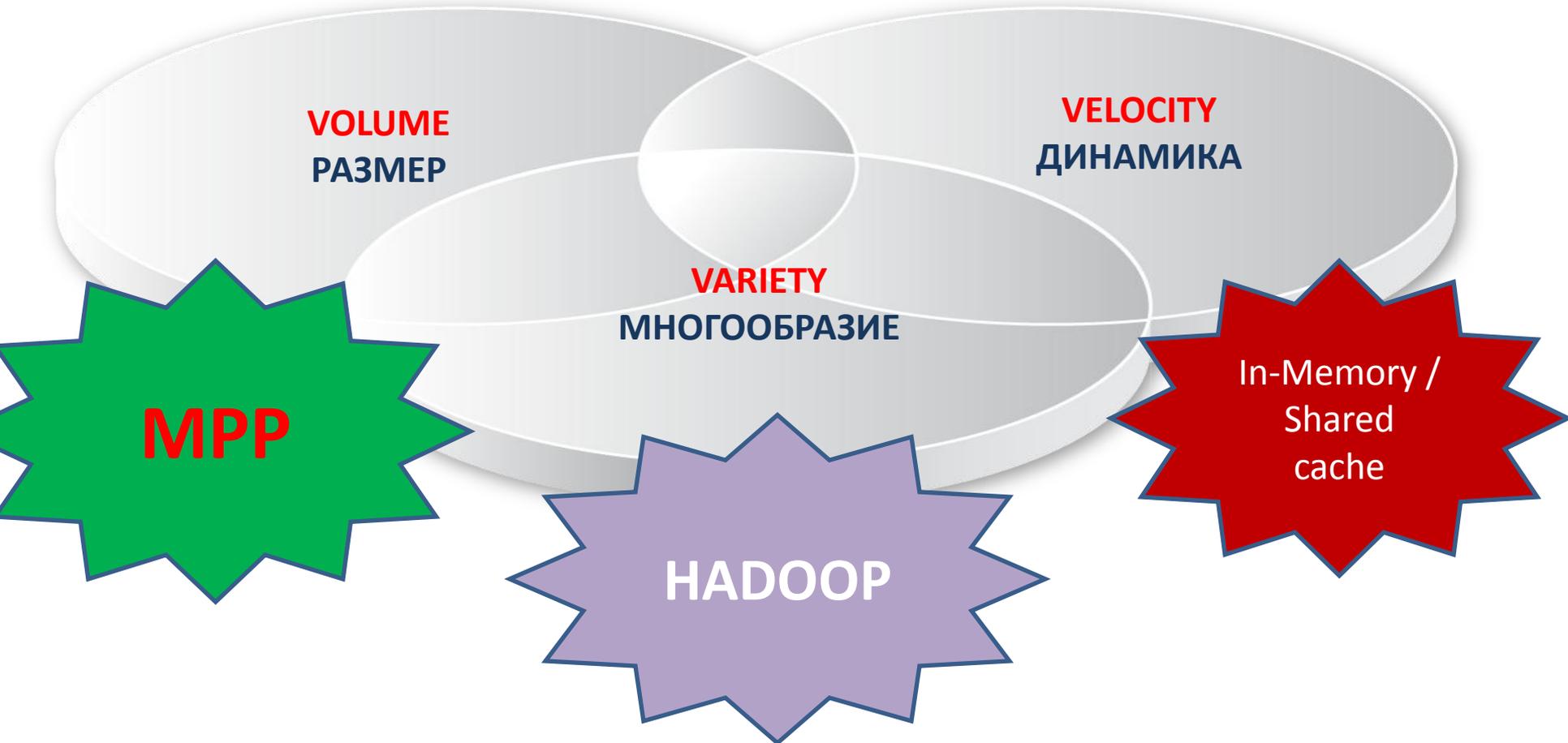


Аналитика: от отчётов – к оптимизации будущего

Ценность для бизнеса

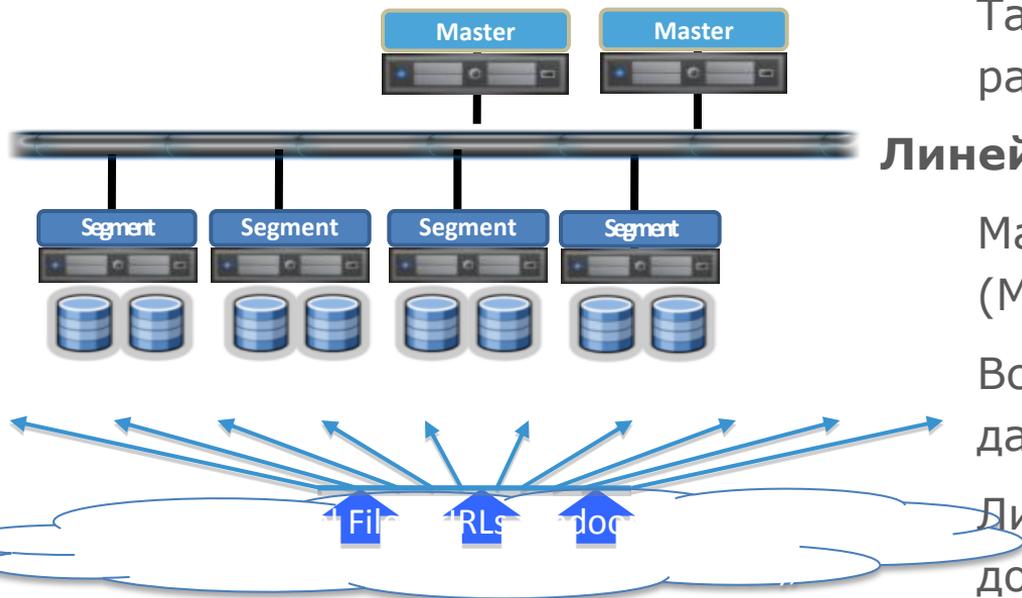


Большие Данные: Больше чем данные



Массово-параллельная архитектура без разделяемых ресурсов

SQL, MapReduce
R, Java, C, Python, Perl
ODBC, JDBC, OLEDB
BI/ETL Tools...



Оптимизирована для BI и аналитики

Простота и автоматизация

Просто загрузи и запрашивай, как любую СУБД

Таблицы автоматически распределяются по всем узлам

Линейная масштабируемость

Массово-параллельная архитектура (MPP) без разделения ресурсов

Все узлы считывают и обрабатывают данные параллельно

Линейная масштабируемость простым добавлением узлов в систему

Корпоративная Архитектура Данных



Большие Данные: Или НЕ-Большие Данные

Получаем ответы на те же вопросы – только быстрее и дешевле



Количество переходит в качество



Выбираем новые платформы-инструменты /мигрируем

Превращаем данные в информацию для анализа

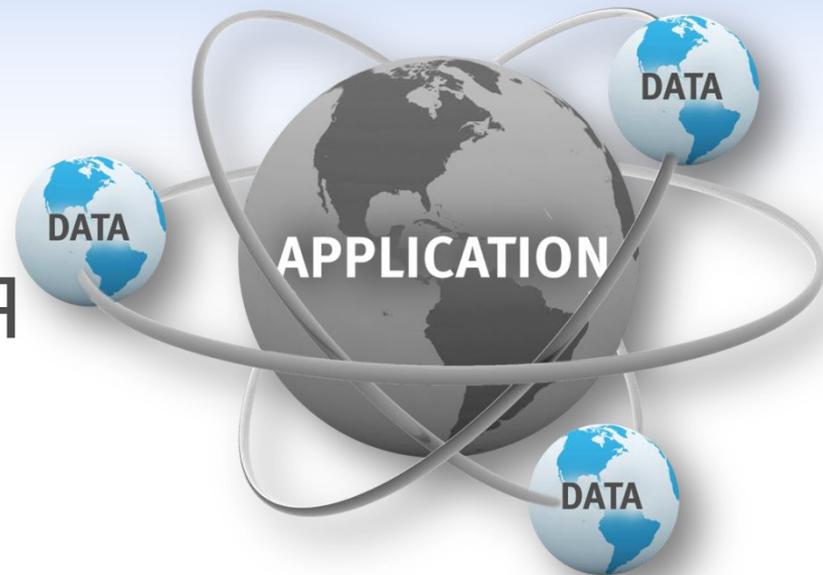


Отвечаем на новые вопросы, которые ежедневно задает жизнь....



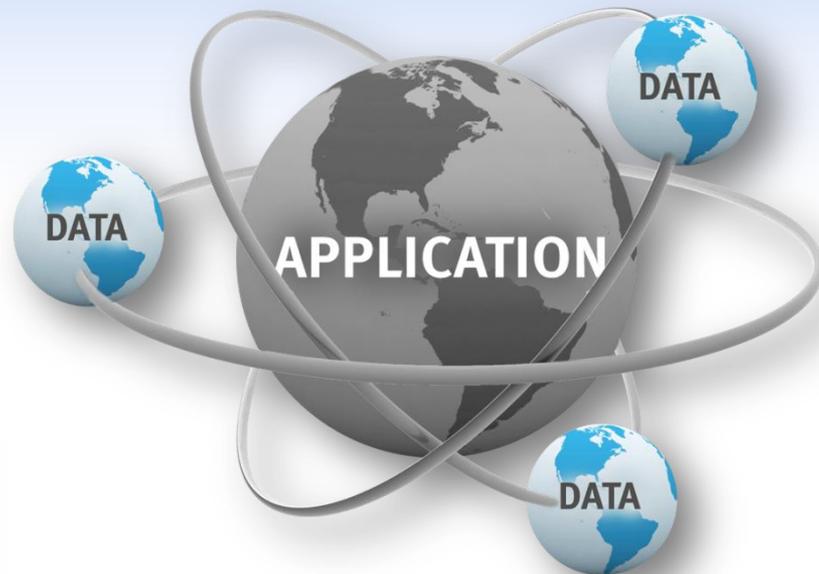
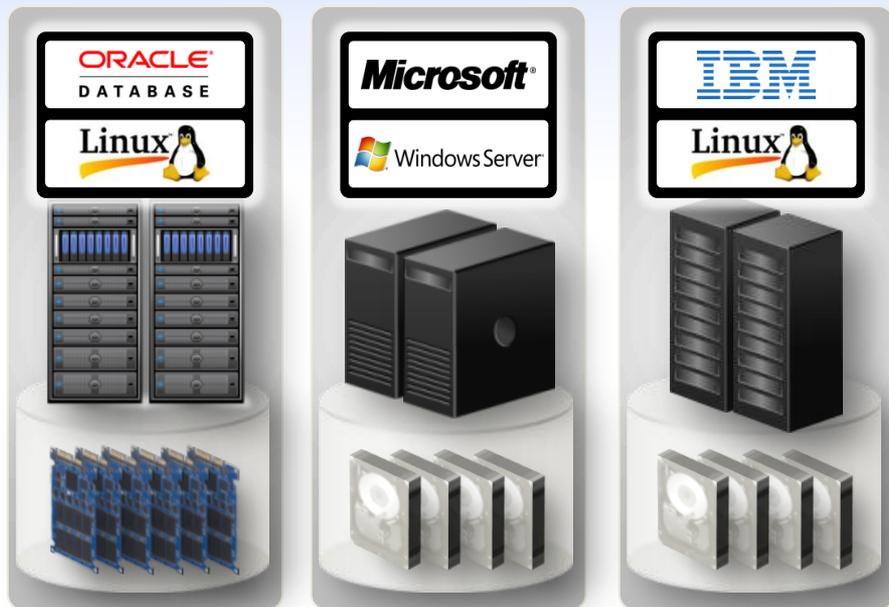
Определяем бизнес-инициативы -источники данных/строим прототип

ПРИЛОЖЕНИЯ
РАНЬШЕ БЫЛИ
ЦЕНТРОМ ПРИТЯЖЕНИЯ



Приложения > Данные

Притяжение создало выделенную инфраструктуру



ДААННЫЕ

ЭТО НОВЫЙ ЦЕНТР ПРИТЯЖЕНИЯ



Данные > Приложения

Типовые сценарии использования Больших Данных по отраслям



Интернет и телекоммуникации

Отчетность и статистика



Банки и страховые компании

Клиентская аналитика



Розничная торговля

On-line сервисы



Здравоохранение и медицина

Сценарии использования Больших Данных по отраслям



Производство

- Контроль состояния оборудования
- Контроль и анализ процессов производства



Государство

- Анти-терроризм
- Контроль траффика
- Визуализация стат.данных



Энергетика и нефтегаз

- Распределенные «умные» сети/Smart Metering
- Контроль регламентов и процессов
- Визуализация состояния сетей на ГИС системах

American Express выбирает Hadoop

Бизнес-задачи:

- Формирование профиля клиента на базе информации о платежах
- Повышение лояльности клиентов за счет программ поощрения
- Расширение методов борьбы с мошенничеством

2РВ



Expedia.com Онлайн Бронирование

Бизнес задача:

- Задержки ~8 сек. при работе сайта. Клиенты уходят.

Решение:

- Внедрены решения обработки в памяти

Эффективность внедрения:

- Ускорение в 10 раз => рост коэффициента конверсии
- Каждый 1 % роста коэффициента конверсии это: + \$60M т.к. 600,000 бронирований (в среднем \$100/ночь)
- При марже 10% рост прибыли на \$6M



Посетителей в
месяц: **90M+**



Бронирований ~\$22
млрд.



Ночей в отеле:
60M+



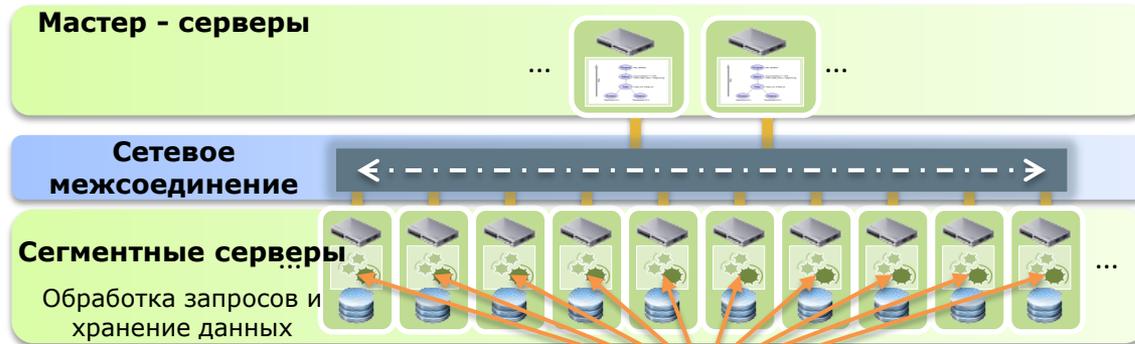
Пример внедрения



Или как предотвратить веерные отключения электроэнергии

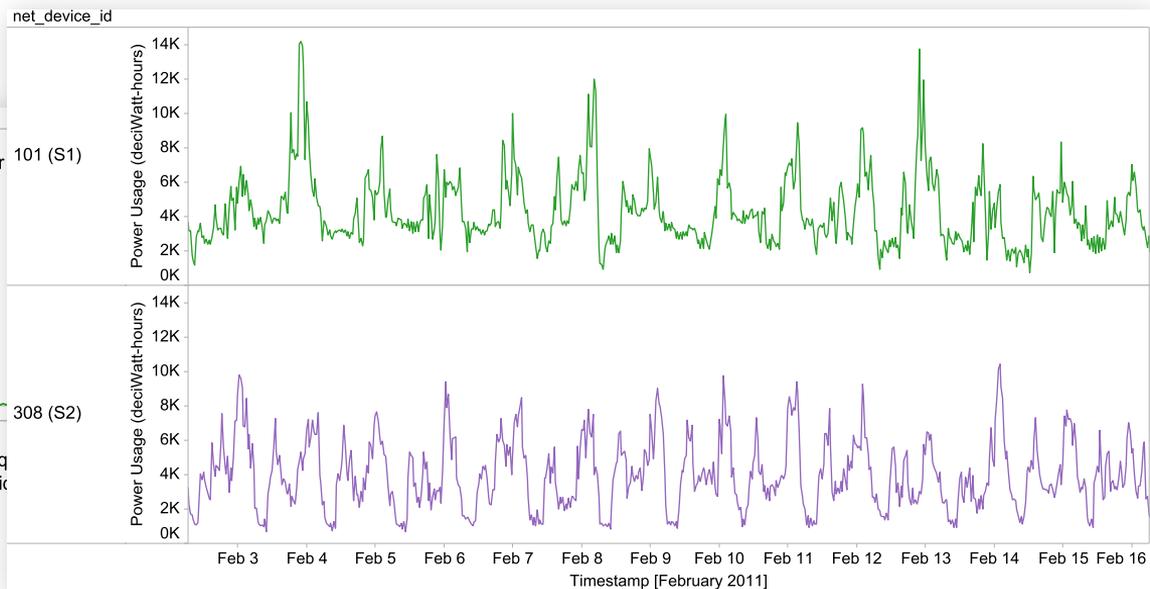
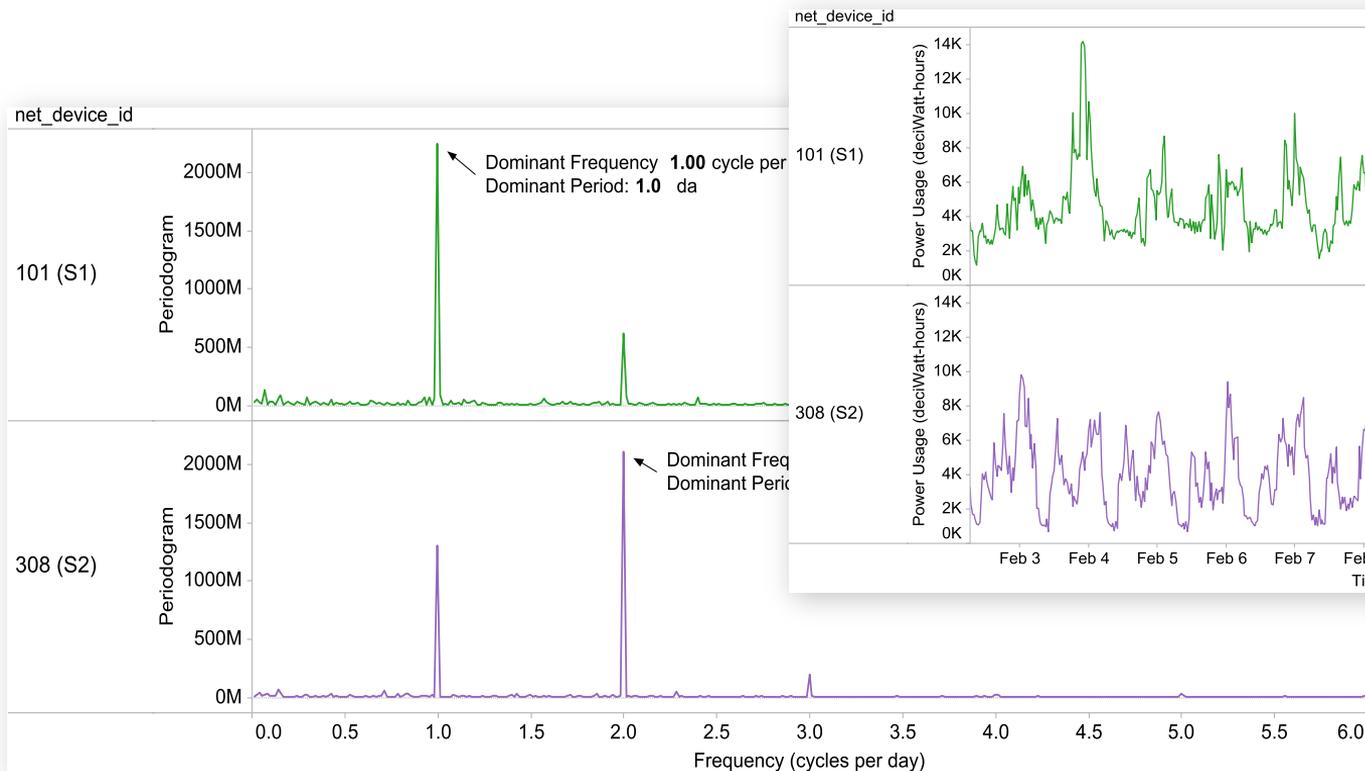
«Умные сети» и МПП СУБД

Массово-параллельная архитектура без разделения ресурсов



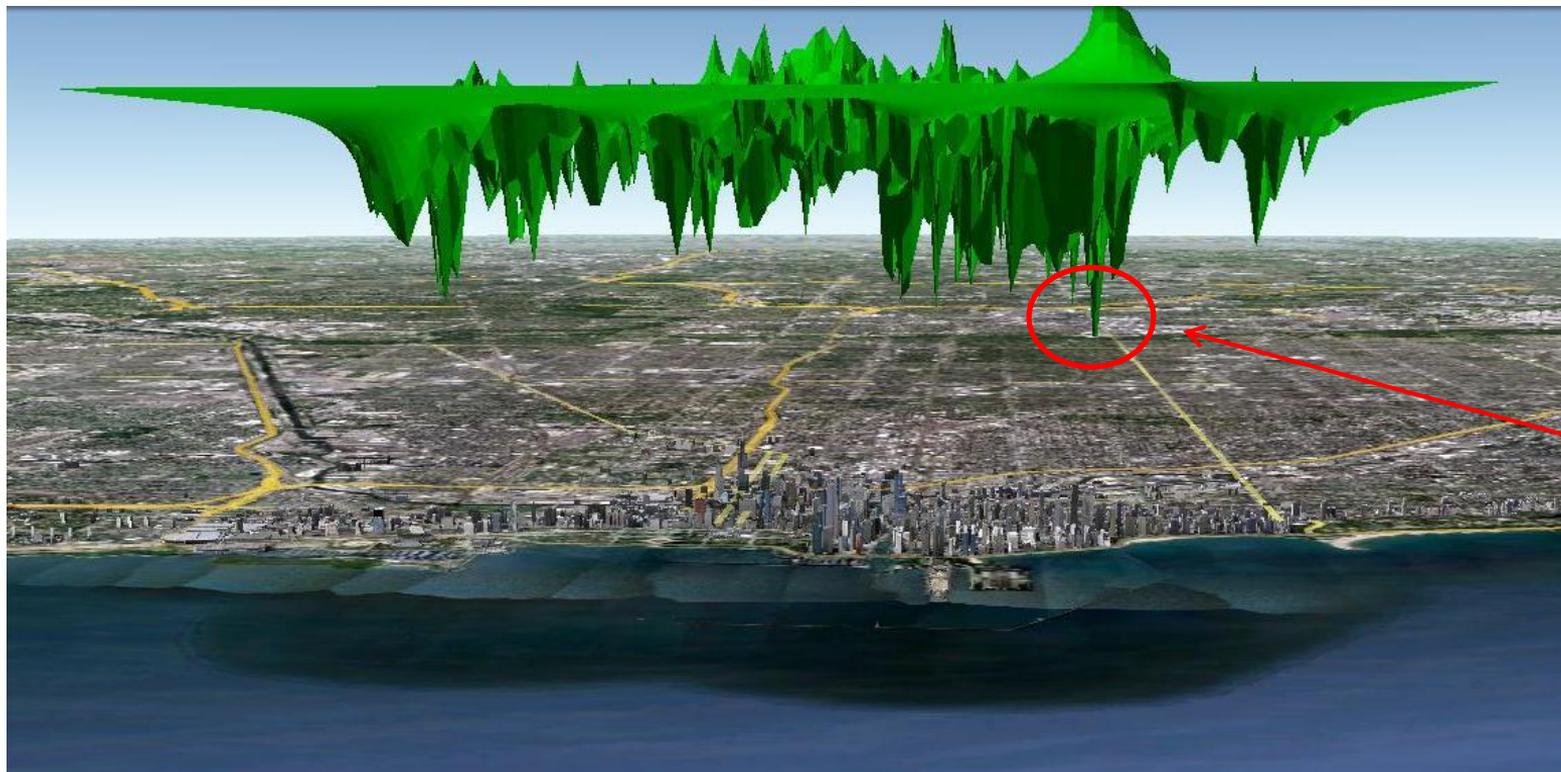
Анализ данных по энергопотреблению

с помощью преобразования Фурье



Данные за 10 недель с 100.000 счётчиков, дискретность 5 секунд.

Визуализация потребления электроэнергии С наложением на обслуживаемую территорию

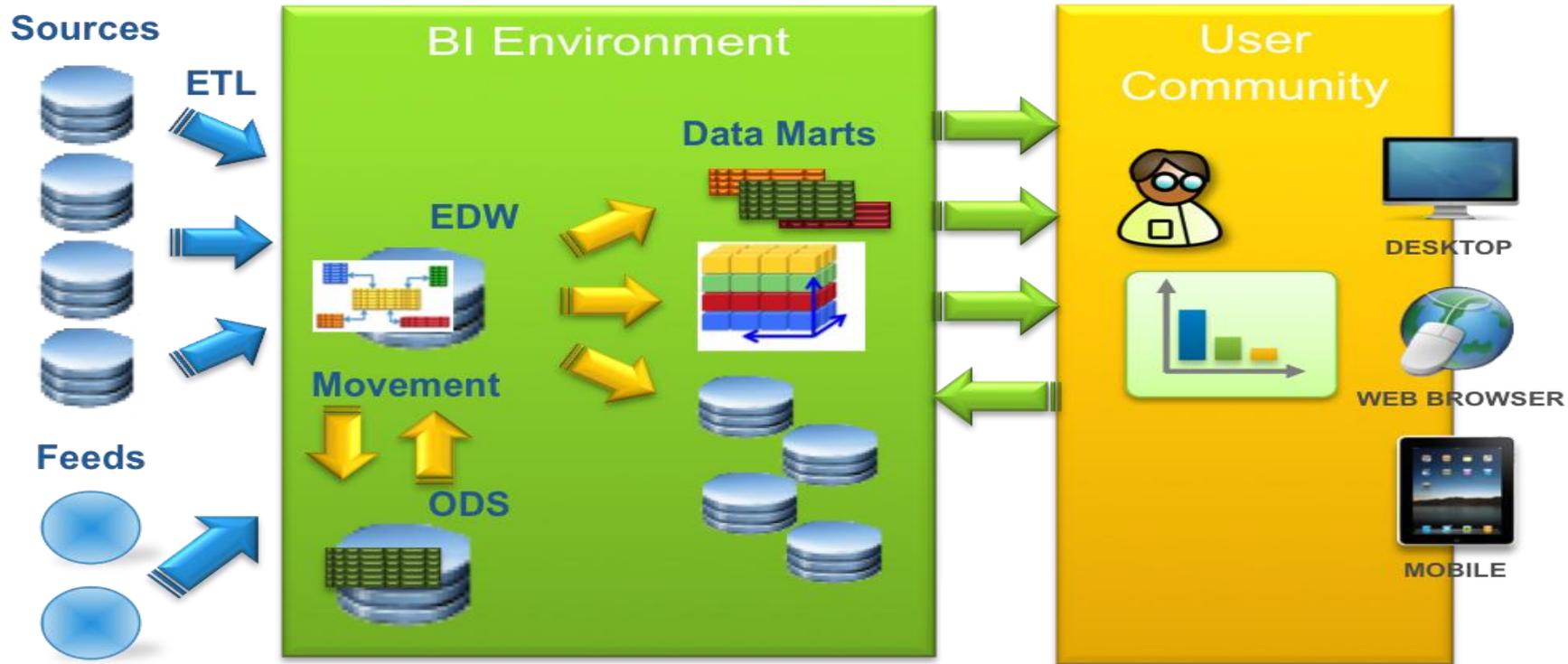


50 датчиков
показывают
напряжение
ниже
минимального
уровня.

Большие Данные для решения «повседневных»
задач ...или с чего начать



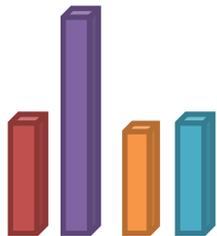
Хранилища данных и BI сегодня



Хранилища и VI сегодня



Суровая реальность :



THE ERA OF BIG DATA IS HERE

“Big Data Is Less About Size,
And
More About Fr

WIRED

FORTUNE

The New York

The
Economist

“More
than big data”

— 451 Group

“It’s Real, It’s Real-
It’s Already
Changing Your World”

—IDC