



ОНЛАЙН-ОБРАЗОВАНИЕ



OLTP vs OLAP

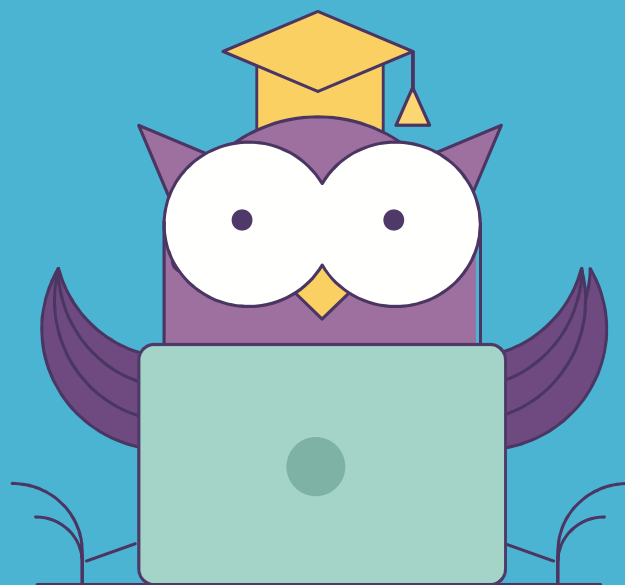
часть 1

Курс «Архитектор высоких нагрузок»

Занятие № 15



Меня хорошо слышно && видно?



Напишите в чат, если есть проблемы!

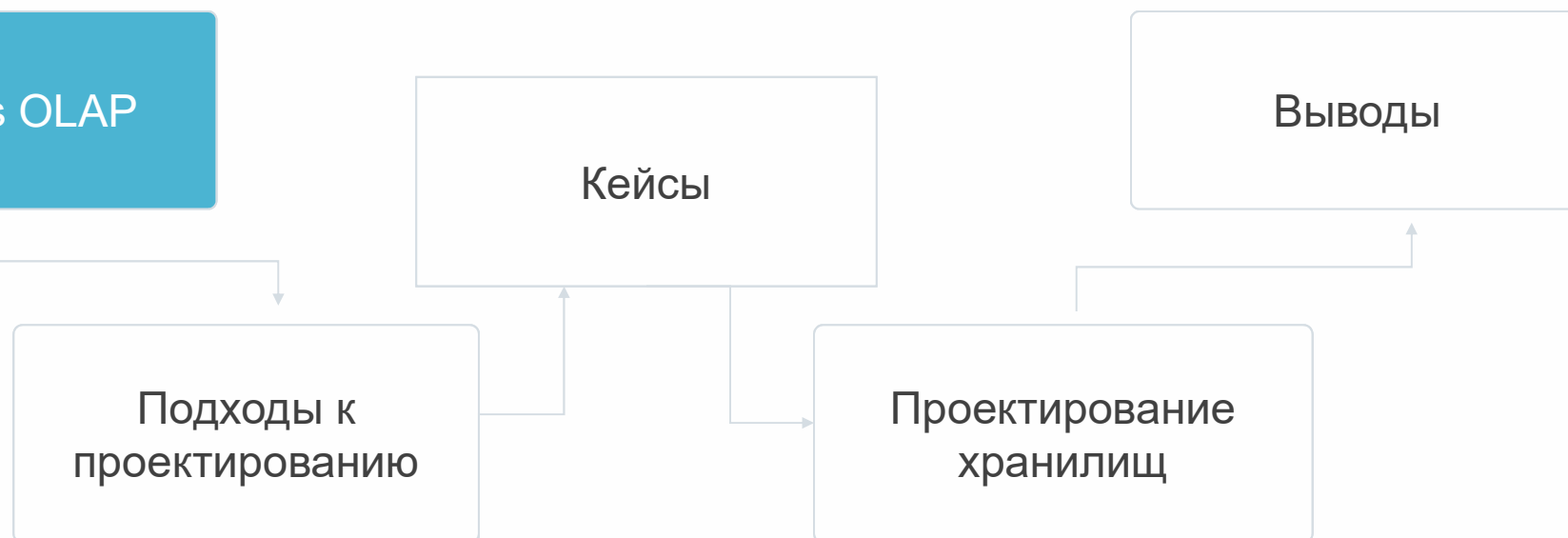
Ставьте + если все хорошо
Ставьте - если есть проблемы

По окончании вебинара вы сможете:

- Объяснить, что такое OLTP и в чем его отличия от OLAP
- Рассказать особенности проектирования для OLTP
- Как проектировать OLAP



OLTP vs OLAP



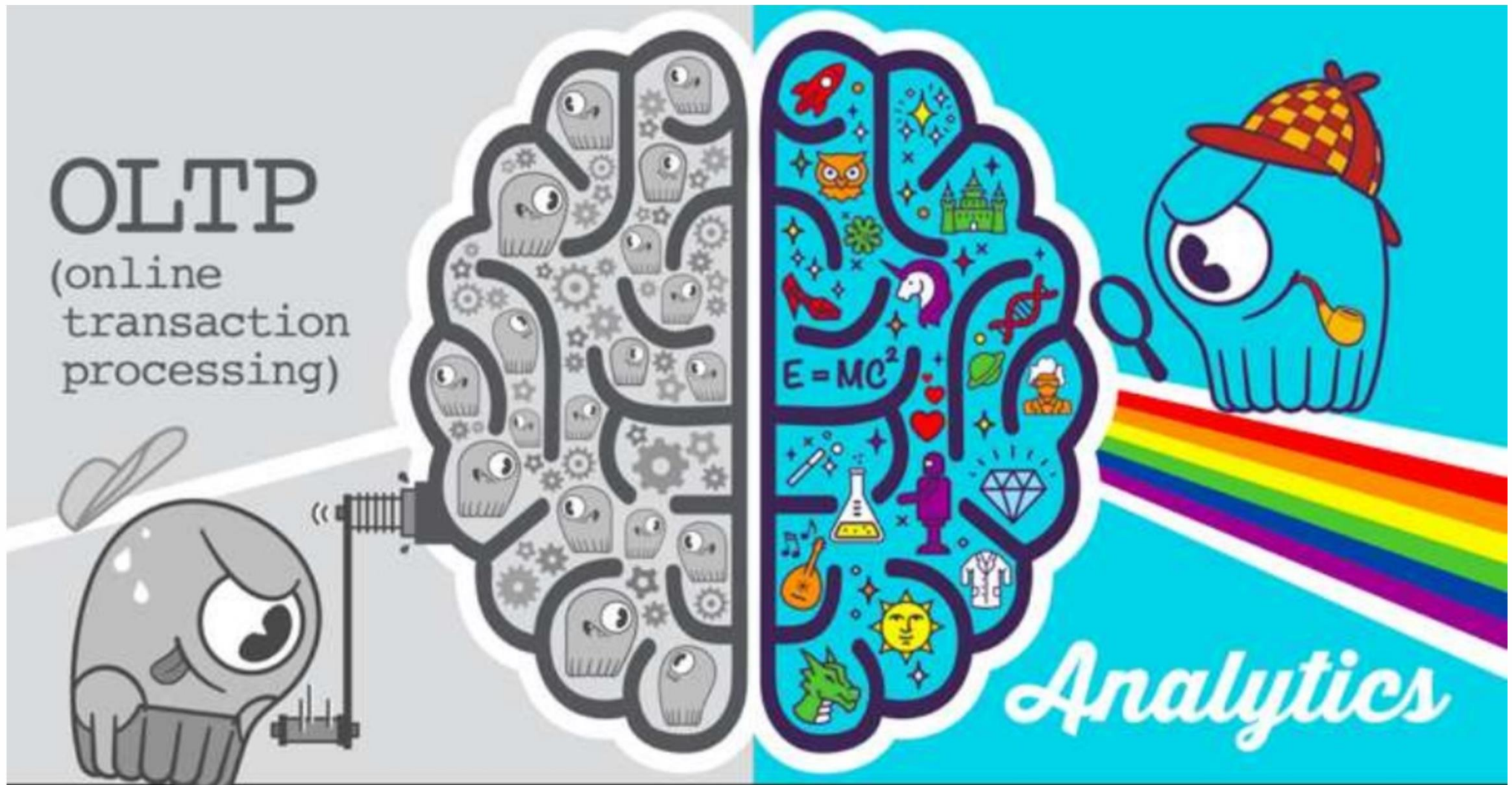
On-line Transaction Processing (OLTP)

- много маленьких коротких транзакций
- Чаще всего выполняются INSERT, UPDATE, DELETE
- Select без GROUP BY и возвращают 1 – несколько строк

On-line Analytical Processing (OLAP)

Select выполняются значительно чаще вставки\изменения

Select охватывают большие периоды – месяц, год



https://www.slideshare.net/ScyllaDB/scylla-summit-2018-olap-or-oltp-why-not-both?qid=73c1c247-ac85-413a-a19e-56bc93f545a8&v=&b=&from_search=4

- Продажа товаров пользователям
- Поиск зависимостей по товарам, которые пользователи покупают вместе
- Информация о продажах по городам за 3 года
- Построение аналитики по шаблонам платежей по группам пользователей
- Прием платежей за сотовую связь

- ✓ Продажа товаров пользователям
- ✗ Поиск зависимостей по товарам, которые пользователи покупают вместе
- ✗ Информация о продажах по городам за 3 года
- ✗ Построение аналитики по шаблонам платежей по группам пользователей
- ✓ Прием платежей за сотовую связь

Есть таблица с данными о состоянии удаленного оборудования, которые приходят раз в 15 минут от каждого терминала.

Страница, которая показывает последнюю пришедшую информацию по всем терминалам. Как вы реализуете такой функционал?

- Разный характер нагрузки
- Стратегии индексирования
- Работа с разными объемами данных
- Поддержка исторических данных на OLTP системах требует дополнительных ресурсов
- Совмещение неэффективно ни для OLTP ни для OLAP

- OLTP работает с операционными данными
- Высокая степень нормализации таблиц
- Более простые запросы
- Объем OLTP базы обычно небольшой – до нескольких Гб
- Почему в OLTP может быть небольшой объем?

1. Разделите OLAP и OLTP нагрузку
2. Не храните данные, которые уже не нужны на OLTP
3. Нормализуйте таблицы
4. Удалите ненужные индексы
5. Подумайте над разделением таблиц (связь 1 к 1), чтобы «сузить их»

- OLTP - In Memory
Redis
Tarantool



redis



- OLAP – Column stored
Vertica
ClickHouse
GreenPlum
Amazon Redshift

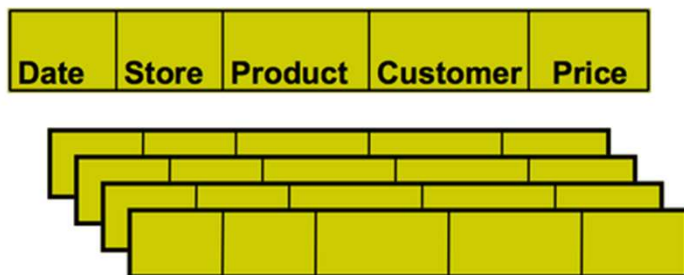


ClickHouse

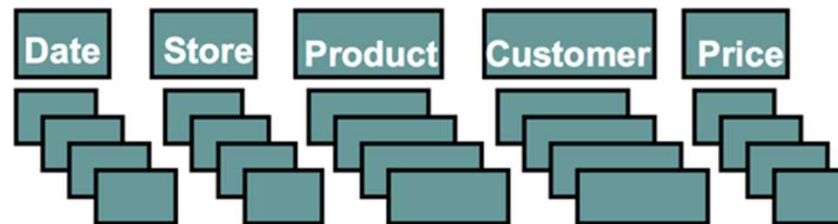


Разница подхода построчного или колоночного хранения

row-store



column-store



ID	Day	Discount
10	4/4/2013	0.25
11	20/4/2013	0.065
12	12/6/2013	0.275
13	25/7/2013	0.0

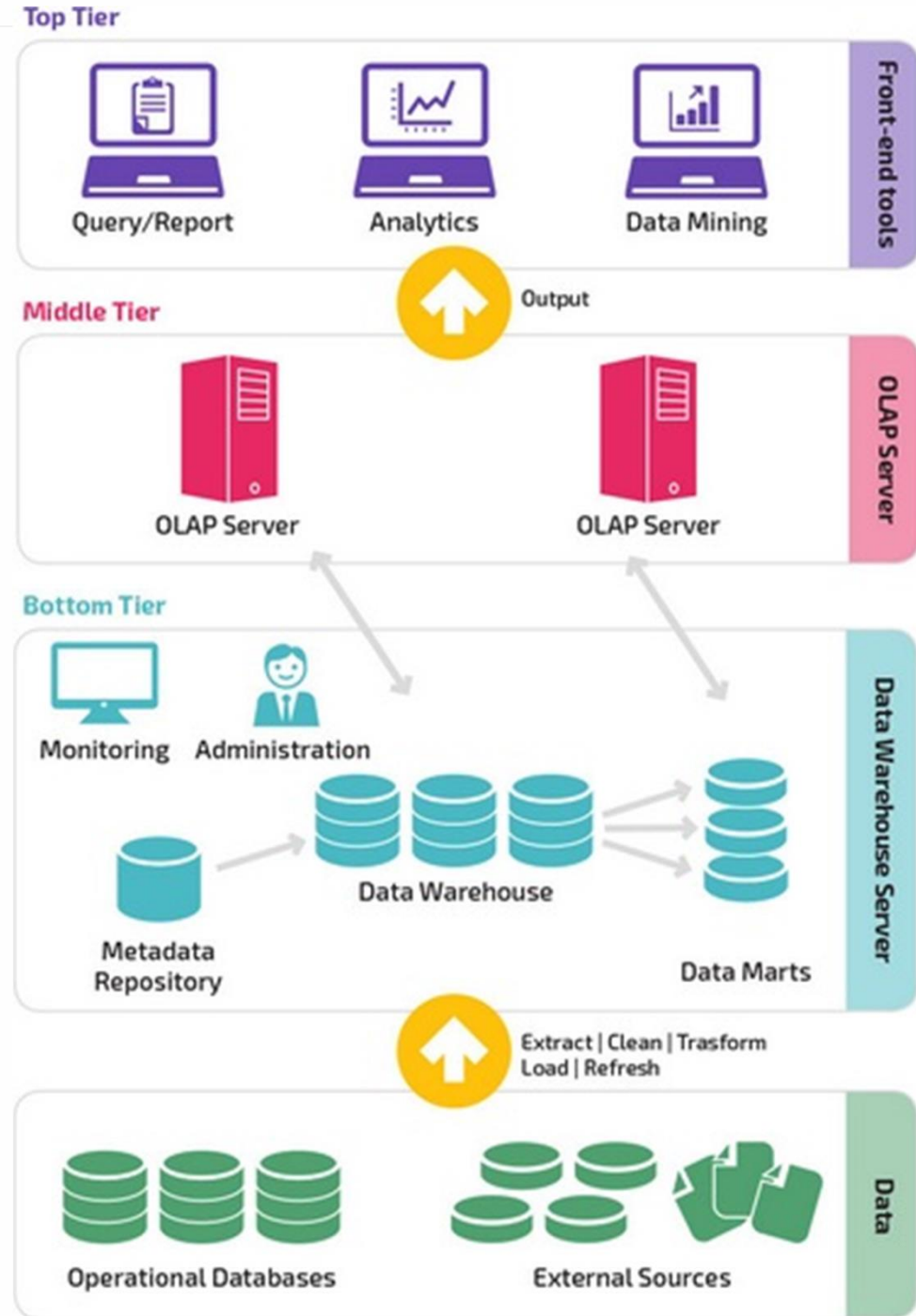
OID	ID
100	10
101	11
102	12
103	13

OID	Day
100	4/4/2013
101	20/4/2013
102	12/6/2013
103	25/7/2013

OID	Discount
100	0.25
101	0.065
102	0.275
103	0.0

- Звезда
- Снежинка
- Data Vault

- Lambda
- Кappa



Bill Inmon
Ralf Kimball

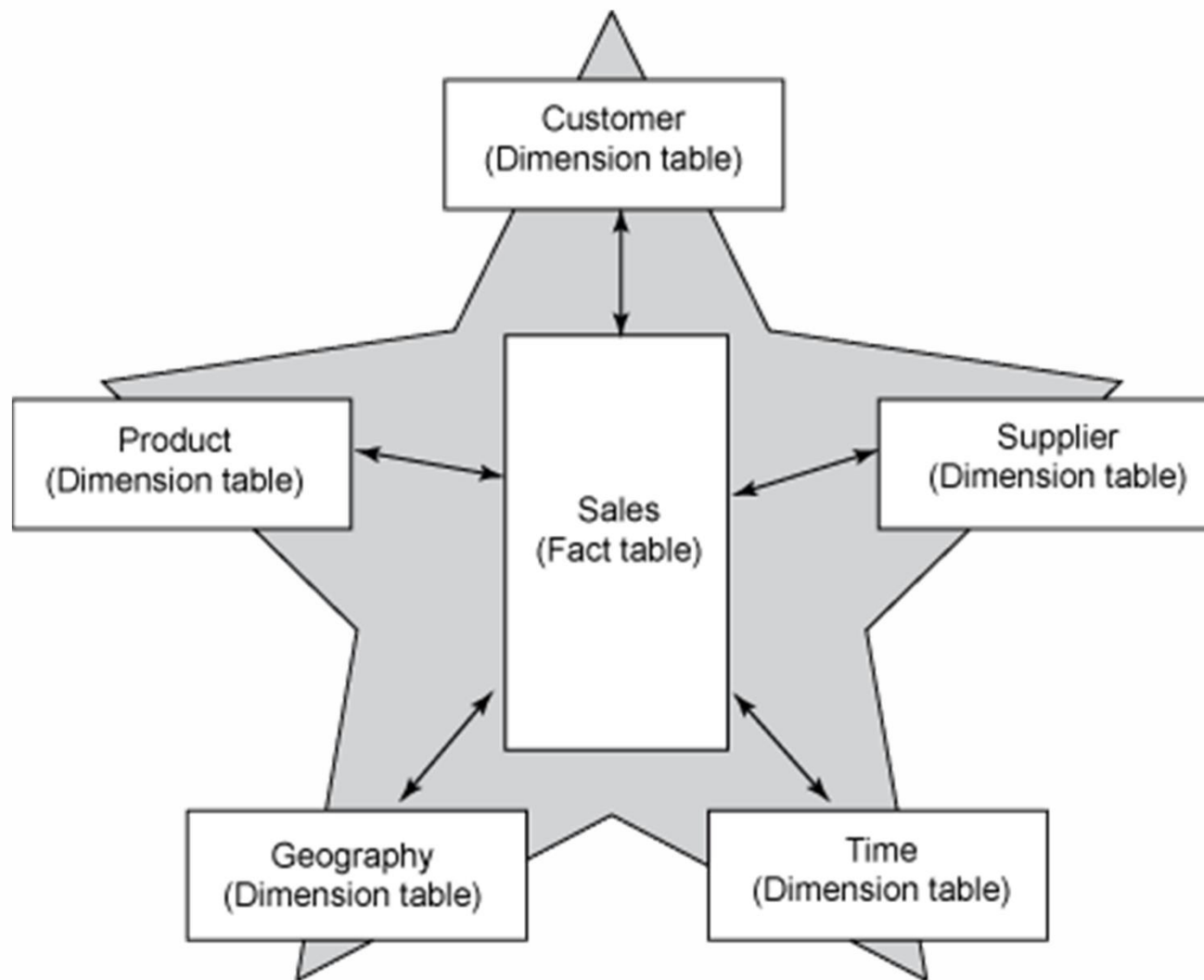
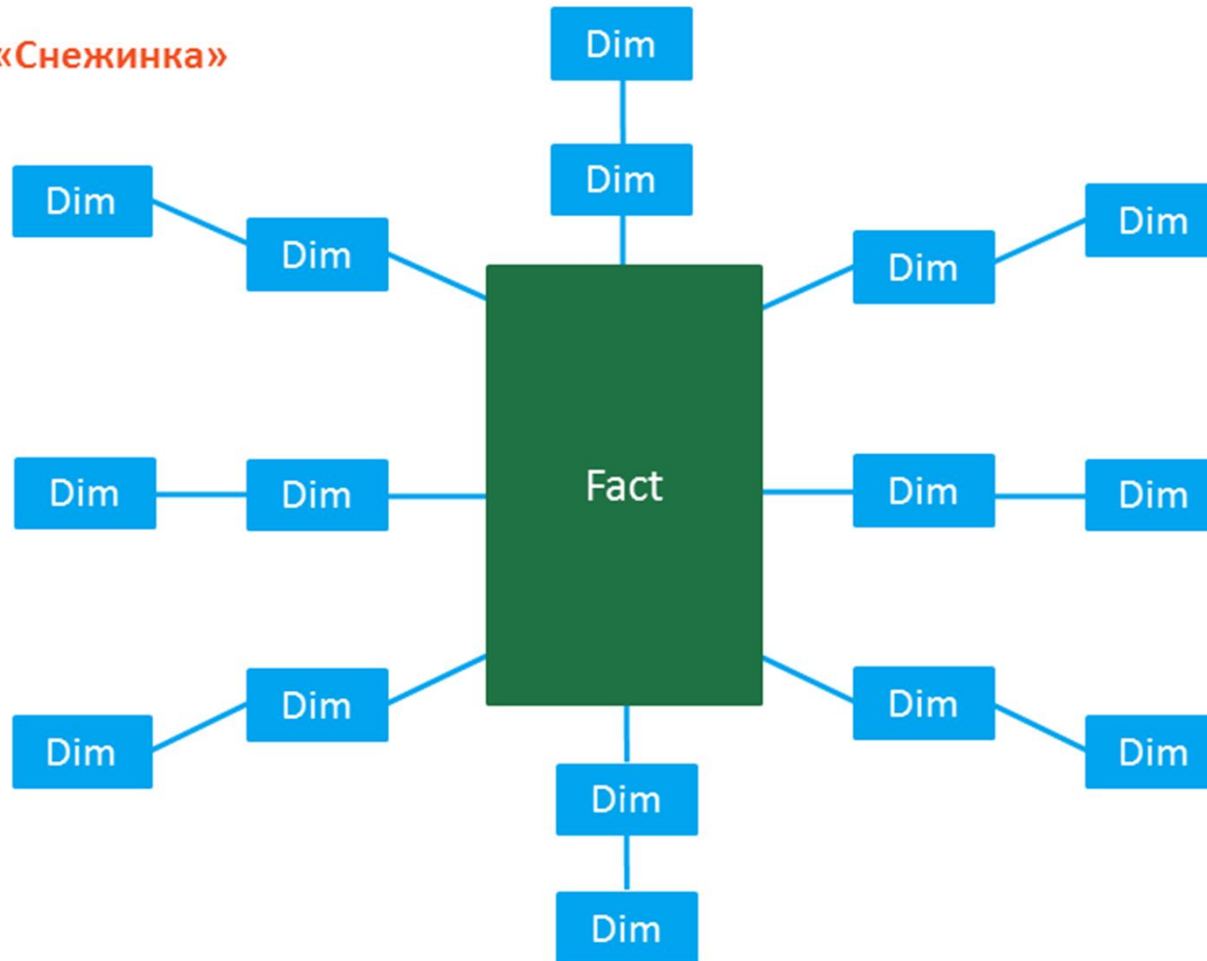
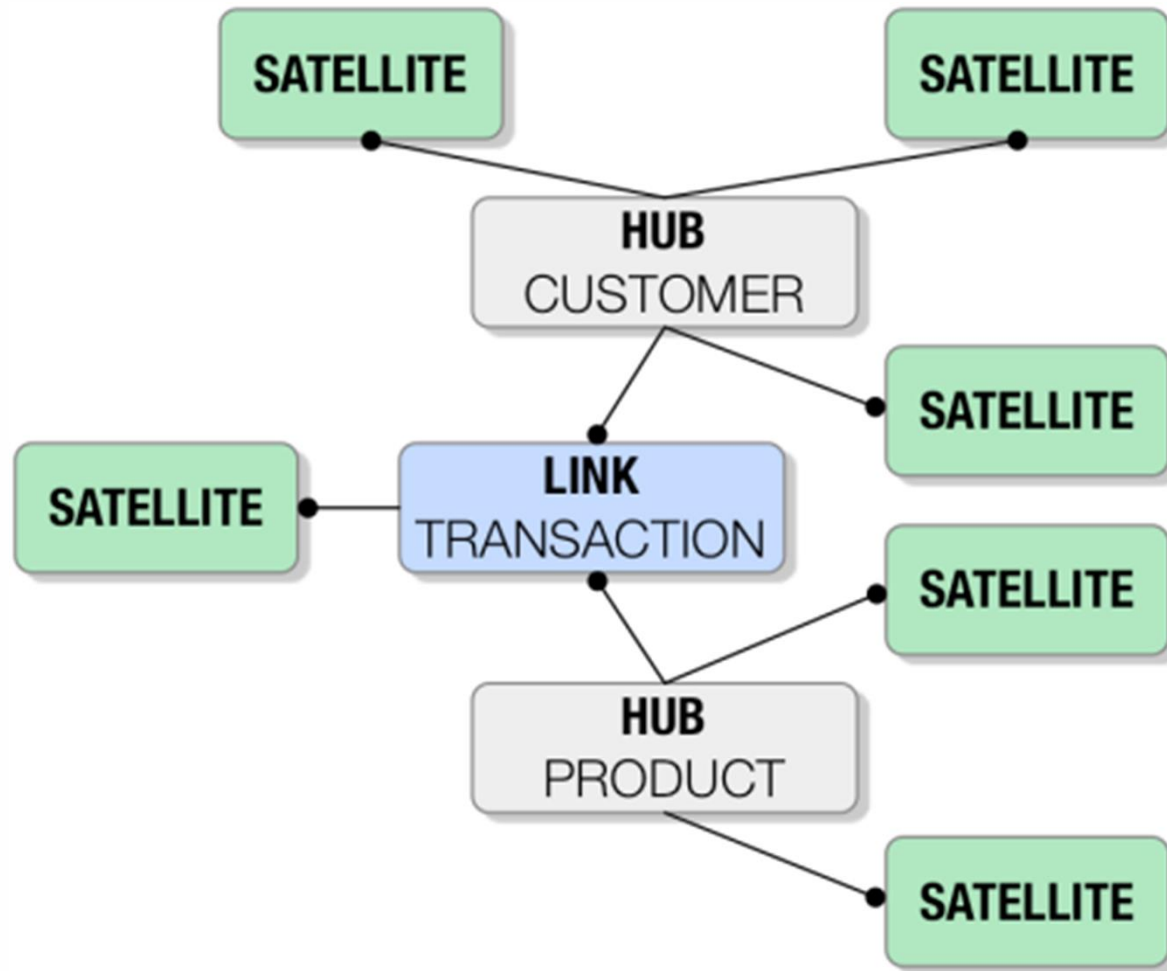


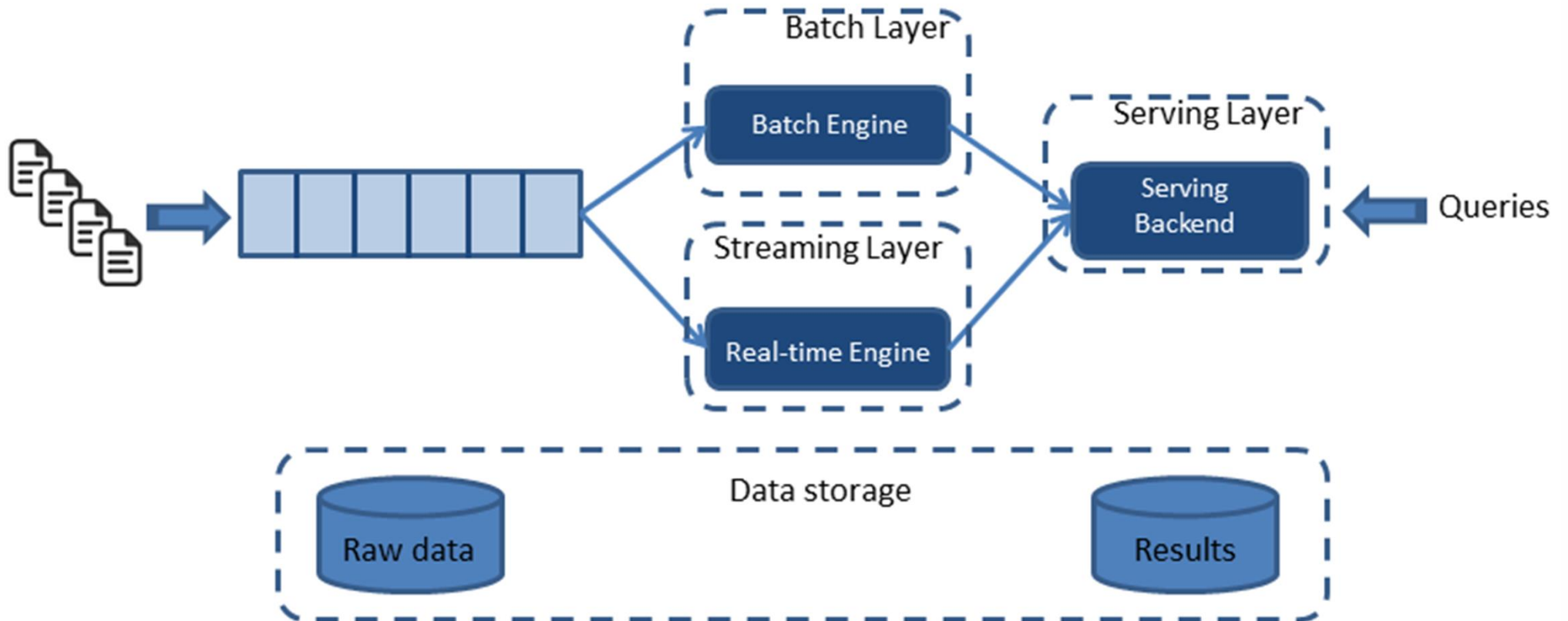
Схема: «Снежинка»



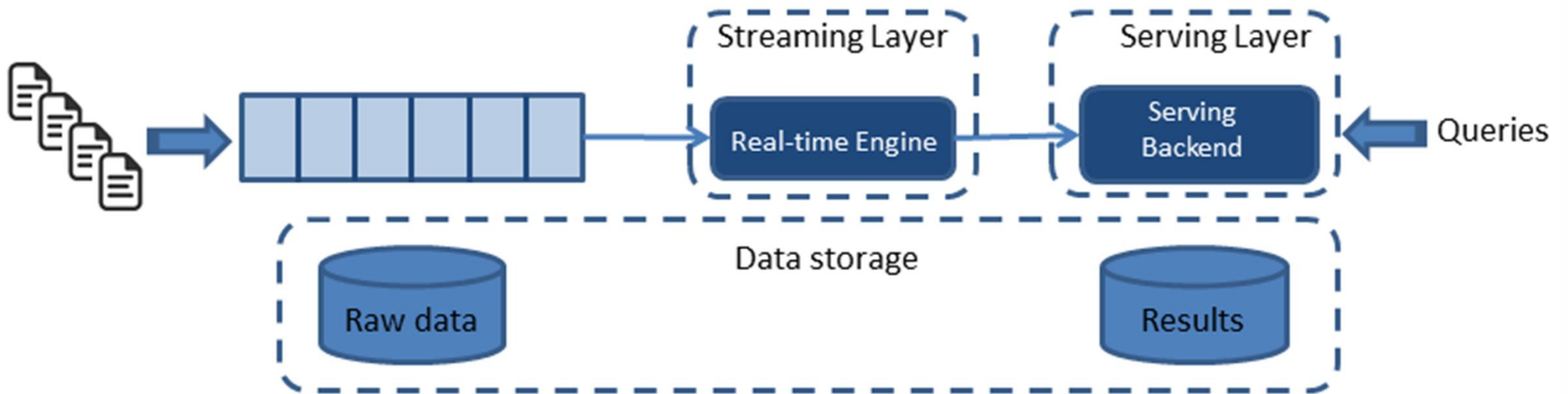
Dan Linstedt, 2001



Nathan Marz, 2013



2014



	OLTP	OLAP
Источник данных	И есть источник данных	Различные OLTP Базы Внешние источники
Цель БД	Фиксация происходящих бизнес процессов	Планирование, аналитика
Проектирование БД	Нормализованные таблицы, данные в 1 экземпляре либо стремятся к этому	Различные представления, агрегированные данные, денормализация
Insertы и Updates	Операции затрагивают обычно 1 строку, много Inserts, update меньше – инициатива от пользователя	Вставка больших блоков данных, изменения обычно затрагивают большие объемы данных
Запросы	Небольшие довольно несложные запросы, обычно возвращают 1 или несколько строк	Сложные запросы, затрагивающие данные за большой период, агрегации данных
Скорость обработки	Очень быстро	Зависит от количества данных
Требования к дисковому пространству	Относительно небольшие объемы данных	Большие так как хранятся агрегации, исторические данные
Резервное копирование и восстановление	Регулярные бэкапы и процедуры направленные на восстановление в случае сбоя	Обычно данные можно загрузить заново при условии хранения исторических данных

Рефлексия

О чем мы говорили сегодня?

- В чем особенности OLTP?
- Зачем разделять OLTP и OLAP нагрузку?



Рефлексия

Напишите, пожалуйста, свое впечатление о вебинаре.

- Отметьте 3 пункта, которые вам запомнились с вебинара.
- Что вы будете применять в работе из сегодняшнего вебинара?



Заполните, пожалуйста,
опрос в ЛК о занятии



Спасибо
за внимание!

До встречи в **Slack** и на вебинаре

