВСКОЕ ВРЕМЯ

21:28:07

Лондон
18:28:07

Нью-Йорк
13:28:07

Лос-Анжелес
10:28:07

Токио
02:28:07

ВИД ОПЕРАЦИИ

сумма в млн руб. за последний час



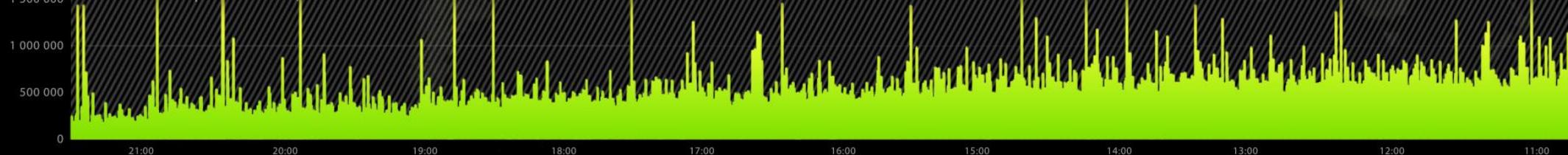
КОЛИЧЕСТВО ТРАНЗАКЦИЙ

5 159 844

СУММА ТРАНЗАКЦИЙ

12 702 210 329 руб.

ОБЪЕМ ОПЕРАЦИЙ В МИНУТУ, РУБ.



Классические базы данных vs. NoSQL-системы

- Классические РСУБД не в состоянии работать с Big Data
- Для хранения столь больших объемов данных крупные компании используют нереляционные NoSQL-системы

- MySQL:

- 300 ms запись
- 350 ms чтение

- Apache Cassandra:

- **0.12** ms запись
- **15** ms чтение



Идея: NoSQL + гибридные суперкомпьютеры

NoSQL:

Высокопроизводительные,
но умеют только хранить
данные

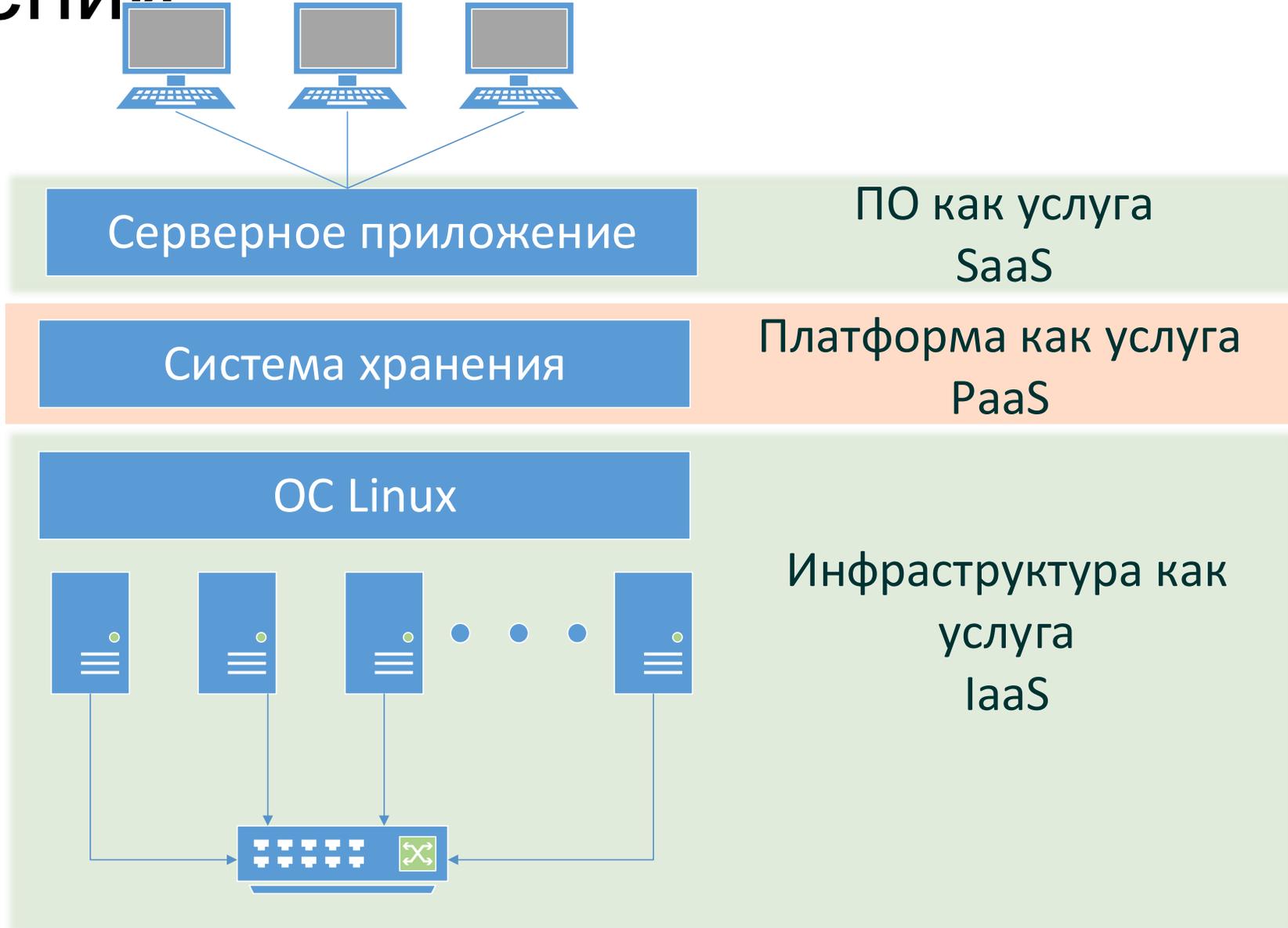


Суперкомпьютеры: с 2008 года все
чаще используют технологию
ускорения вычислений на
дискретном сопроцессоре, т.е.
имеют гибридную архитектуру

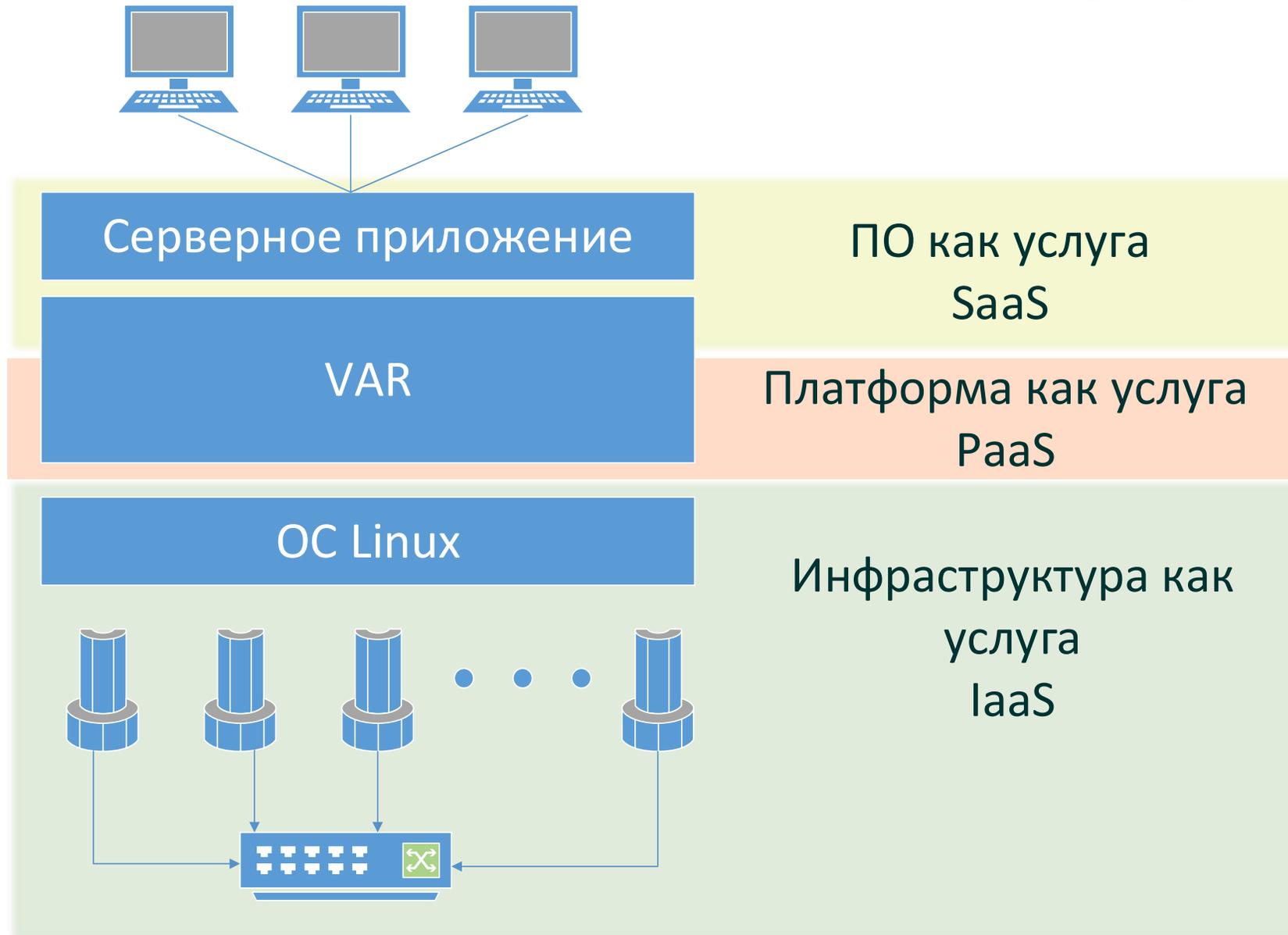


Высокопроизводительные системы поиска и хранения данных, позволяющие выполнять операции дополнительной обработки информации (шифрование, сжатие, анализ, вычисление сигнатур и т.д.) при ее сохранении

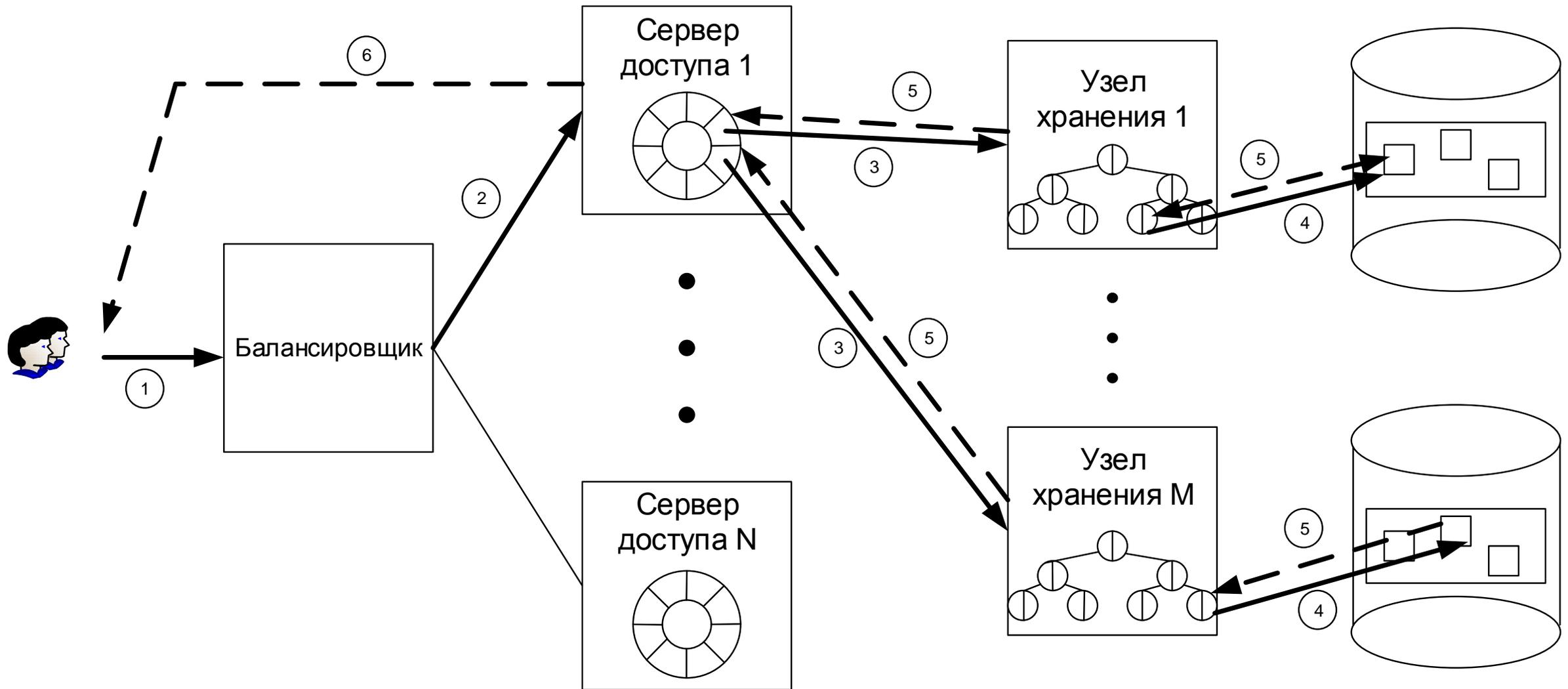
Архитектура классических BigData-приложений



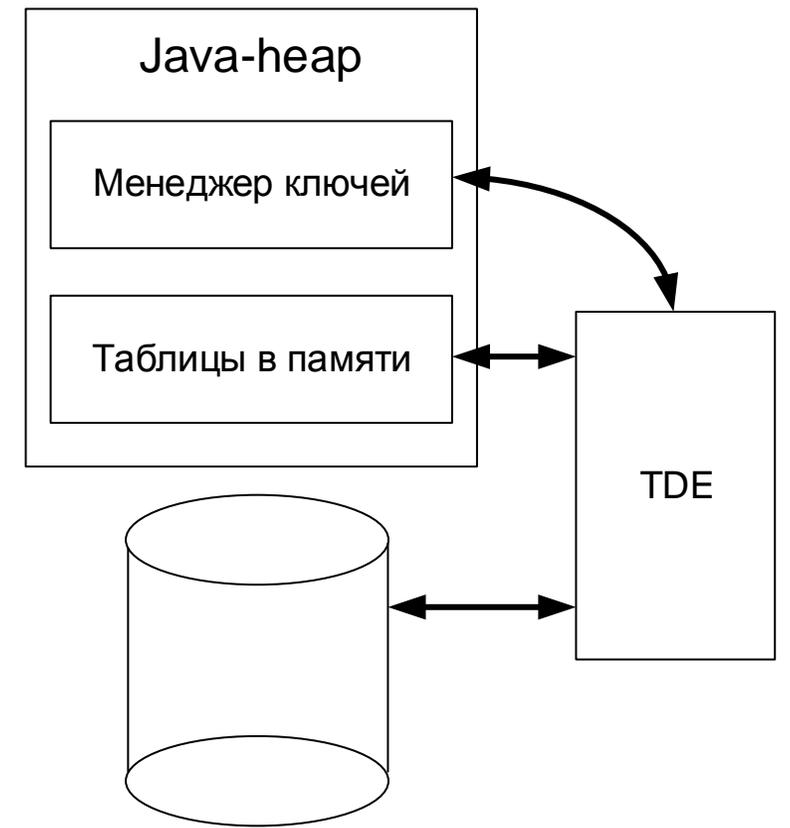
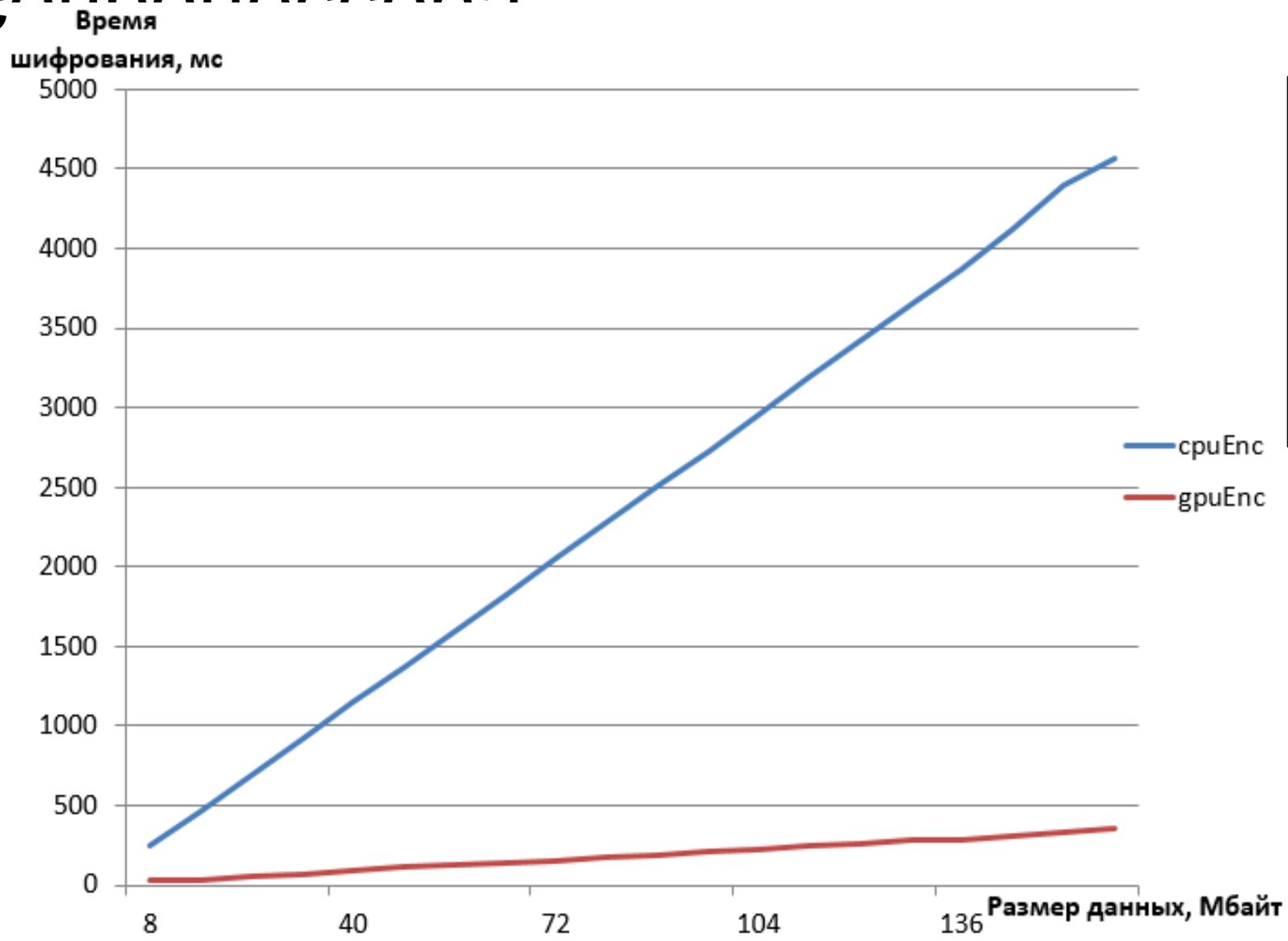
VAR – НОВЫЙ ВЗГЛЯД НА АРХИТЕКТУРУ BigData



Программная платформа VAR

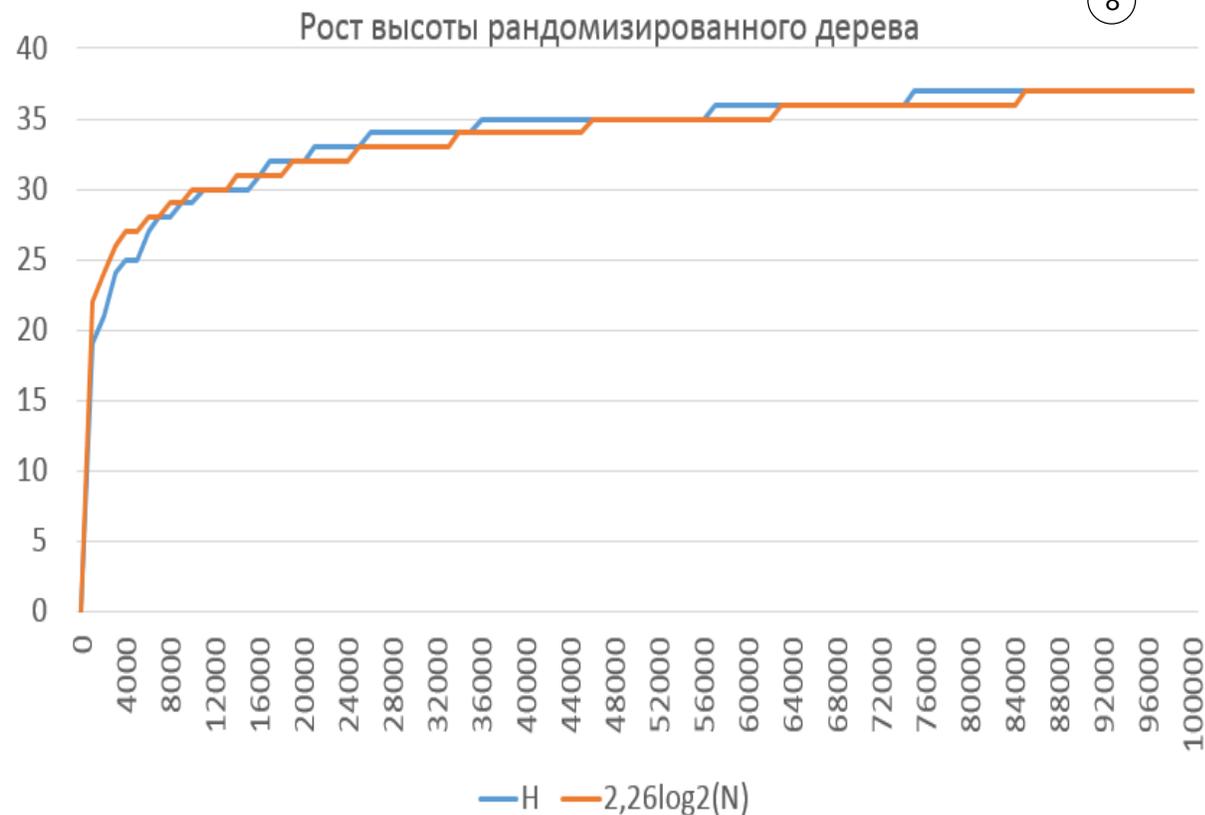
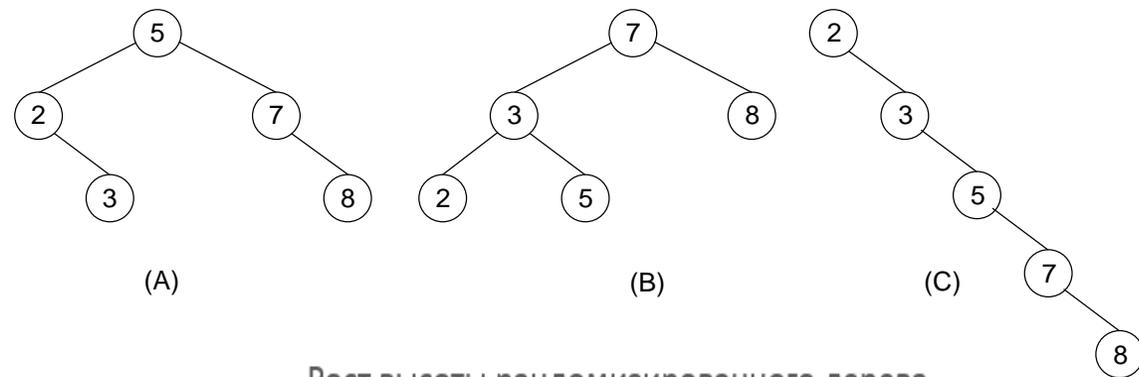
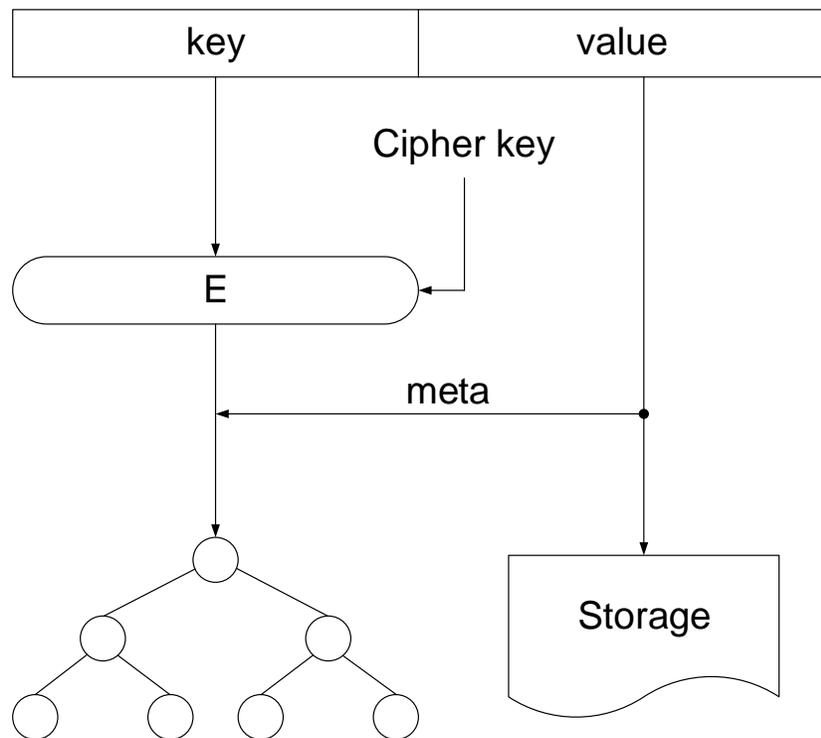


Технологии. TDE по алгоритму ГОСТ 28147-89 с применением GPGPU-сопроцессора



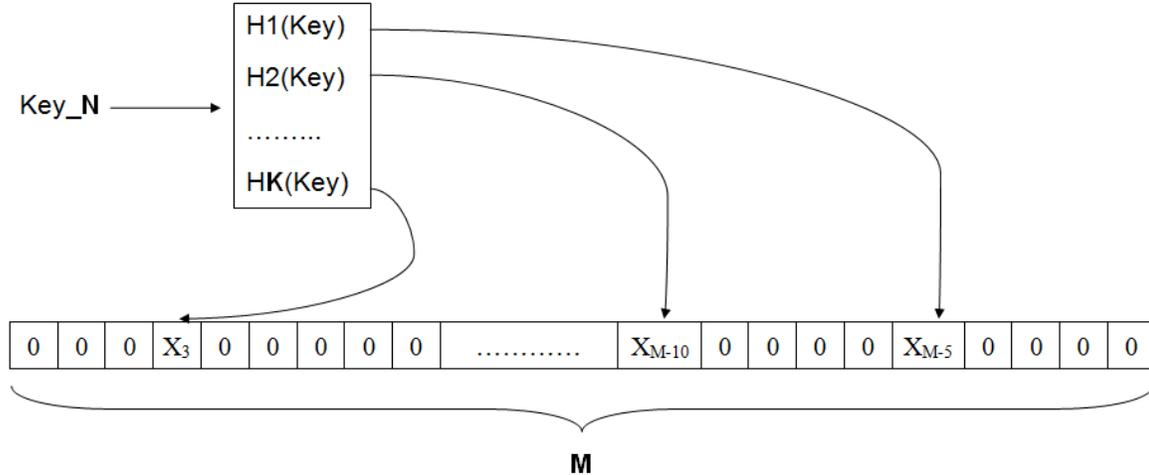
ТЕХНОЛОГИИ.

стохастическое бинарное дерево поиска для GPGPU-СИСТЕМ



- Математическое ожидание высоты случайного бинарного дерева поиска с n ключами равно $O(\log_2 n)$
- В качестве E применяется GPU-версия ГОСТ 28147-89

Технологии: модифицированный фильтр Блума в гибридной CPU/GPU реализации



- Фильтр Блума - вероятностная структура данных, которая отражает сведения о добавленных элементах в хранилище данных в формате «ключ-значение»
- Решить проблему удаления элементов можно, используя модификацию фильтра Блума со счетчиками. При удалении счетчик уменьшается на единицу или не изменяет своего состояния, если его значение равно нулю или максимуму
- В CUDA-реализации каждая нить вычисляет K хеш-функций для каждого запроса. Массовый параллелизм нитей CUDA позволяет достичь 5-кратного роста производительности

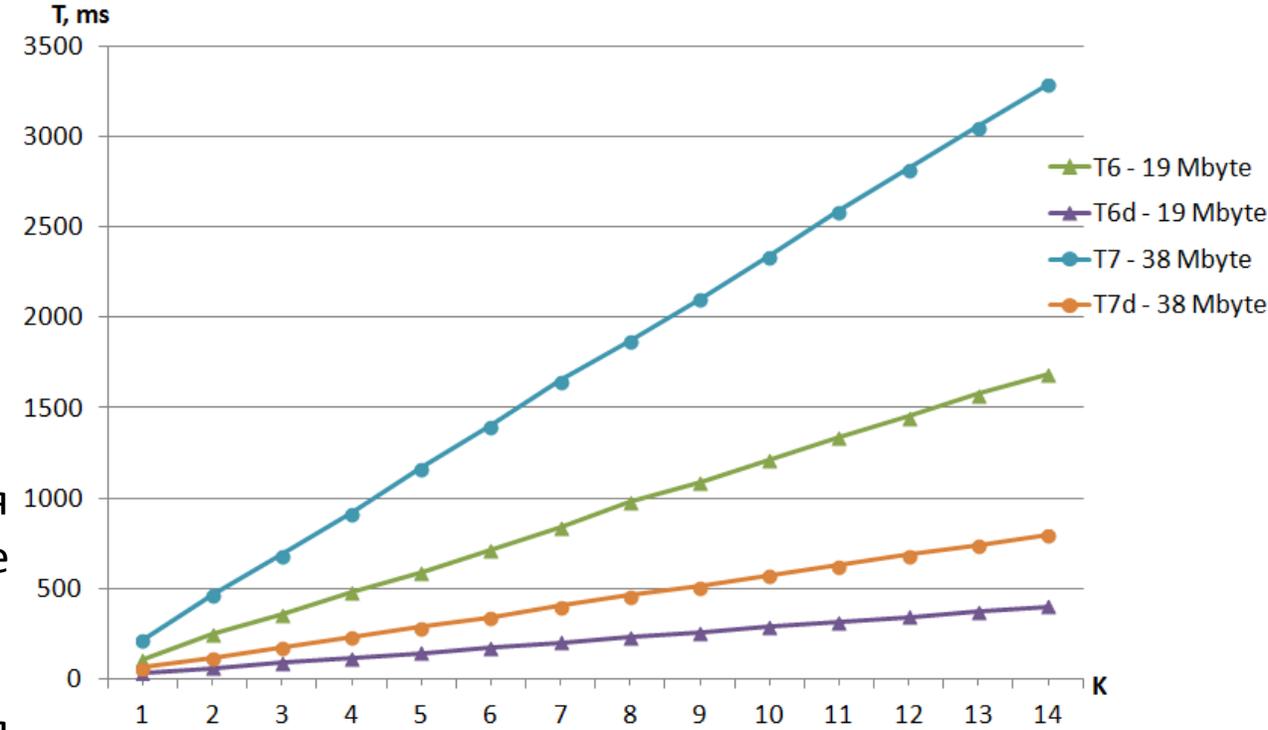


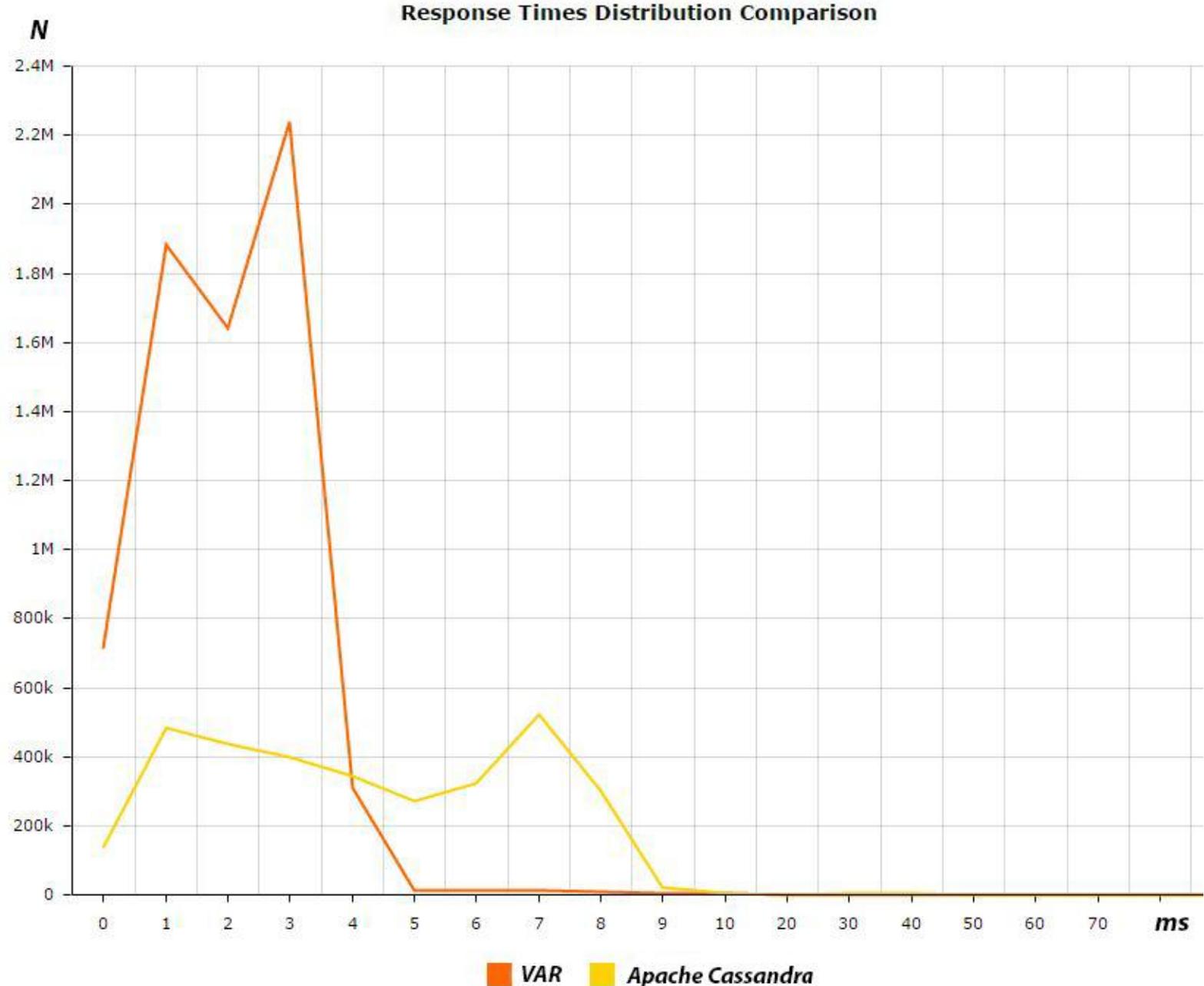
График зависимости времени выполнения программы от количества хеш-функций для CPU (графики T6, T7) и GPU (графики T6d, T7d) версий

Система VAR (предпосылки)

- Работа по созданию прототипа велась в течение трех лет
- В рамках кандидатской и докторской диссертаций
- В инициативном порядке
- Разработаны методы и средства поиска данных в гибридных суперкомпьютерных системах
- Разработана архитектура системы
- Создан прототип

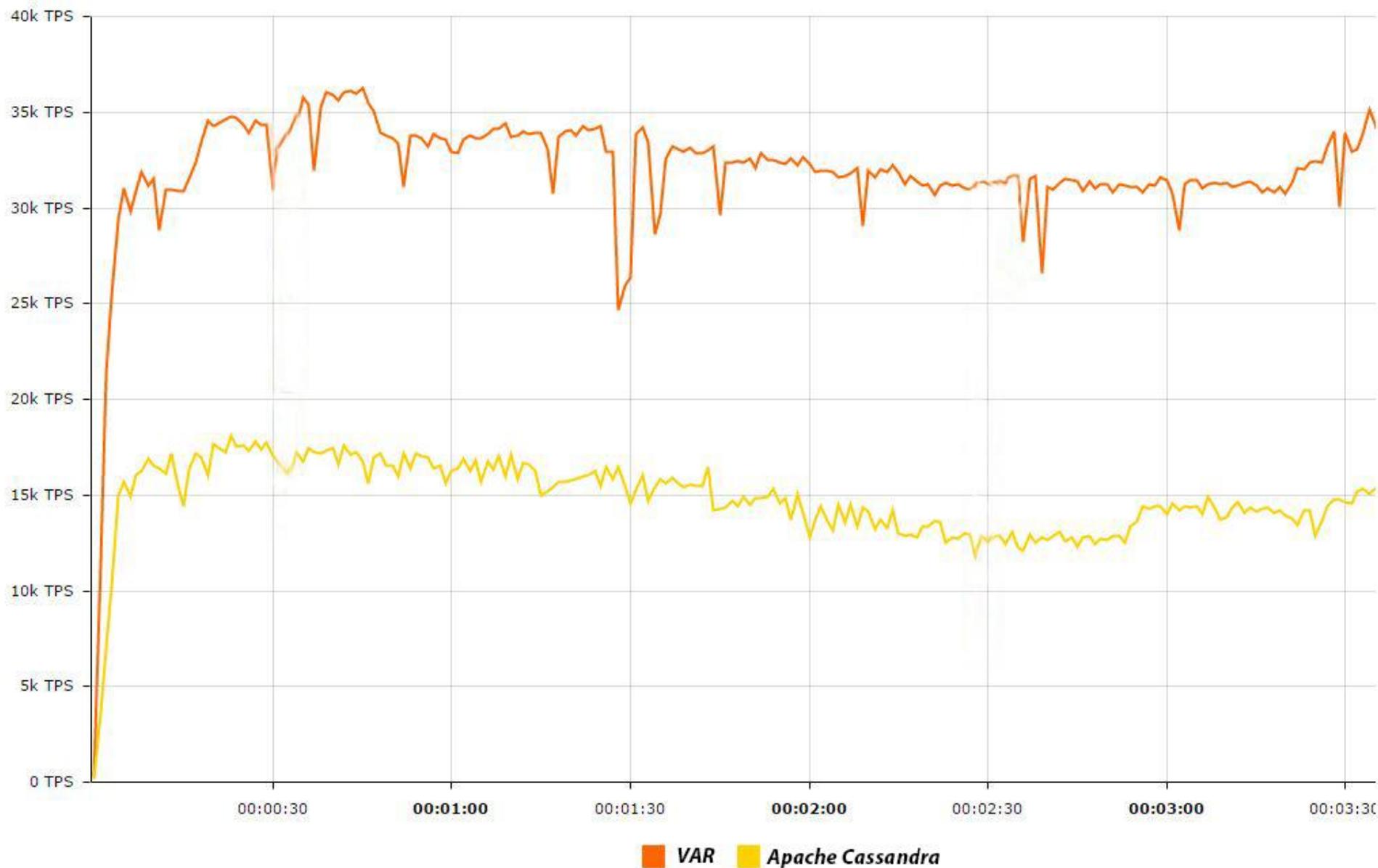
Система VAR (время отклика)

- Оборудование:
 - Intel Core i7 – 3770K
 - 32 Gb DDR3
 - NVIDIA GeForce GTX680 (4Gb GDDR5)
- Имитация нагрузки:
 - Утилита нагрузочного тестирования Apache Jmeter
- Что тестировалось:
 - Узел хранения программной платформы VAR
 - Узел распределенной NoSQL-системы Apache Cassandra



Система VAR (производительность, транзакции в секунду)

TPS Timeline Comparison



Система VAK (преимущества)

- Работает **в 2 раза быстрее** ближайших зарубежных аналогов

Краткосрочный метеопрогноз может быть просчитан не за 120 мин, а за час

- Имеет встроенную **систему шифрования** по алгоритму ГОСТ 28147

Физический доступ к дискам может оказаться бесполезным для злоумышленника

- При использовании в гибридных суперкомпьютерных системах потребляет на **25-30% меньше электроэнергии**, чем аналогичные классические решения

Согласно данным из рейтинга TOP500 гибридные суперкомпьютеры обладают наилучшими показателями производительность/потребляемая мощность

- Система построена **только при участии российских разработчиков** на базе технологии Java без использования проприетарных программных модулей и может быть использована в рамках **стратегии импортозамещения** ПО

Система VAR (состояние разработки)

- Пройдена стадия НИР
- Создан прототип системы (готовность основной системы 30%)
- Прототип может быть преобразован в промышленный образец, в течение 1,5 лет
- VAR является принципиально новым **российским** программным продуктом

Система VAR (версии продукта)

- Свободно-распространяемая
- Коммерческая с подключаемыми модулями
- Специализированная под заказчика (в т.ч. для государственных критически важных задач)