



EMC FORUM 2012
ТРАНСФОРМИРУЙ
ИТ+БИЗНЕС+СЕБЯ

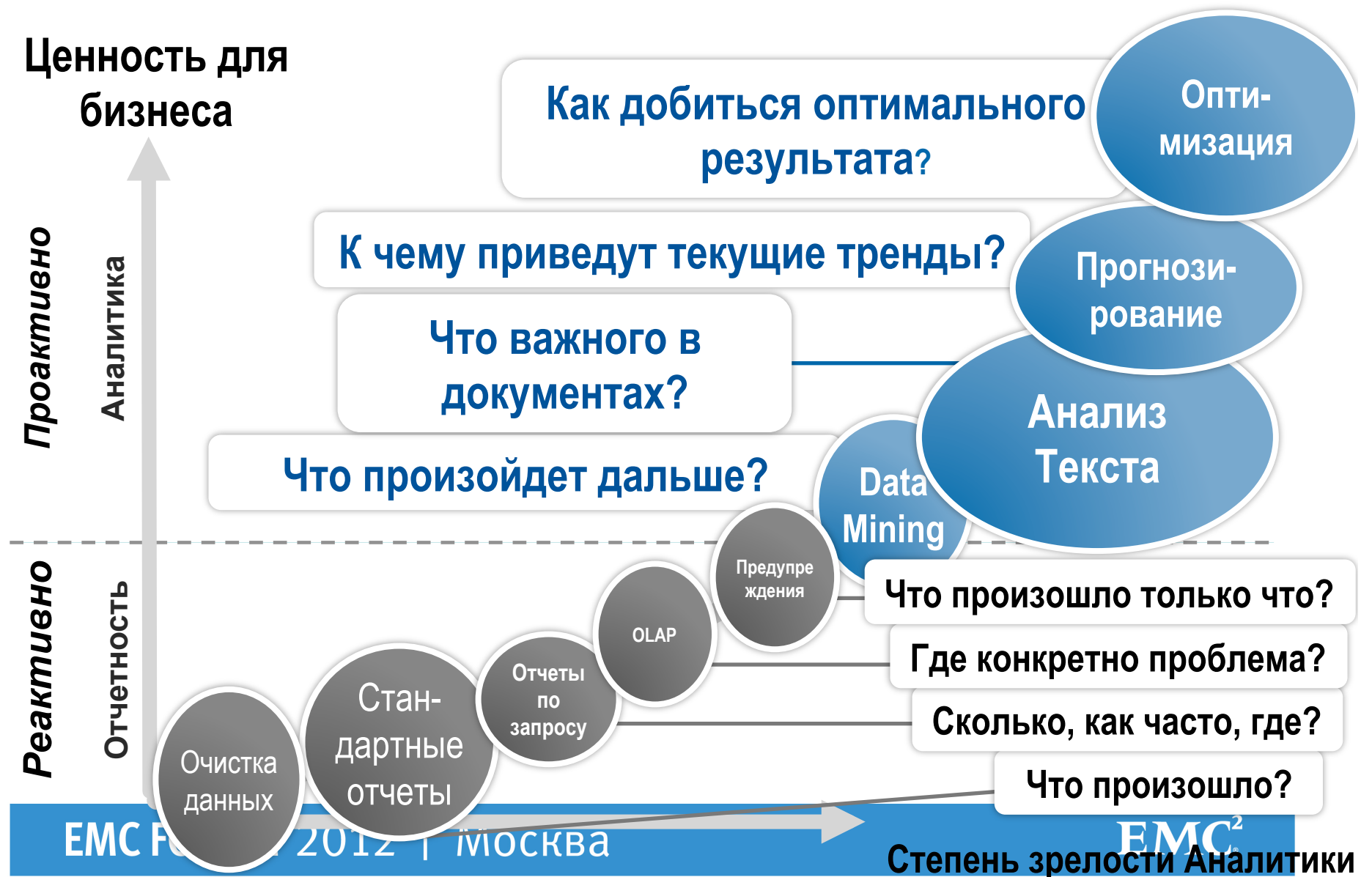
Что объединяет Большие Данные и Smart Metering.

Сергей Золотарёв
Директор по продажам Россия и СНГ,
EMC Greenplum

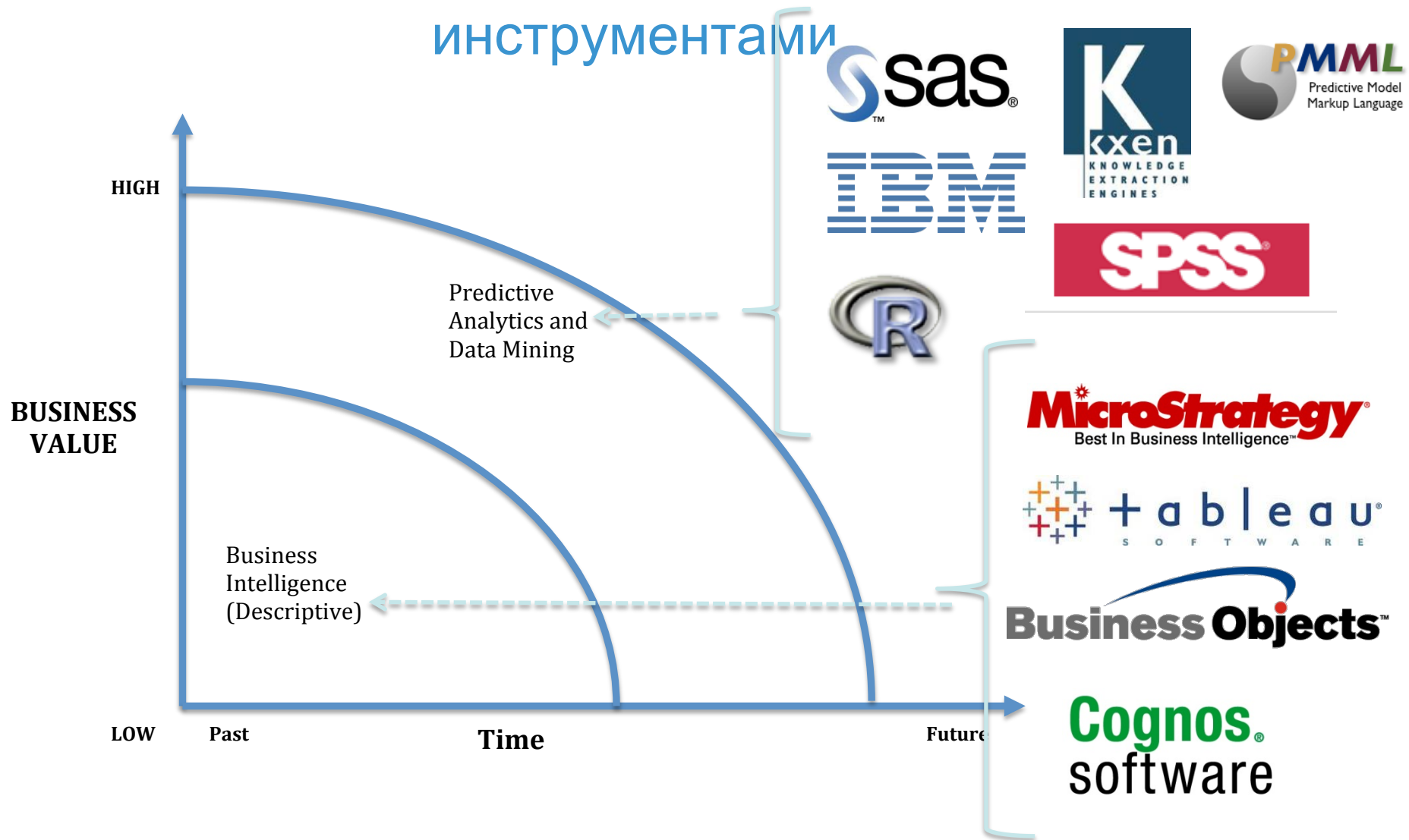
Большие Данные: Больше чем данные



Аналитика: от отчётов – к оптимизации будущего



Интеграция с различными аналитическими инструментами



Новый бизнес на Больших Данных



1,000,000,000
Запросов / День



250,000,000
Новых фото / День



290,000,000
Обновлений / День



4,000,000
Обращений / День



2,800,000,000
Сделок / День



31,000,000,000
Итераций / День

Лидеры рынка первыми начинают и выигрывают....



Типовые сценарии использования Больших Данных по отраслям



Интернет и
телекоммуникации

Отчетность и
статистика



Банки и страховые
компании

Клиентская
аналитика



Розничная
торговля

On-line
сервисы



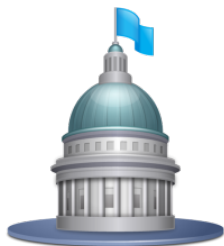
Здравоохранение и
медицина

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

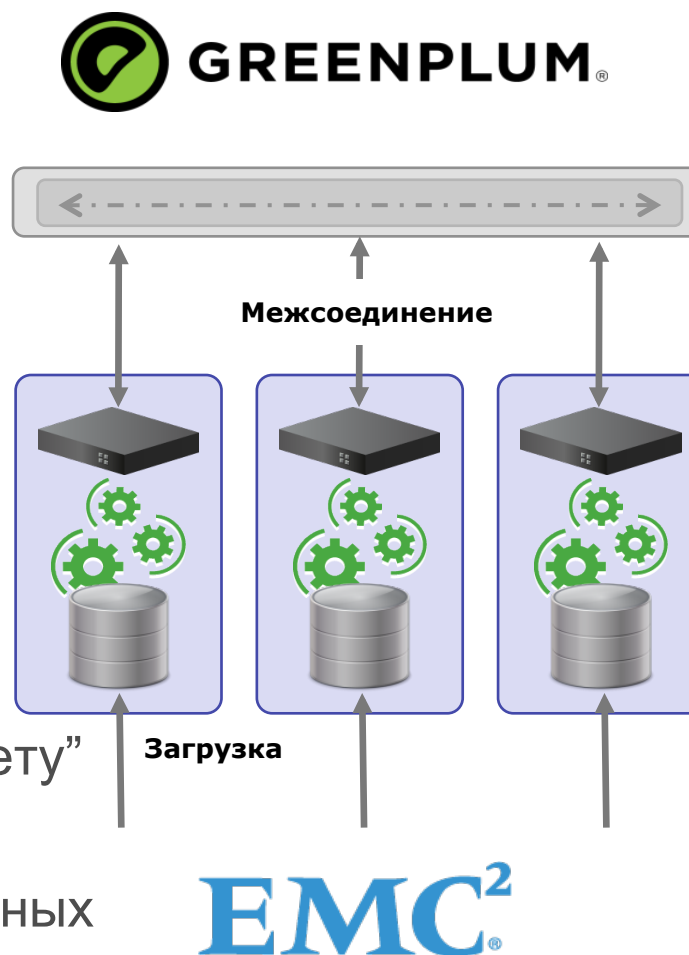
Что такое ?



Производительность и параллелизм

Горизонтально-масштабируемая архитектура, построенная на стандартных серверах

- Массивно-параллельная архитектура
- Оптимальное соотношение цена/производительность
- Минимальное время внедрения
- Линейное расширение системы до ПБ объёмов вместе с ростом скорости
- Сложные преобразования данных “на лету”
- Поддержка одновременной загрузки объёмных пакетных заданий и оперативных данных из бизнес систем



Несколько примеров из жизни Больших Данных

BIG DATA
TRANSFORMS BUSINESS

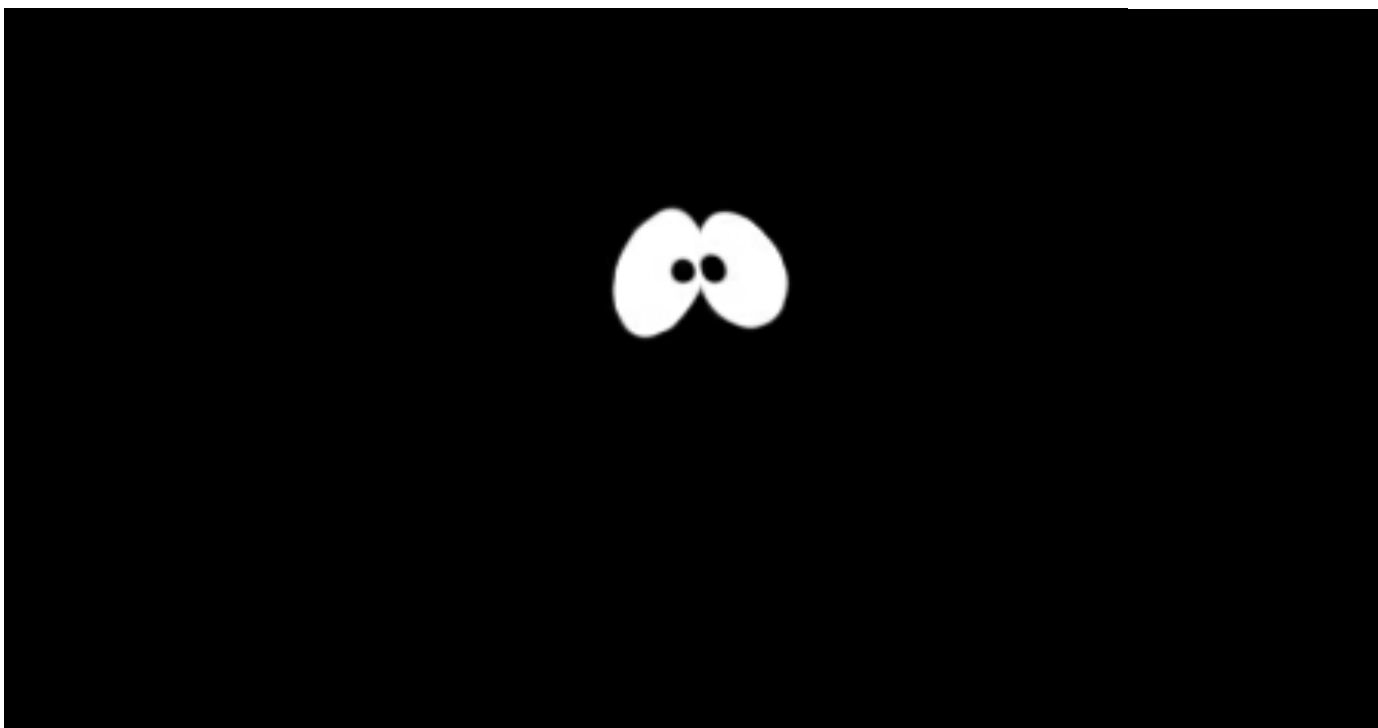


EMC²

«Умные электросети»/Smart Metering



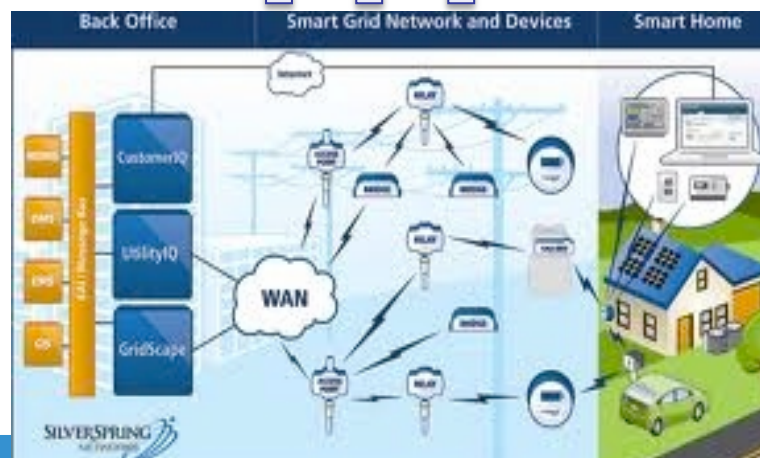
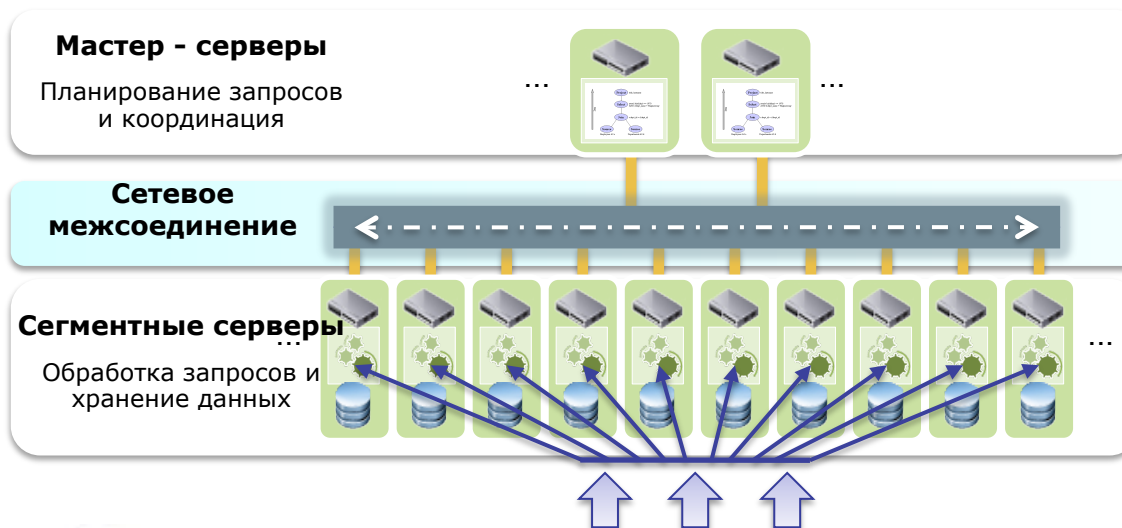
Предотвращение веерных отключений электроэнергии



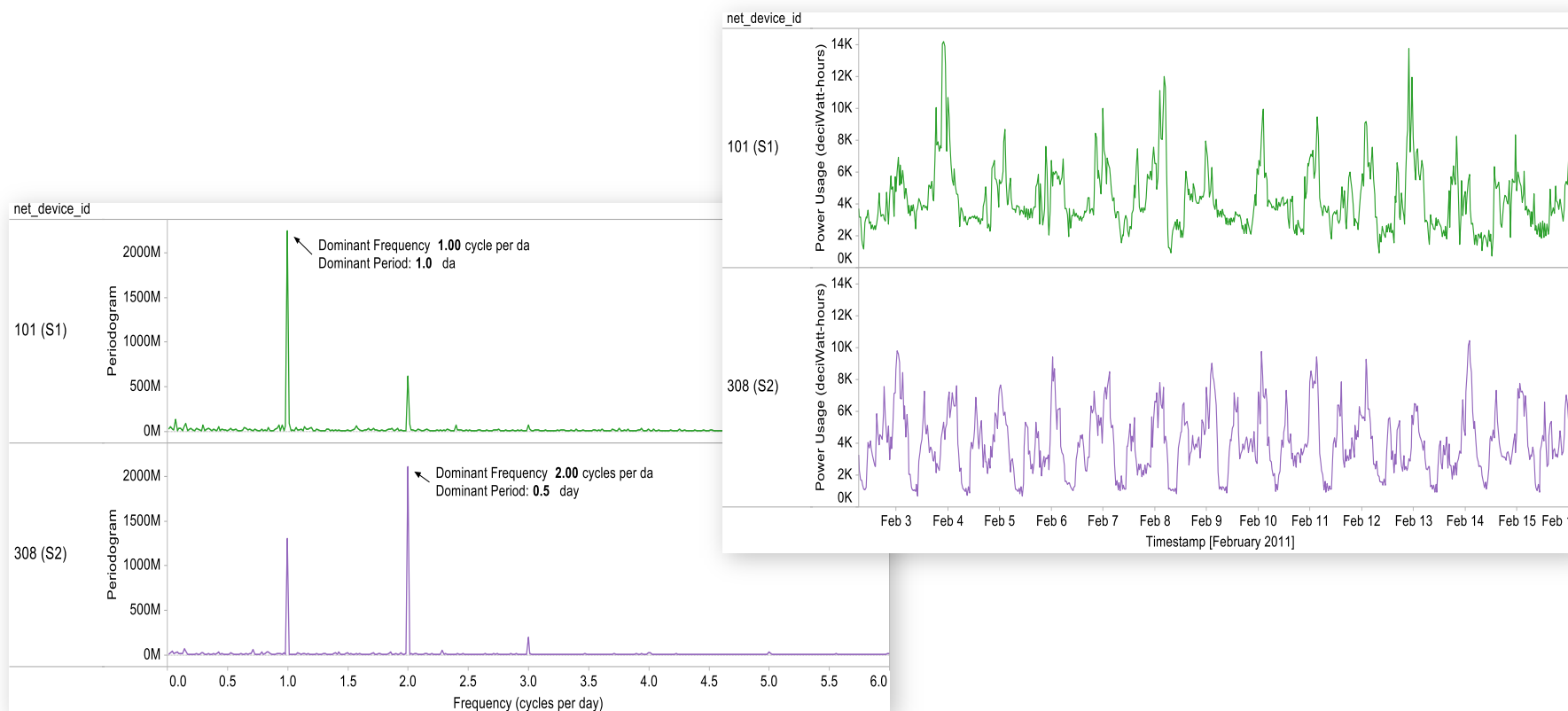
Нужно знать, каким будет потребление электроэнергии завтра... и на следующей неделе... и в следующем месяце.

«Умные сети» и EMC Greenplum

Массово-параллельная архитектура без разделения ресурсов



Анализ данных по энергопотреблению с помощью преобразования Фурье



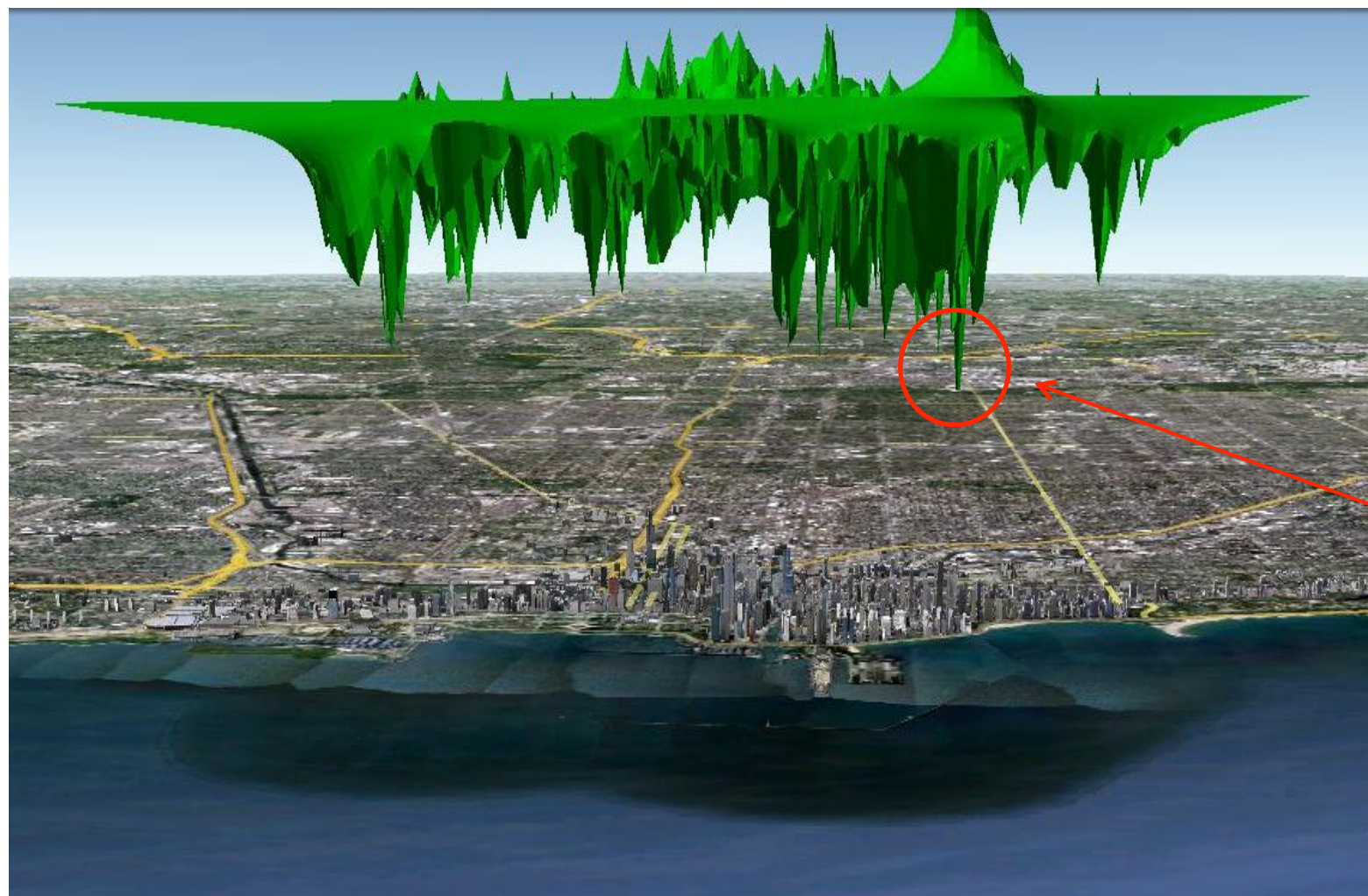
Данные за 10 недель с 100.000 счётчиков, дискретность 5 секунд.

Обслуживание на основе контроля состояния

Причинно-следственный анализ происшествий

The screenshot displays the 'Outage Restoration' software interface. At the top, there are control buttons: 'Disable Click', 'AutoRefresh', 'Refresh', 'Apply Regions', 'Time', and 'Expand'. On the left, a sidebar lists various data layers: Topography, Distribution Circuit, Congestion, Vegetation, Interruption Metric, NERC, Natural Events (with sub-items EarthQuake, Fire, Hurricanes, Live Fire, Lightning, Radar, Weather, Weather Stations), and Dynamic Alerts. The main window shows a 3D map of California with a network of power lines. A date stamp 'Aug 23, 2005 12:00 am' is visible. Three specific locations are marked: TX-6549, TX-3928, and TX-2288. A callout box points to the map with the text: «Перемотка назад» позволяет визуализировать удары молний, явившиеся причиной деградации оборудования. Below the map, there are two panels. The left panel, titled 'Info Frame', shows a line graph with a date range from 03/29/2008. A callout box over this panel states: Причинно-следственный анализ происшествий позволяет «перемотать назад» время и восстановить контекст на момент события. The right panel, titled 'Dashboard: RT Graphic', shows a detailed schematic diagram of a power system with components like 'Generators', '115kV Bus 1', '115kV Bus 2', and '115kV Bus 3'. The Windows taskbar at the bottom shows the system time as 100%.

Визуализация потребления электроэнергии С наложением на обслуживаемую территорию



50 датчиков
показывают
напряжение
ниже
минимального

Найдено
исключительно
через
Greenplum

Примеры внедрений в России



Банк: Он-лайн аналитика для он-лайн банка

Кредитные карты и вклады | Банк Тинькофф Кредитные Системы - Windows Internet Explorer
http://www.tcsbank.ru/#convenience

File Edit View Favorites Tools Help

Кредитные карты и вклады | Банк Тинькофф К...

Тинькофф
Кредитные Системы

Банк работает здесь:
[укажите ваш город](#)

Рус Eng

[О банке](#)
[Контакты](#)

Кредитные карты Вклады Дебетовая карта **Страхование** Друзья Интернет-банк

ДЕБЕТОВАЯ КАРТА
TINKOFF BLACK

- ✓ 10% cashback за покупки
- ✓ 10% на остаток по карте
- ✓ Наличные без комиссии

Подробные условия в тарифе и на странице «Дебетовая карта». Тинькофф Блэк, кэшбек

Дебетовая карта Tinkoff Black

[Кредитная карта «Тинькофф Платинум»](#)

[Вклады](#)
Если деньги нужно приумножить

[Пригласите друзей](#)
и получите до 10 000 рублей!

ОНЛАЙН ОБСЛУЖИВАНИЕ
и никаких очередей по всей России

СОВРЕМЕННЫЕ ТЕХНОЛОГИИ
обеспечивают высокий уровень сервиса

ГАРАНТИЯ БЕЗОПАСНОСТИ
ваших сбережений системой страхования

ДАЛЬШЕ ДЕЙСТВОВАТЬ БУДЕМ МЫ

Тинькофф Кредитные Системы — первый и единственный в России банк, обслуживающий **более 2 000 000 клиентов** без отделений по всей России от Калининграда до Владивостока.

Ваше общение с банком происходит только по удобным для вас удаленным каналам обслуживания:

3:42 PM
10/3/2012

Банк: Хранилище данных для корпоративной отчетности



- **Компания:** Тройка Диалог
- **Проект:** Хранилище Данных для Отчётности/2010 год
- **Внедрение позволило:**
 - Ускорение отчётности. Запросы в 10-100 раз. Результаты за минуты, не дни.
 - Разгрузка БД бэк-офисной системы
 - Линейно-масштабируемая платформа для развития нового функционала, например алгоритмической торговли, системы мониторинга нормативных требований, количественного анализа
 - Оптимизация затрат на ИТ инфраструктуру

Телеком: Снижение оттока абонентов

Моделирование поведения абонентов осуществляется на основе математической теории и практики анализа социальных сетей

Накладывая на построенную модель информацию о происходящих изменениях рассчитывается вероятностный прогноз персонального поведения абонентов



**Снижение оттока
на 20-25%**

Большие Данные : Не только технологии

- Сотрудничество с ведущими мировыми университетами Berkley, Stanford

- Поддержка открытых проектов

- Собственная среда для аналитиков Chorus

- Команда аналитиков Data Science Team
отраслевая экспертиза в сочетании со знанием инструментов матстистики и языка программирования



КАК СДЕЛАТЬ TEST DRIVE ?



- В тестовой зоне EMC в Москве
- В лаборатории EMC (Корк, Ирландия) – возможность моделирования комплексных гетерогенных систем

	Greenplum DB ½ стойки
Number of Servers	8
Total Number of Cores	96 cores
Total Memory	384 GB
Storage Type	600 GB
Total Number of Storage Drives	86
Useable Capacity (uncompressed)	18 TB
Useable Capacity (compressed)	72 TB
Scan Rate	12GB/Sec
Data Load Rate	5TB/Hour

THE ERA OF BIG DATA IS HERE

“Big Data Is Less About Size, And More About Freedom”

—The

The New York Times

WIRED

FORTUNE

The Economist

For

“Total data bigger than big data”

— 451 Group

...! It's Real, It's
...time, and It's Already
...Changing Your World”

—IDC

EMC FORUM 2012 | Москва

EMC²