

Трансформация архитектуры решения Performance Management в условиях Big Data: предпосылки и вызовы.

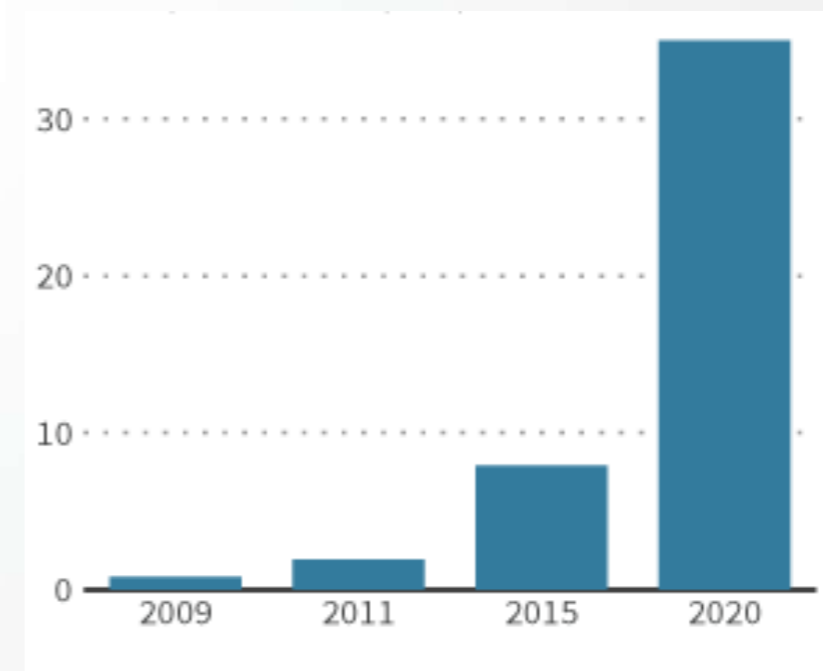
Основная мысль.

Для принятия эффективных решений требуется обрабатывать все больше и больше данных и зачастую в режиме, максимально приближенном к реальному времени, объемы данных растут в геометрической прогрессии и традиционные СУБД в составе решений Performance Management уже не в состоянии соответствовать возрастающим нагрузкам. На смену им приходят СУБД, изначально спроектированные для Big Data и новейшие поколения OSS систем, в качестве хранилищ данных, начинают использовать именно СУБД такого класса.

Предпосылки. Объем данных.

800 000 петабайт было сохранено в 2009 году во всем мире.
По прогнозам к 2020 году объем данных составит уже **35** зеттабайт.

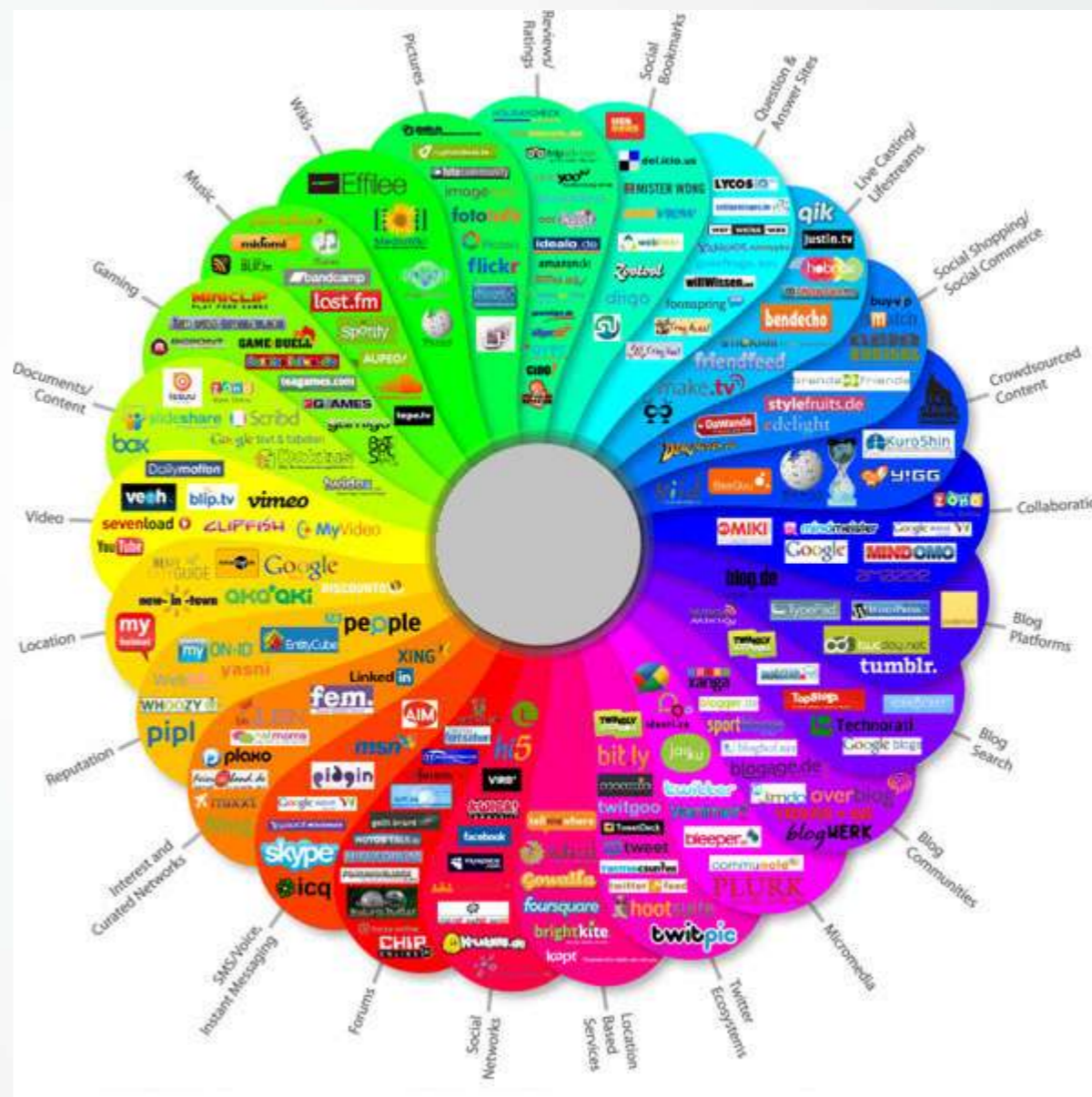
- 1 килобайт (KB) = 1024 байт
- 1 мегабайт (MB) = 1024 килобайта
- 1 гигабайт (GB) = 1024 мегабайта
- 1 терабайт (TB) = 1024 гигабайта
- 1 петабайт (PB) = 1024 терабайта
- 1 эксабайт (EB) = 1024 петабайта
- 1 зеттабайт (ZB) = 1024 эксабайта
- 1 йоттабайт (YB) = 1024 зеттабайта



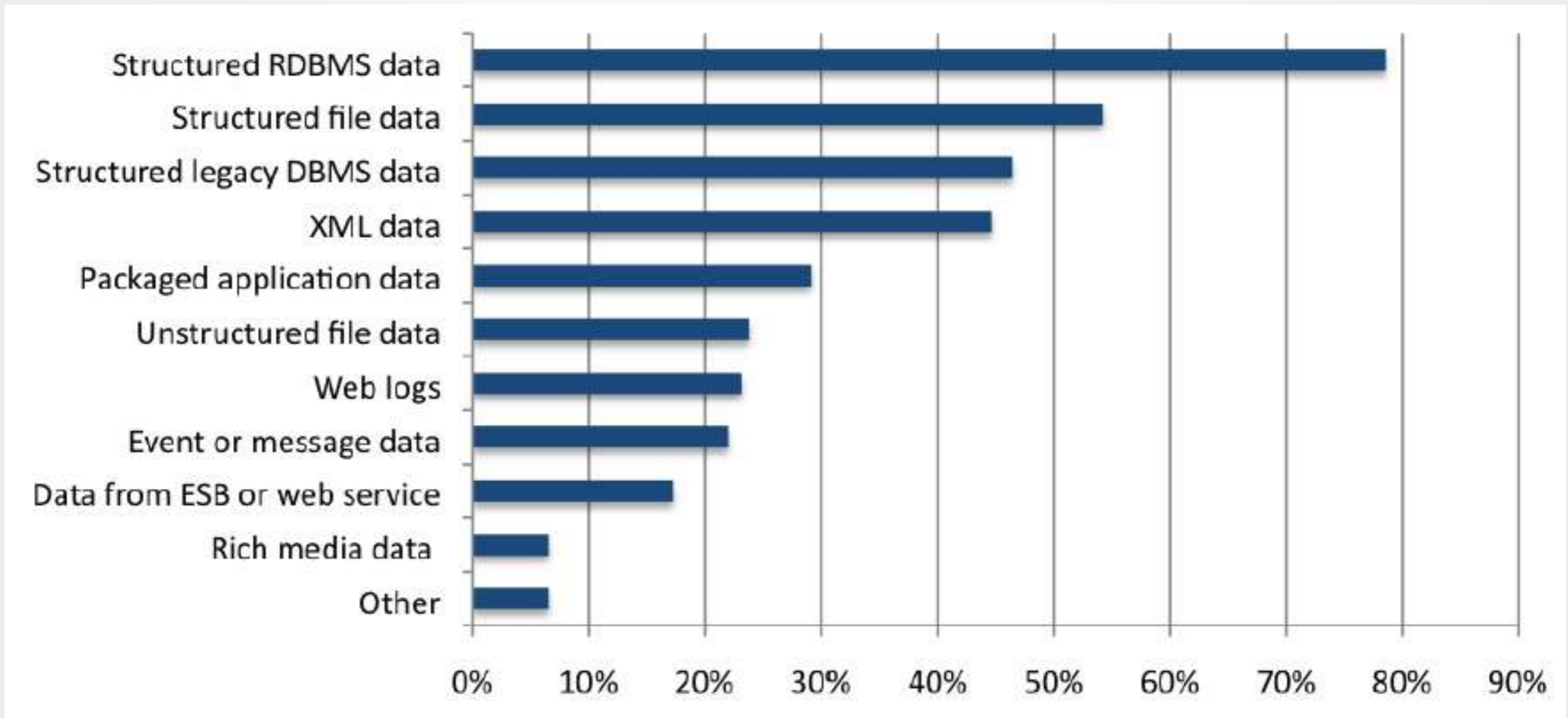
Предпосылки. Многообразие данных.

Структурированные и неструктурированные данные.

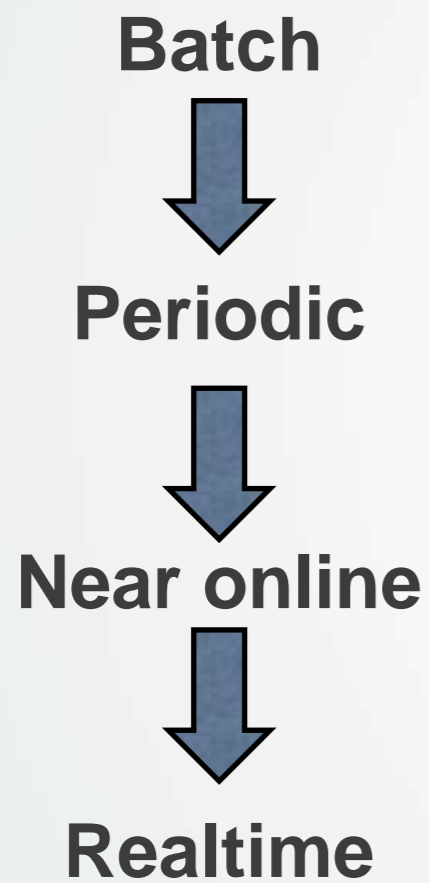
- Информация с датчиков
- Информация из логов
- CDR и т.п.
- Статистика
- Поточковый контент, аудио и видео
- Информация из социальных сетей, эл.почты, SMS
- Неструктурированные документы – формы, заявления, отчеты, отсканированные изображения



Предпосылки. Многообразие данных.

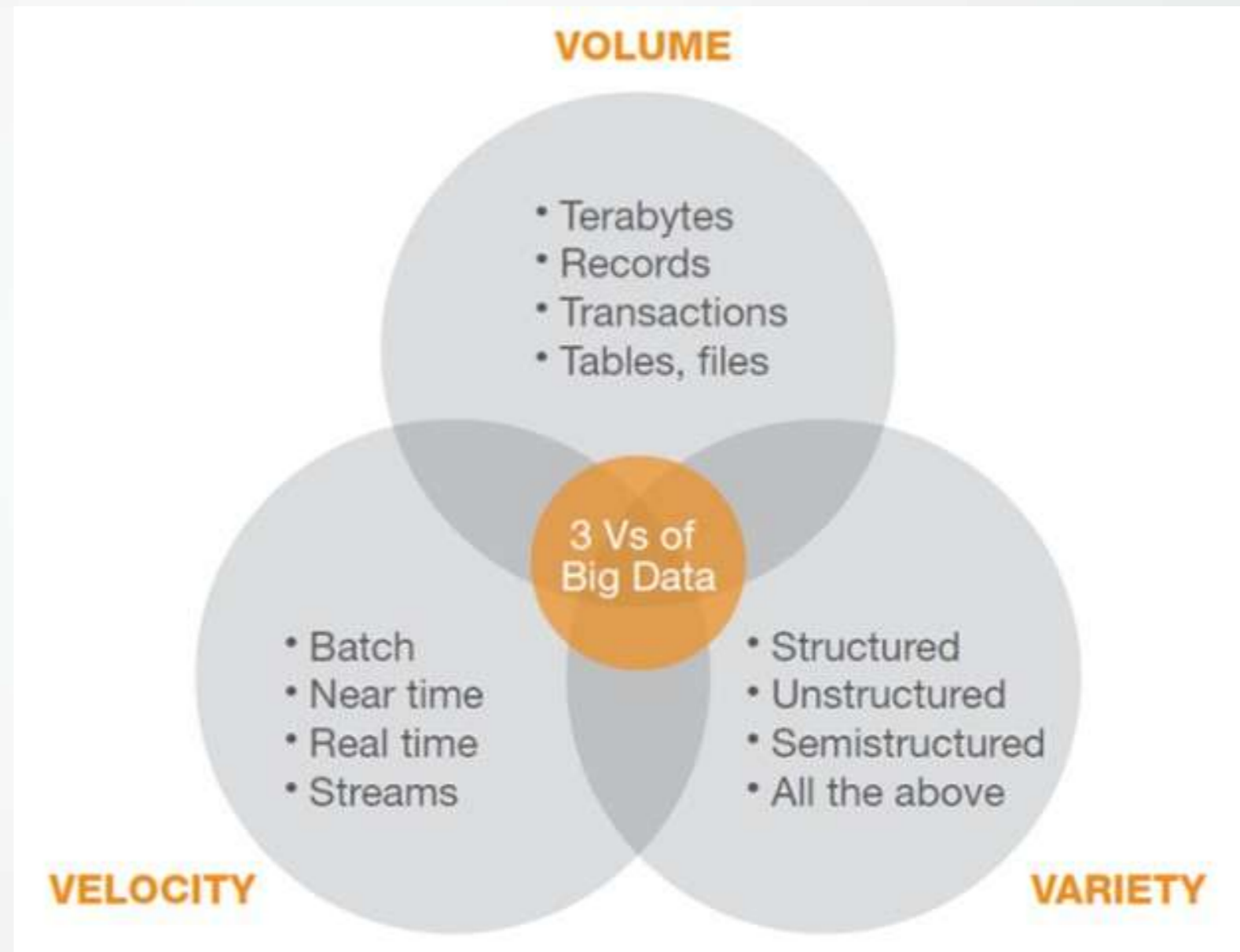


Предпосылки. Скорость обработки и анализа.



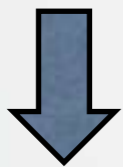
Что такое Big Data?

Массивы данных, размер и тип которых не позволяют эффективно обрабатывать и анализировать их с помощью традиционных технологий и инструментов.

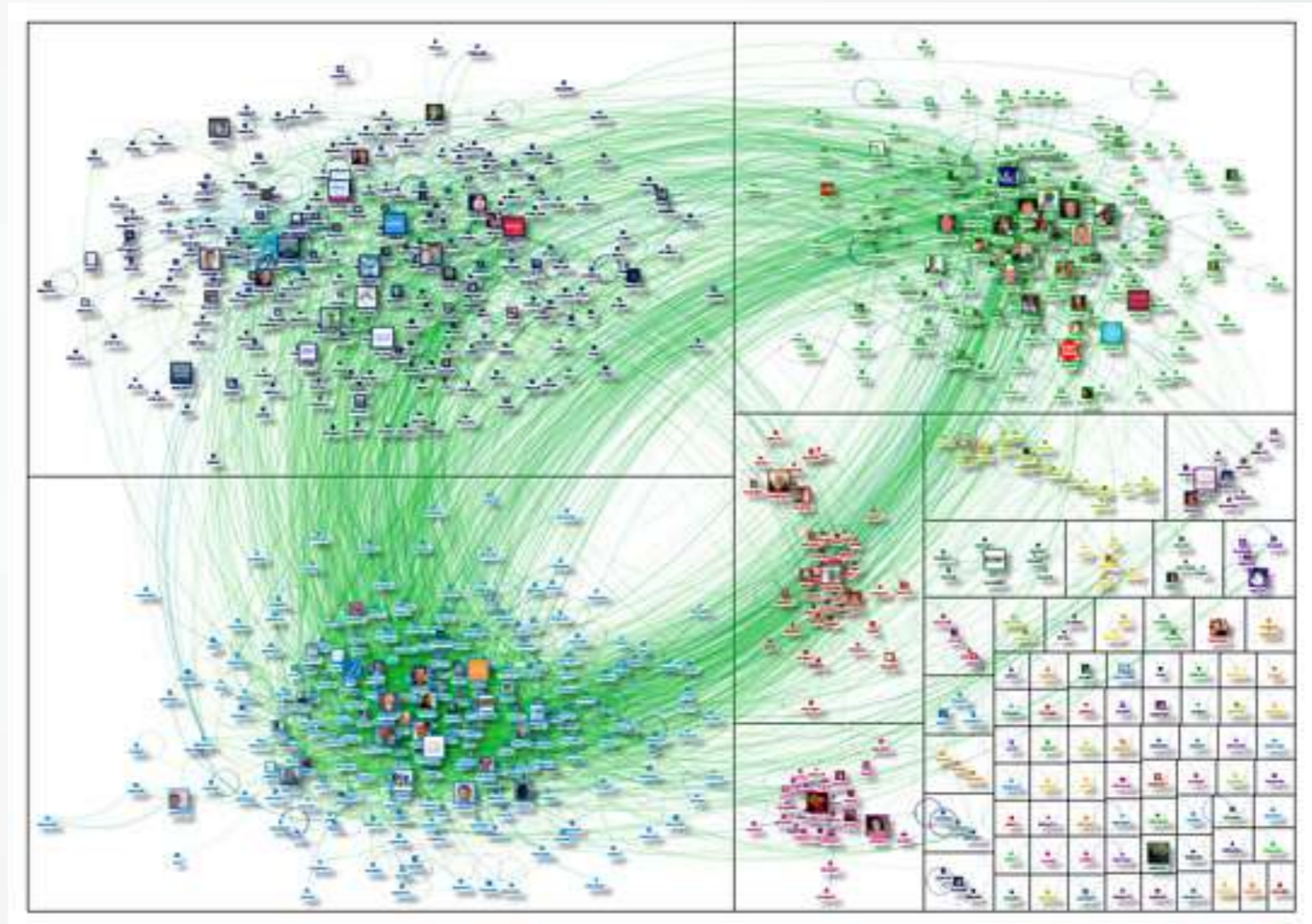


Проблемы существующих систем.

- Скорость и объемы загрузки данных
- Объемы и время хранения данных
- Хранение и анализ произвольных данных
- Скорость и объемы анализа

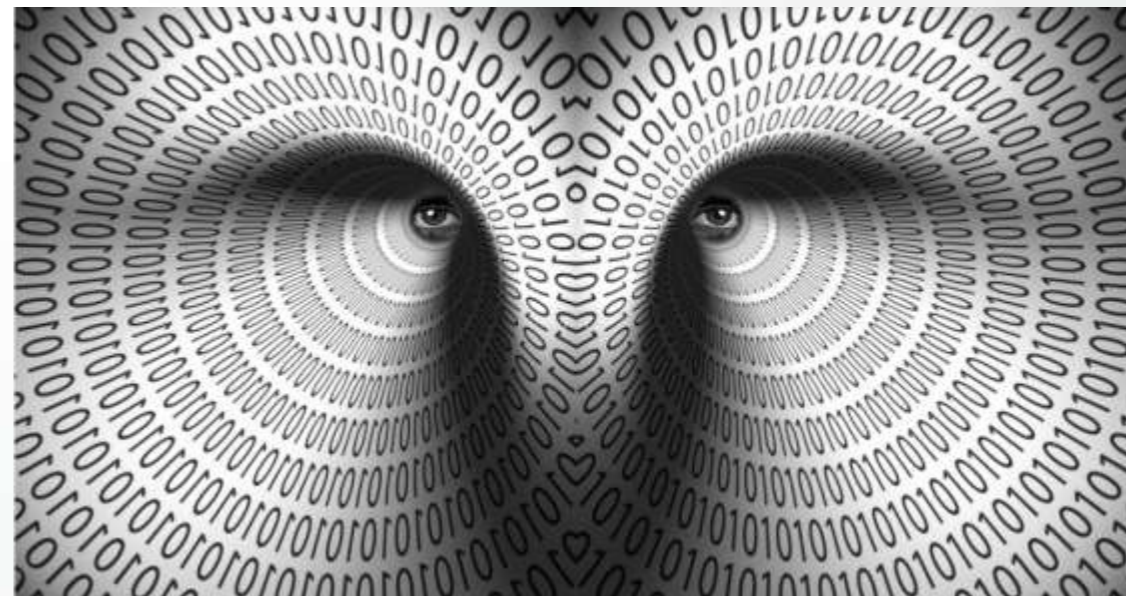


Постоянный поиск компромисса между требованиями бизнеса и возможностью их реализации



Выход?

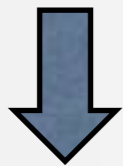
Использовать продукты новейшего поколения, изначально спроектированные с учетом требований Big Data.



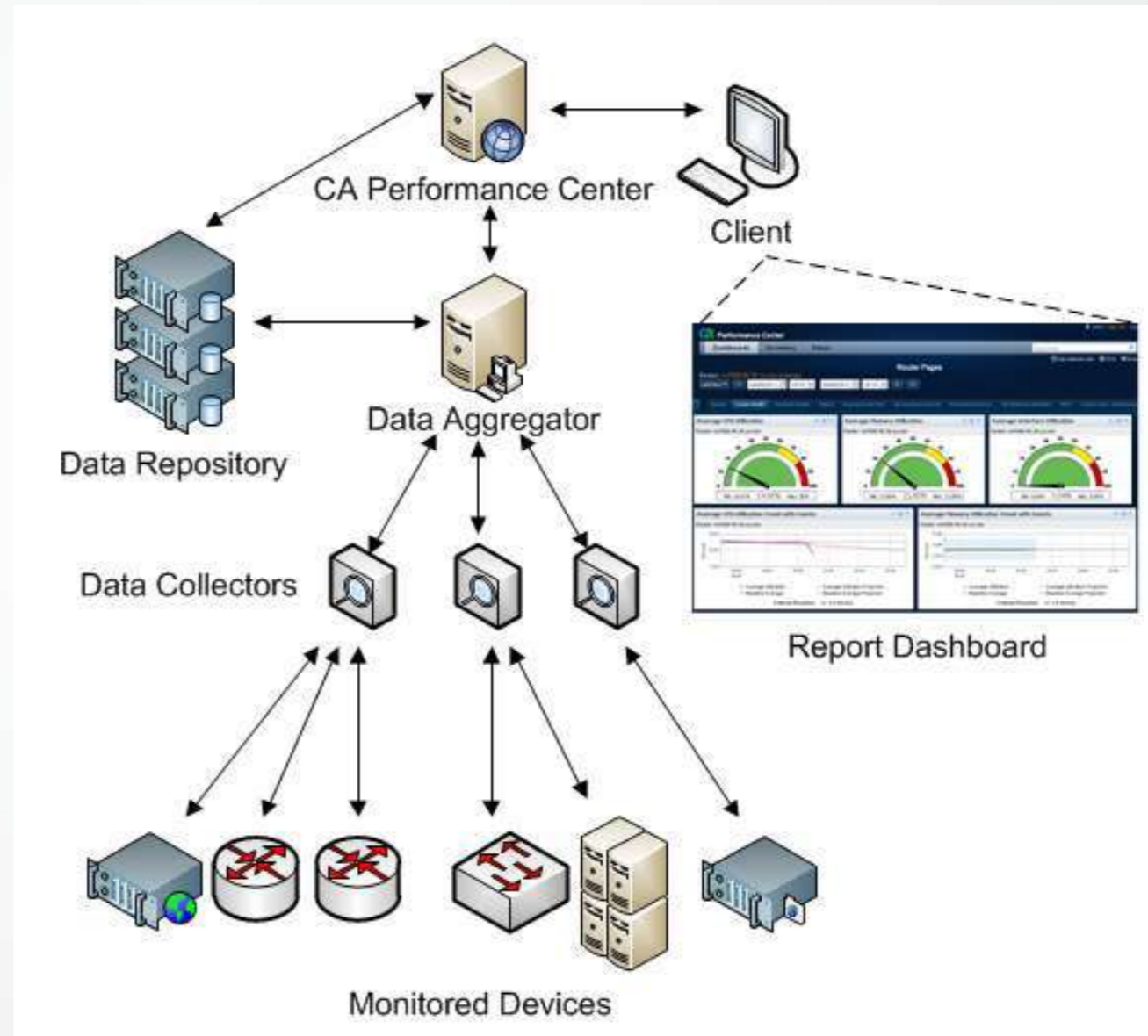
Приведены логотипы производителей ПО, решения которых используются в YOTA

CA Infrastructure Management 2.0

Oracle



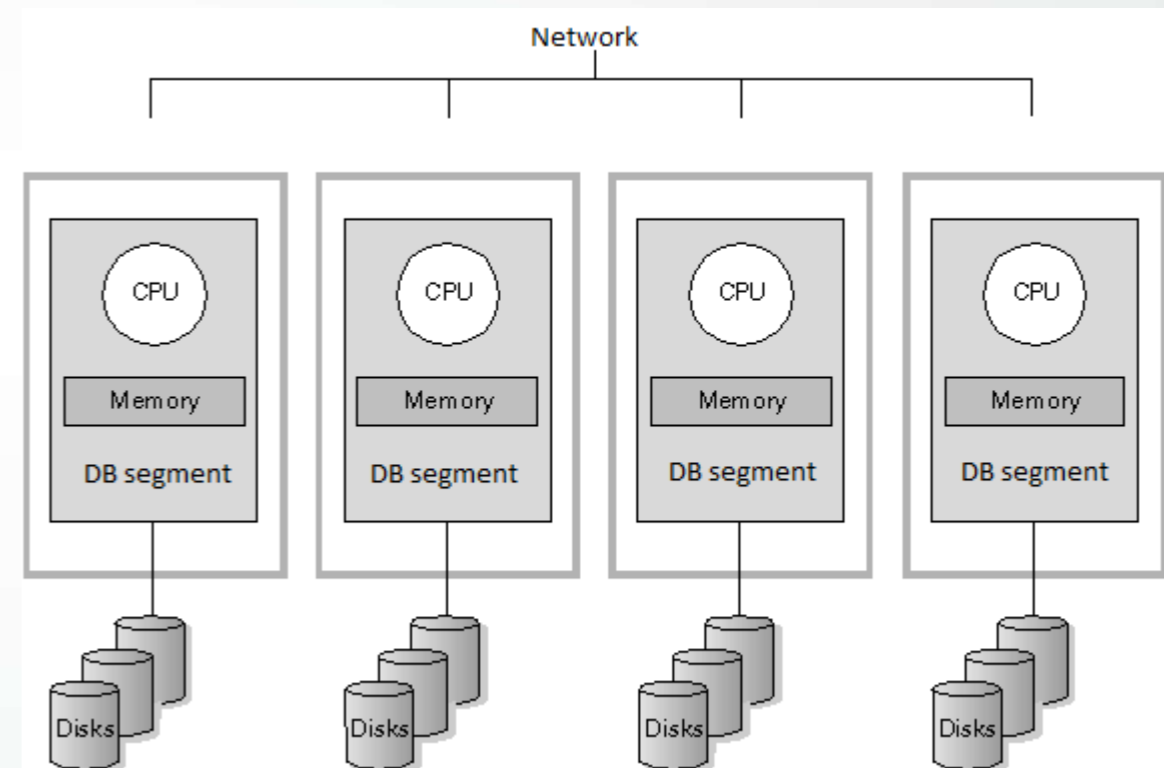
HP Vertica



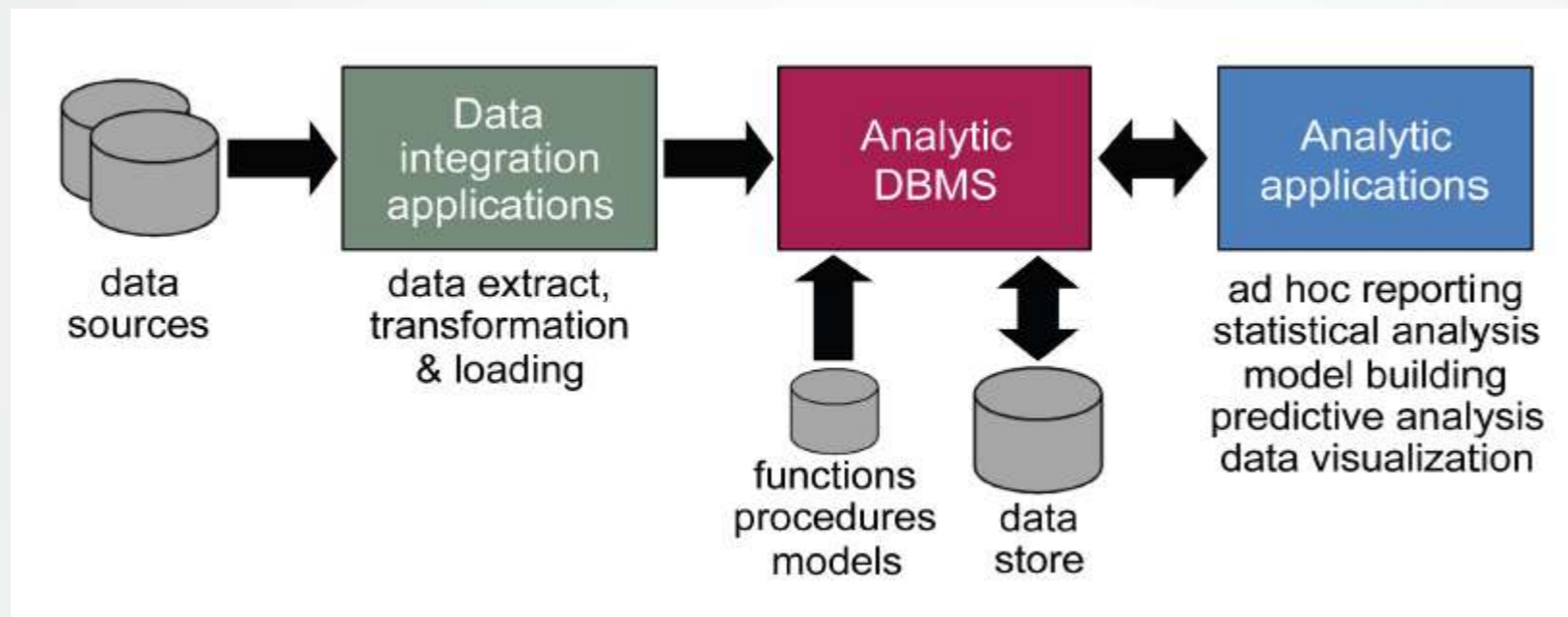
СУБД - основа Big Data

Основные принципы построения Big Data СУБД:

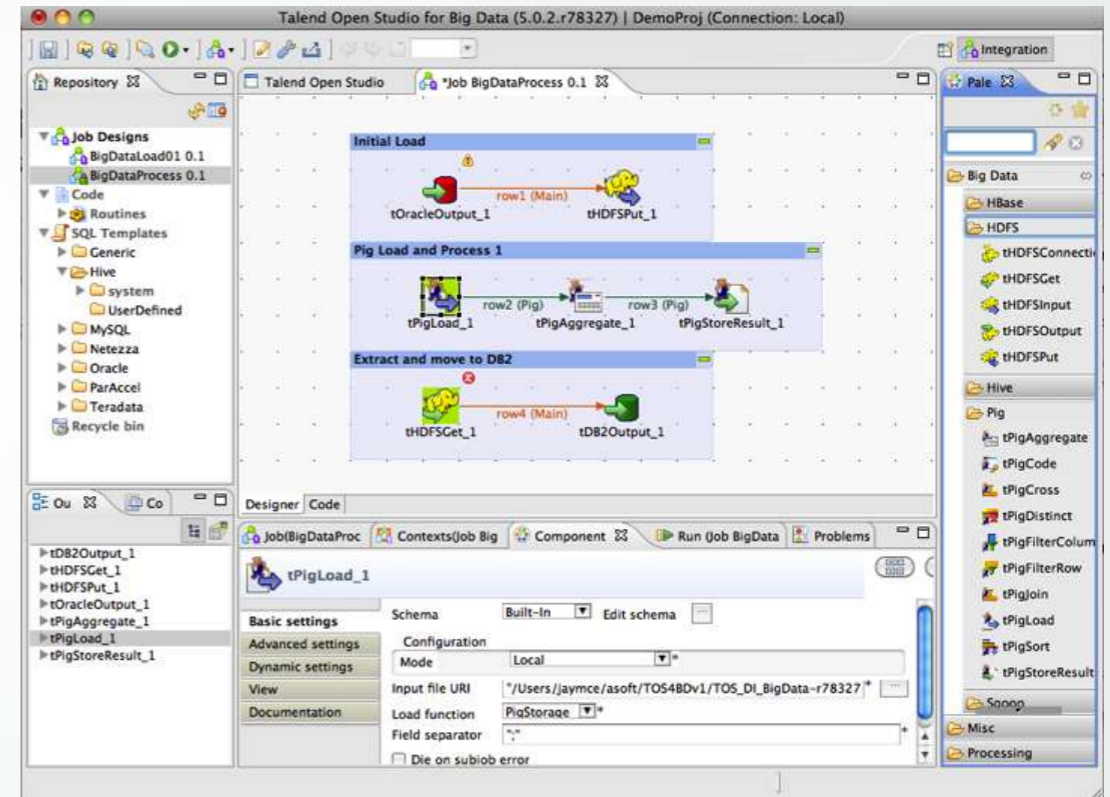
1. Архитектура массивно-параллельных вычислений MPP (Massive Parallel Processing)
2. Отсутствие разделяемых компонентов (Shared nothing)
3. Колонко-ориентированная архитектура (Column oriented)
4. Компрессия данных



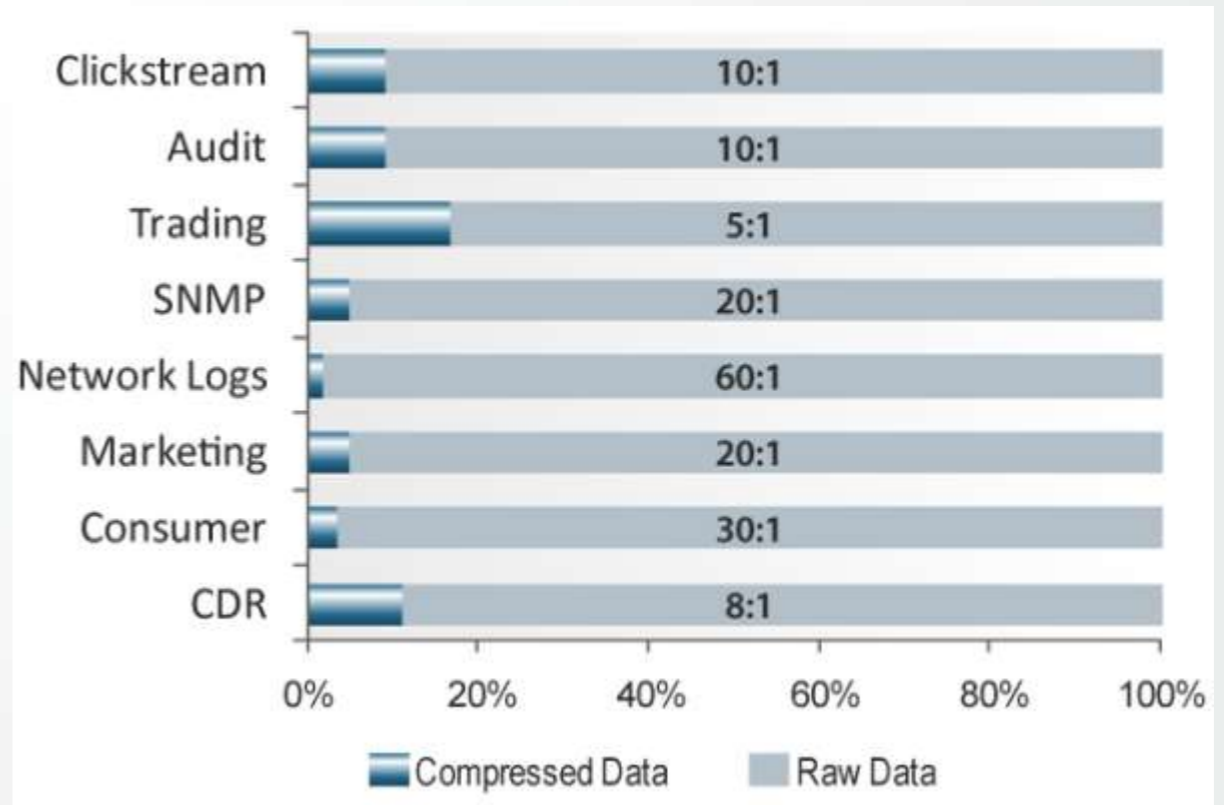
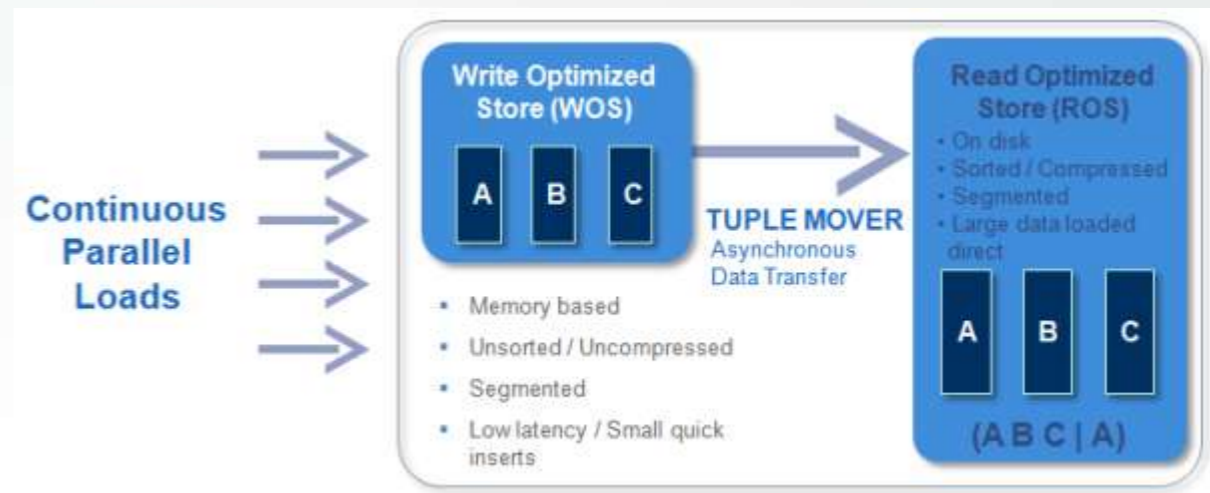
Компоненты аналитической платформы



YOTA case study: интеграция данных



YOTA case study : аналитическая СУБД



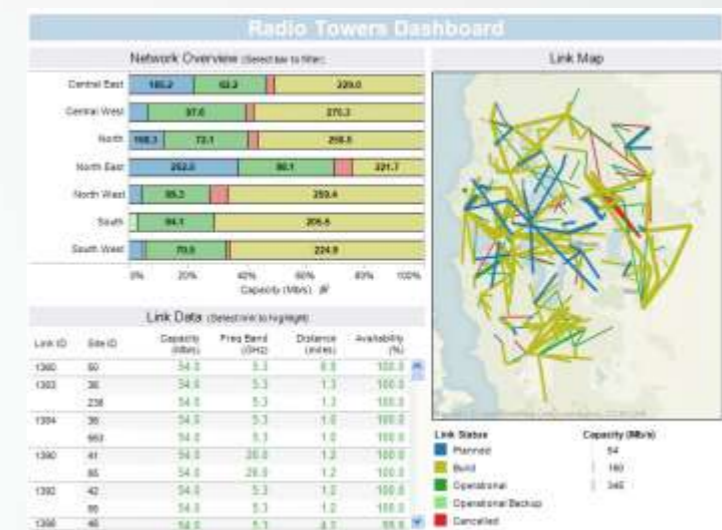
YOTA case study : визуализация и анализ данных



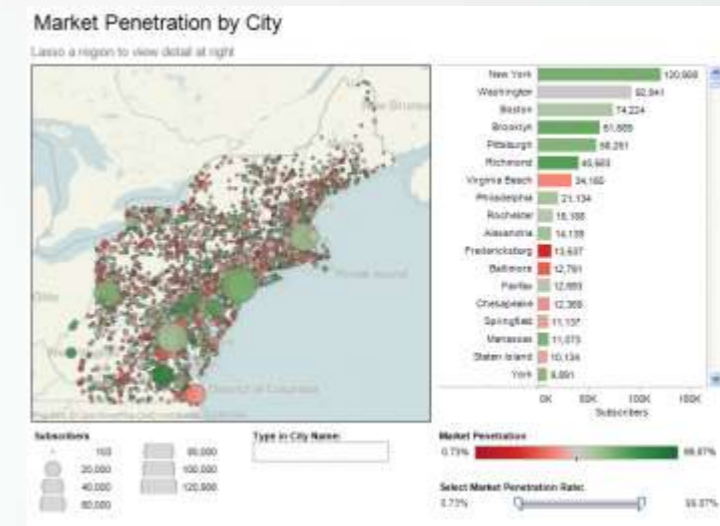
Flexibility to Choose How You Work With Data

Direct connect to fast databases

In-memory accelerates slow databases



YOTA case study : визуализация и анализ данных



YOTA case study - как начиналось

Задача: создание федерального операционного хранилища данных сетевой статистики.

Источники данных:

1. Статистика с RAN
2. Статистика с Core
3. Статистика с IP
4. Алармы

Потребители данных:

1. Технические подразделения
2. OSS (2-3 системы)

YOTA case study - что получилось

Единое федеральное DWH YOTA.

Источники данных (основные):

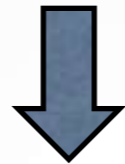
1. Практически вся статистика и алармы
2. Абонентская база (лицевые счета, тарифы и т.д.)
3. CDR
4. EDR PCRF
5. Yota Access статистика
6. Configuration Management
7. Inventory

Потребители данных:

1. Практически все подразделения
2. Более 10 ИС

YOTA case study

PoC на виртуальной платформе



Production system



3 месяца

YOTA case study

Объем данных: **5 ТБ**

Прирост: **25 ГБ в день**

Источники данных: **более 30**

Форматы данных: **csv, xml, hex, json, dblink**

Периодичность загрузки: **near realtime**

Интеграция с другими ИС: **более 10**

Одновременные подключения: **около 100**

Спасибо за внимание!



Владимир Баранов
vbaranov@yotateam.com
YOTA Networks