



РОССИЙСКИЕ ИНТЕРНЕТ-ТЕХНОЛОГИИ
Высокие нагрузки

Масштабирование системы баннерной рекламы с централизованной базой данных



Сергей Нековаль
Компания «Грамант»



Что масштабируем?

- • **Японская система баннерной рекламы**
более 450 млн. показов в сутки
- **Система на базе DART Enterprise**
более 50 млн. показов в сутки

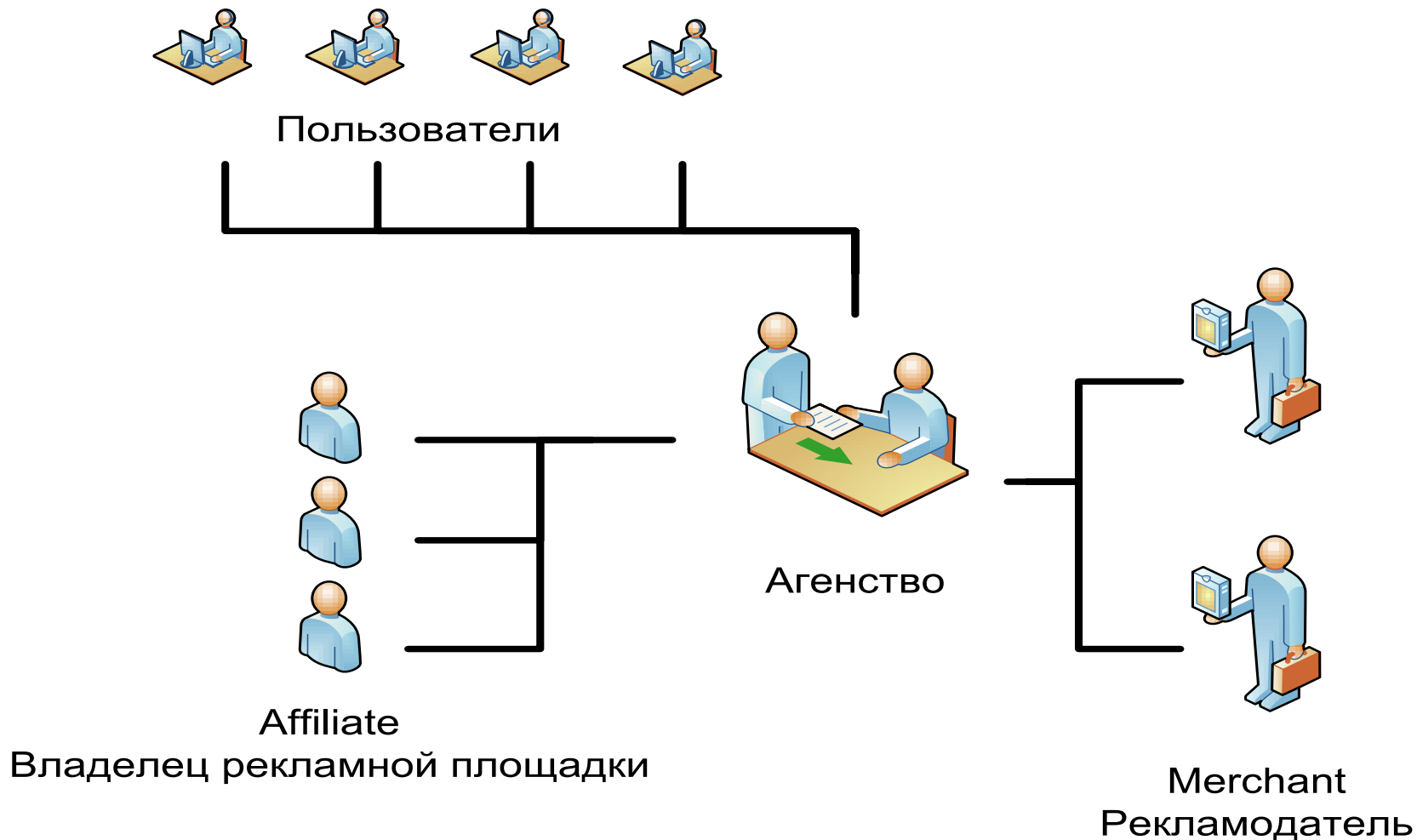


Что рассказываем?

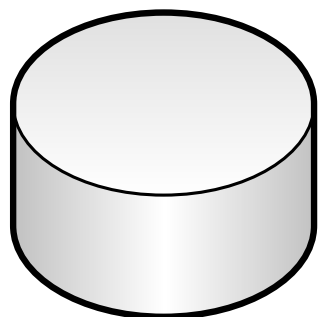
- Эволюционное развитие архитектуры системы или *Как мы докатились до такой жизни*
- Выгоды и проблемы централизации данных
- Capacity Planning и риски



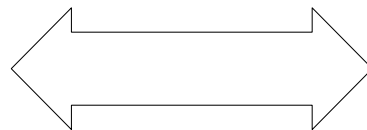
Бизнес-модель



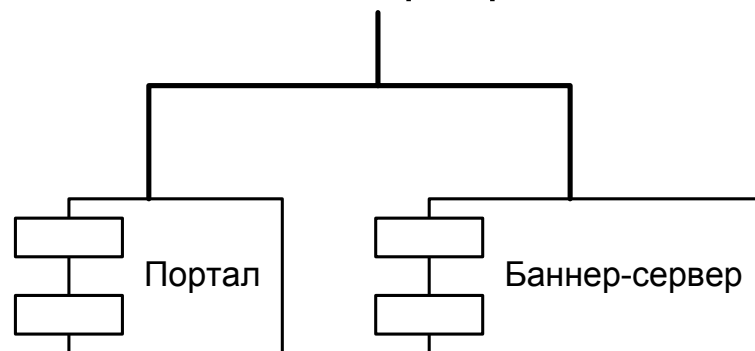
Начало пути



Oracle



Веб-сервер



- СУБД Oracle
- Веб-сервер

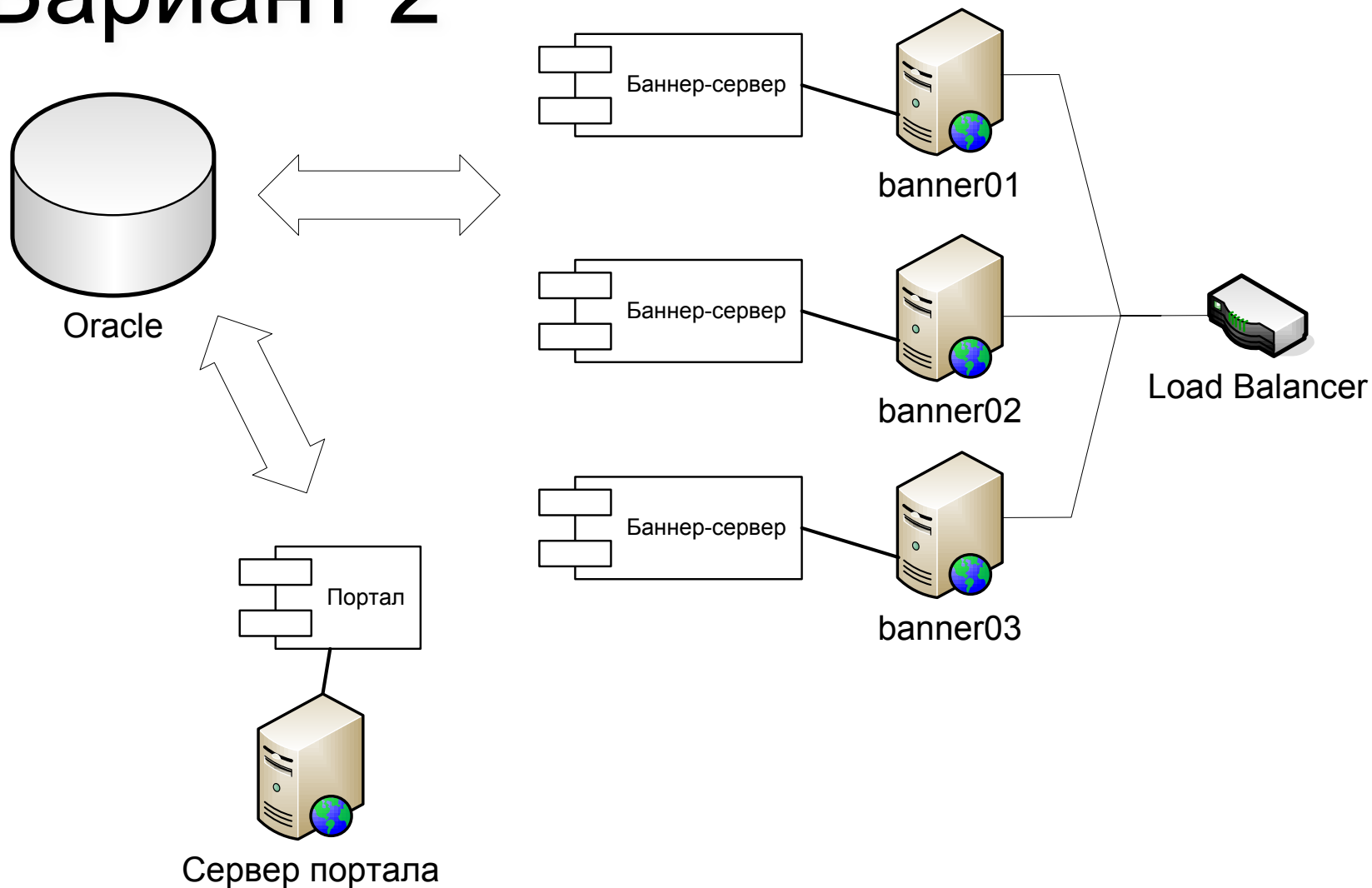


Проблемы

- Нагрузка на баннер-сервер многократно превышает нагрузку на Портал
- Что будет, если баннер-сервер сломается?

теряем деньги!

Вариант 2





Преимущества

- Балансировка загрузки
- Отказоустойчивость баннер-серверов
- Возможность независимого обновления баннер-серверов

Проблемы

- Мы балансируем *загрузку CPU!*?
- BLOBs

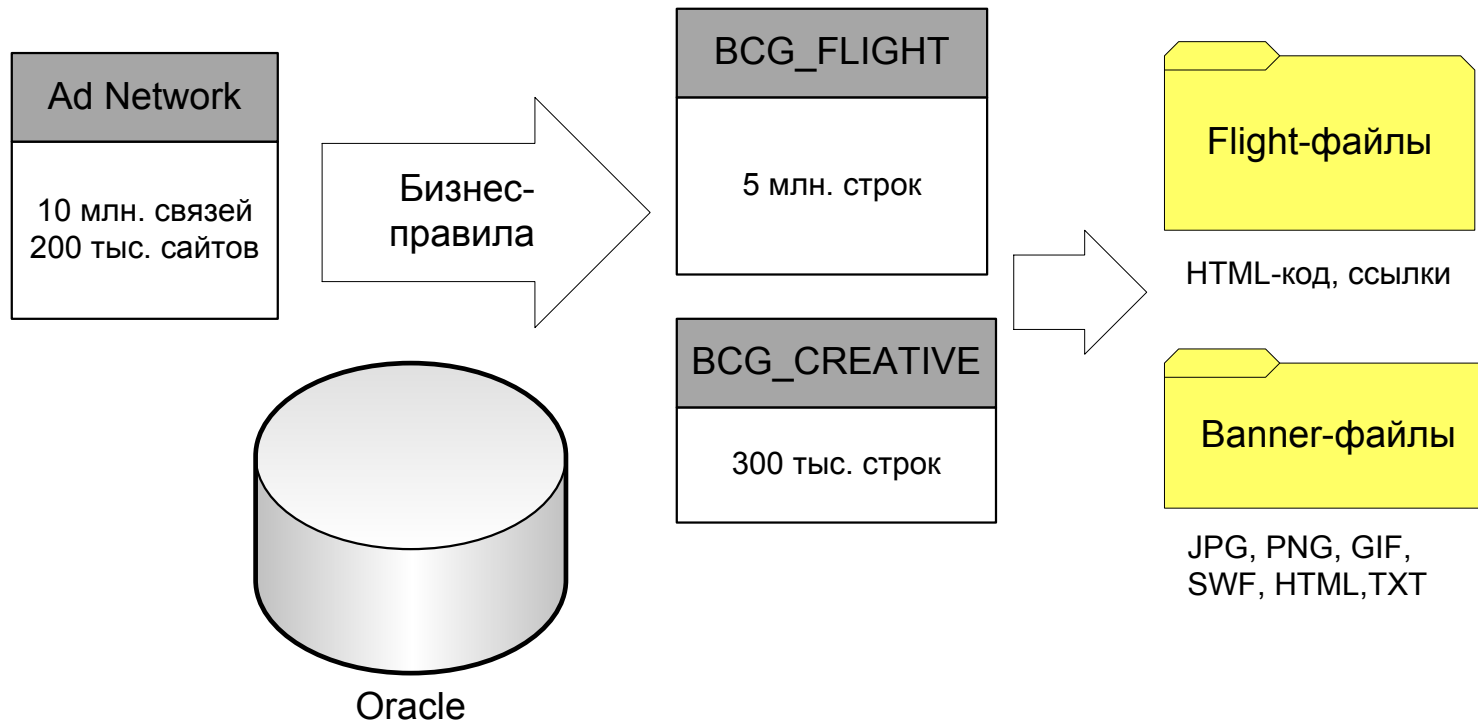


за сутки

События

- | | |
|-----------------------|----------|
| • Impression (Показ) | 450 млн. |
| • ClickThrough (Клик) | 1,5 млн. |
| • Sale (Продажа) | 35 тыс. |

Исходящие данные



- Генерируем «свернутое» представление баннеров (*BCG_**)
- Создаем файловую структуру



Игнорируем...

- Задержку данных (20 мин)
- Случайную рассогласованность данных между серверами

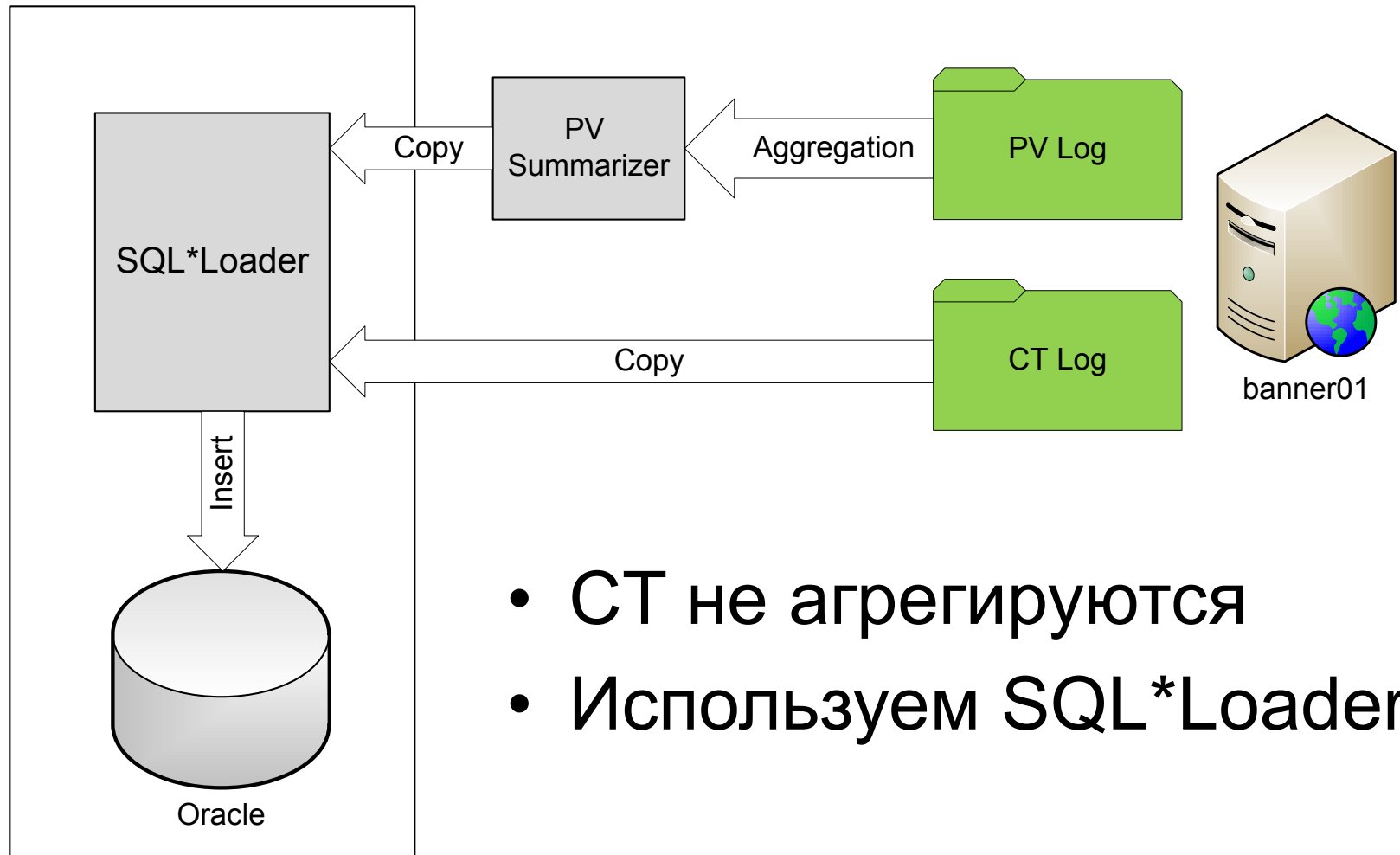


Альтернативы

- Проксирование баннеров
- Content Delivery Networks (Akamai и т.п.)



Входящие данные

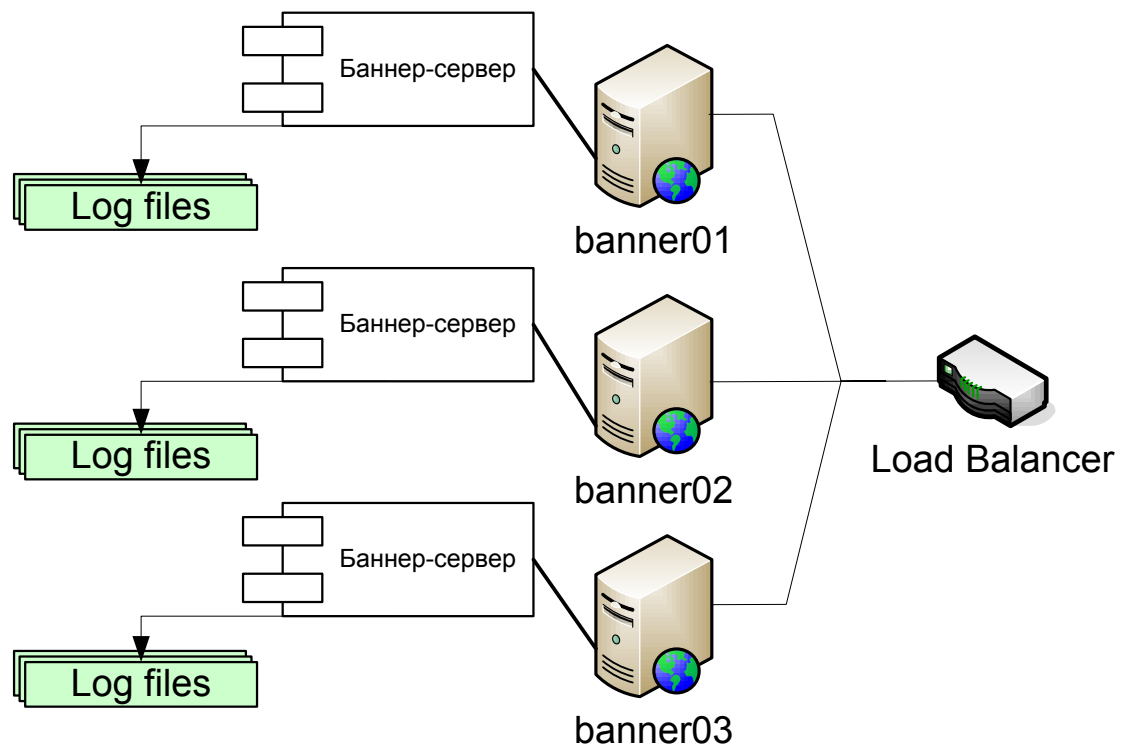
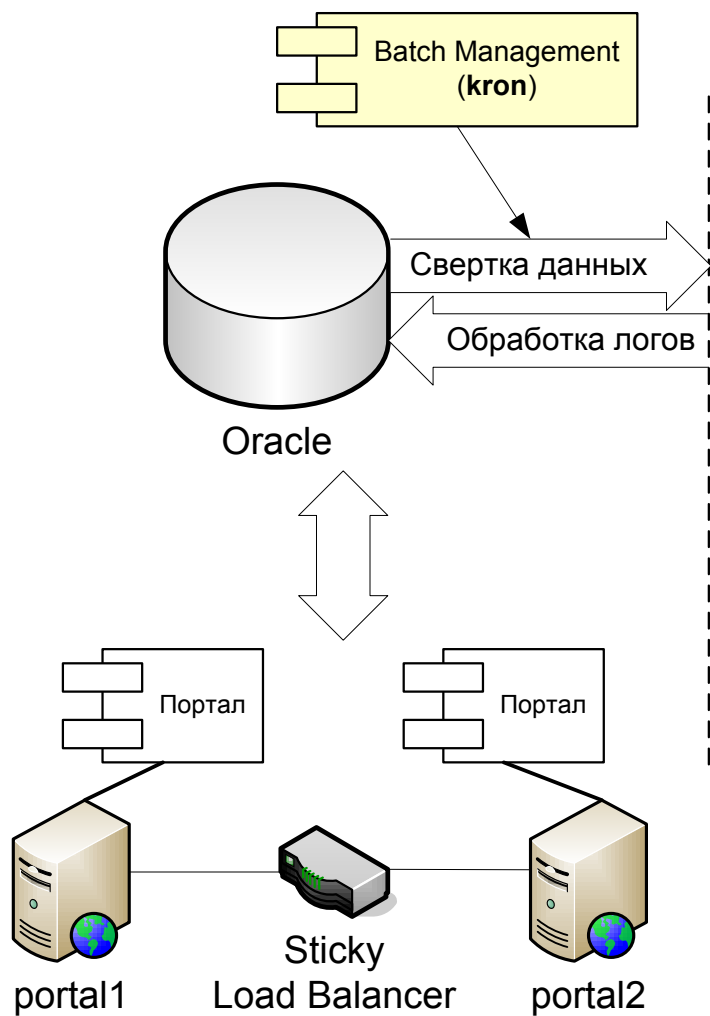




Цикл загрузки

- Сколько PV суммировать?
- Загрузку СТ задерживать нельзя!
- Late Arrivals

Вариант 3





Проблемы объема

- Размер БД растет
- Оперативные данные составляют не более 10%



Решения

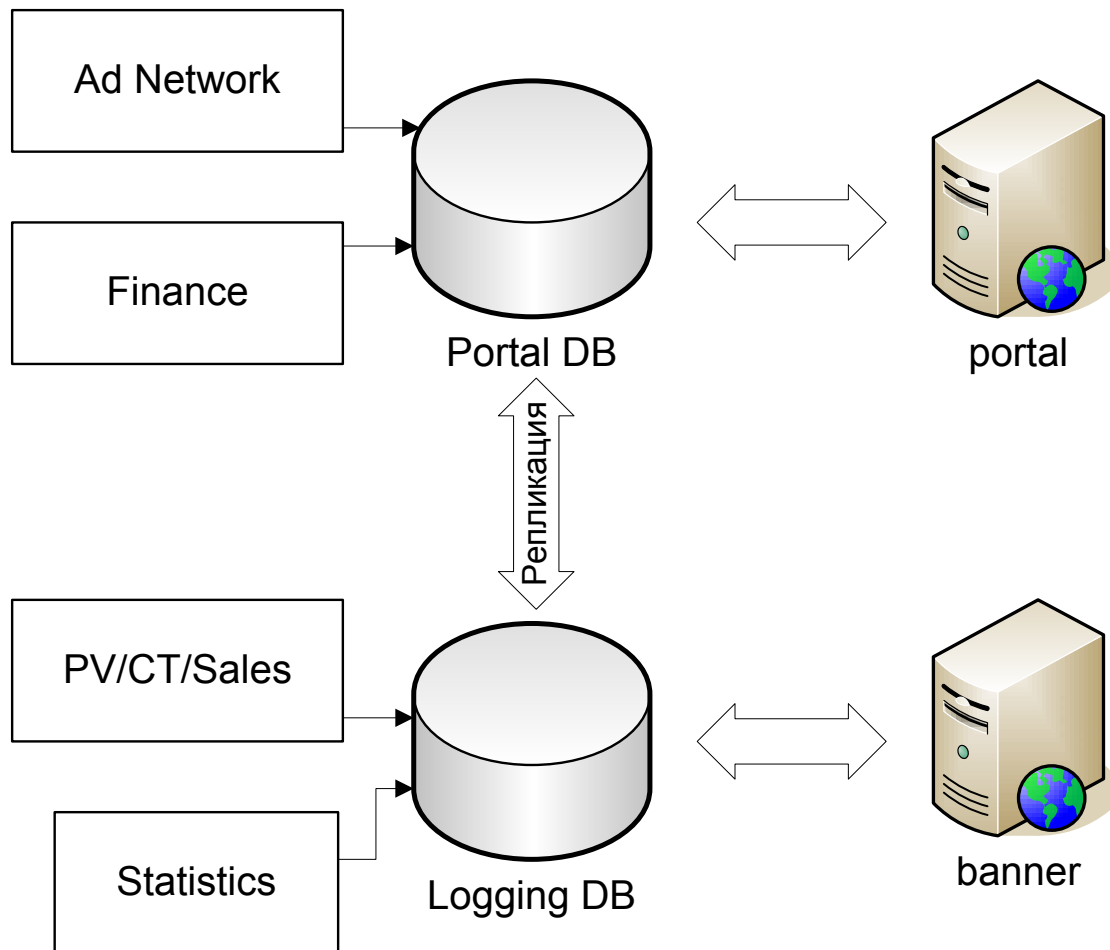
- Разделение на несколько БД
- Partitioning
- Архивирование



Одна база / много баз?

- Связанность данных
- Объем данных для репликации
- Сложность схемы

Разделение данных

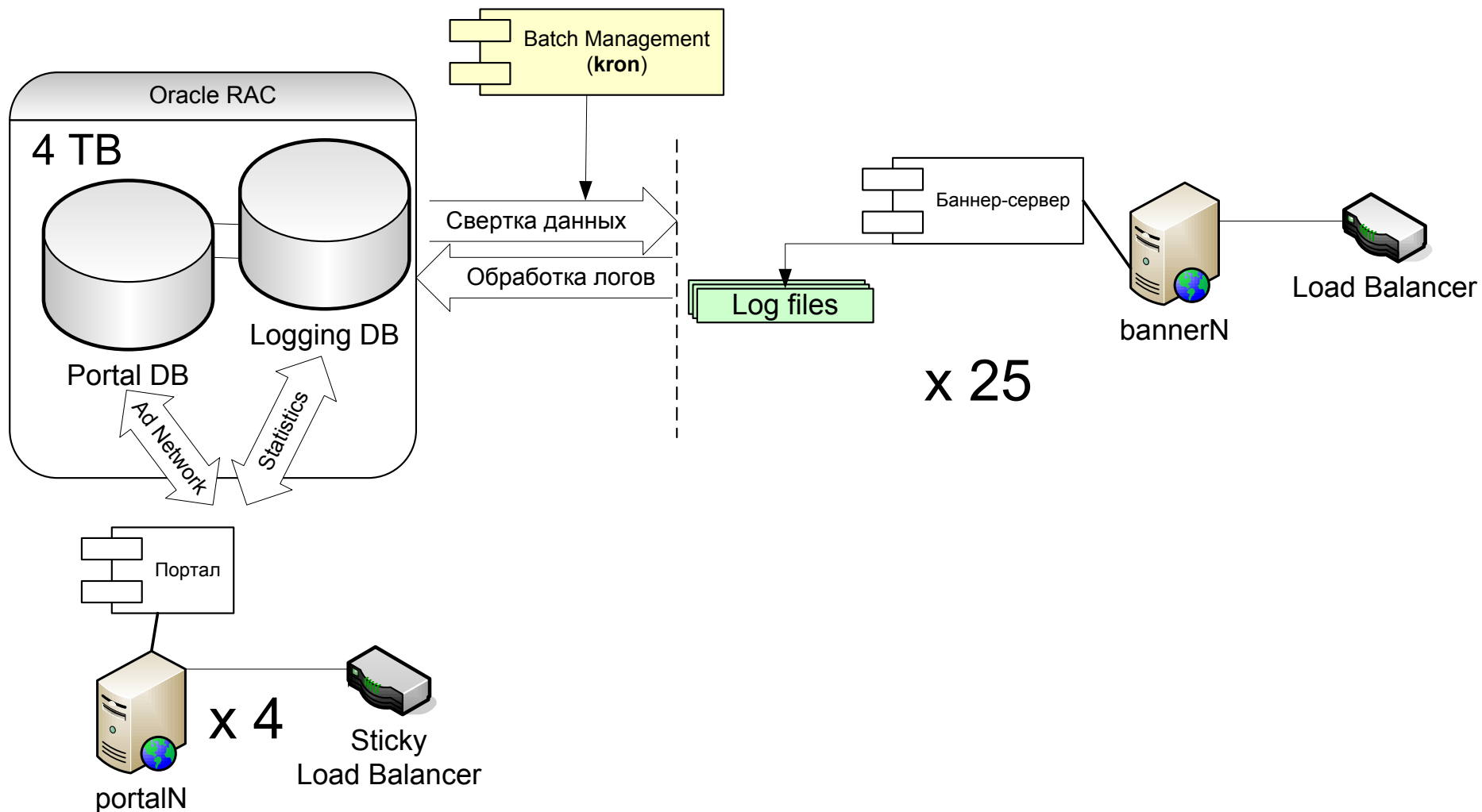




Capacity Planning

- Оценка загрузки СУБД
- Оценка загрузки баннер-серверов

Последний вариант

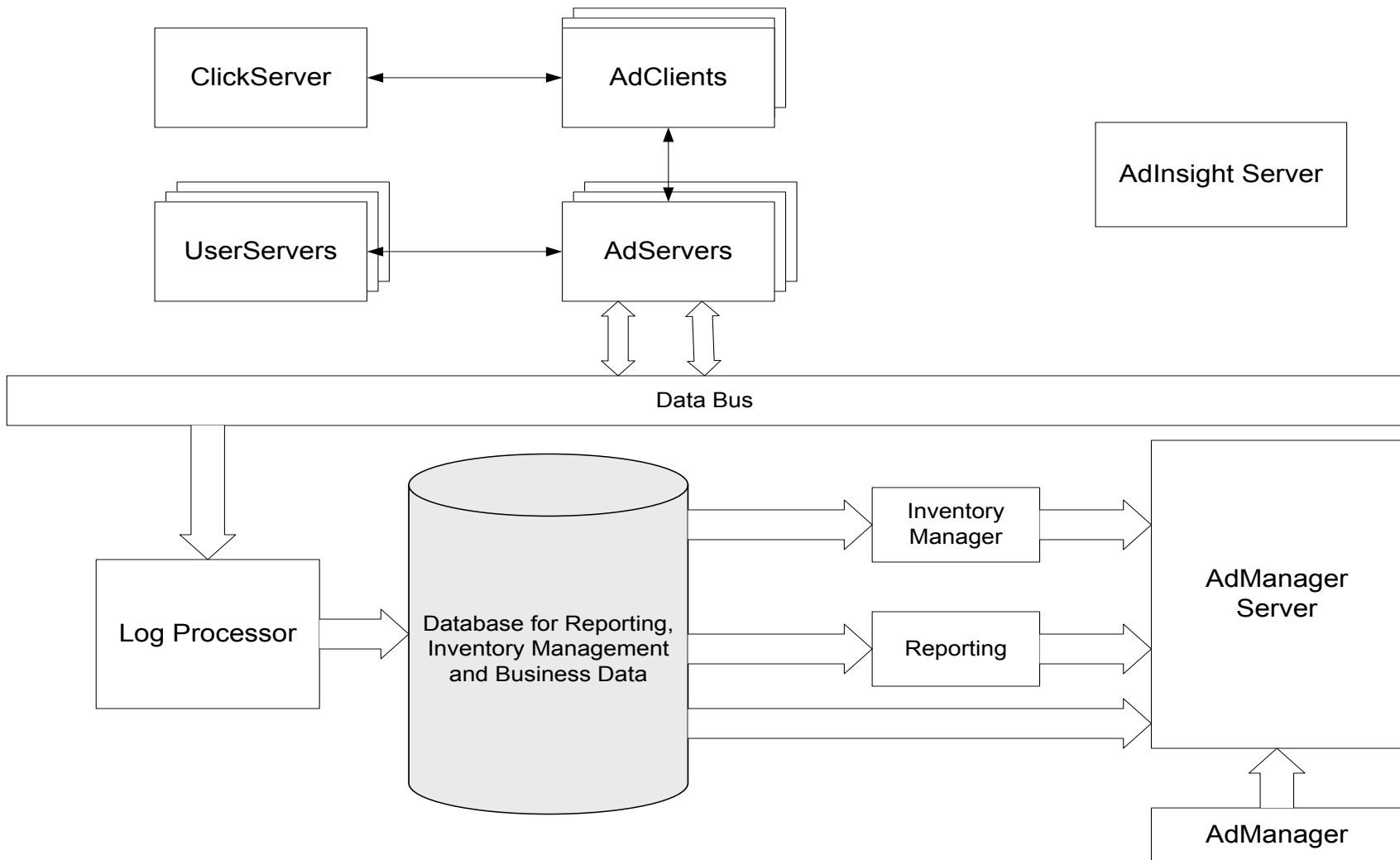




Сравним с DART Enterprise

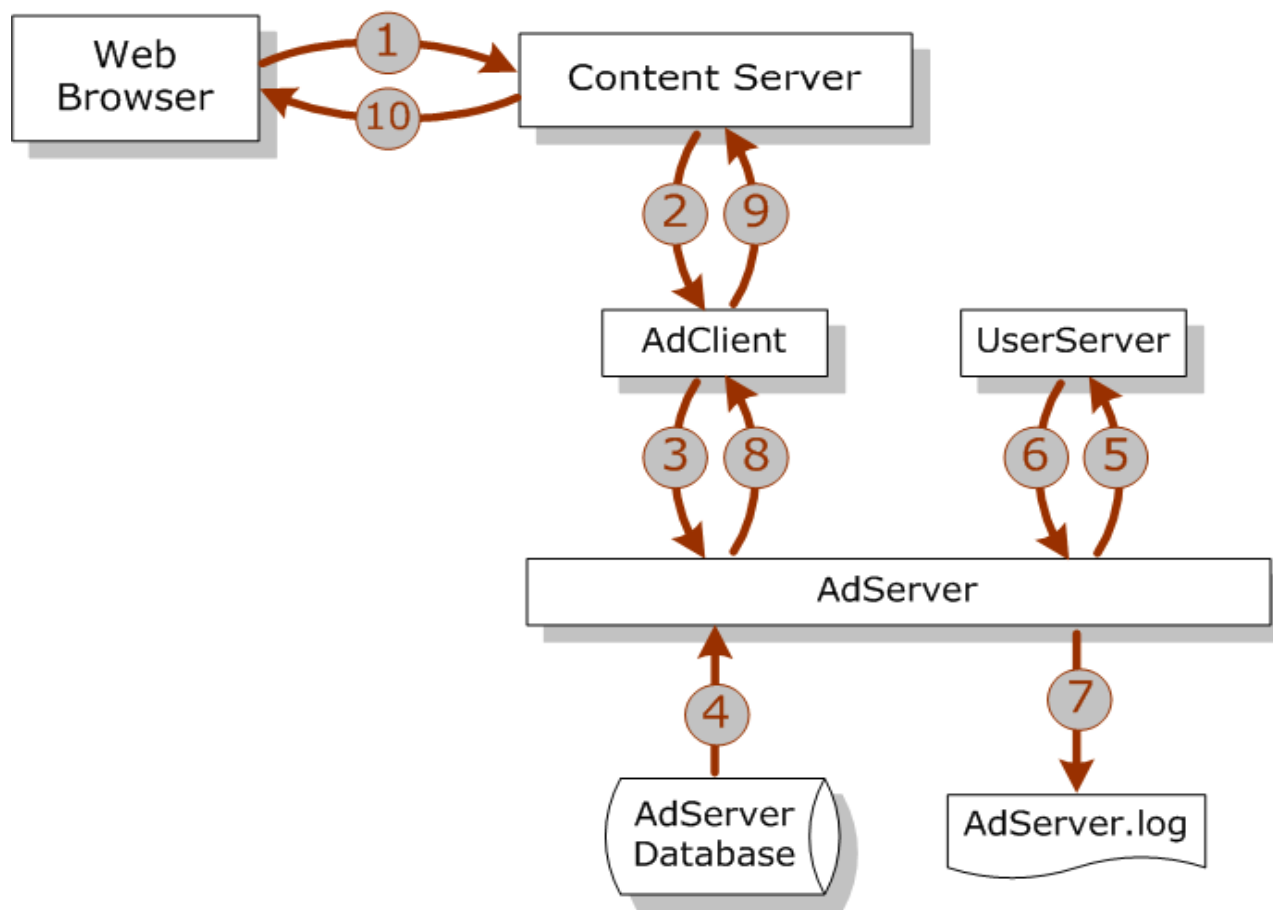
- Построен на базе Oracle
- Подсчет уникальных событий
- Нет поддержки Sales (хотя есть возможность модифицировать DART)
- Четкое разделение функций между серверами

Компоненты DART





Обработка запроса в DART





РОССИЙСКИЕ ИНТЕРНЕТ-ТЕХНОЛОГИИ
Высокие нагрузки

Вопросы?

Сергей Нековаль

Компания «Грамант»

snekoval@gramant.ru

www.gramant.ru/co/highload2008.ppt

