

Отчёт по заданию 8

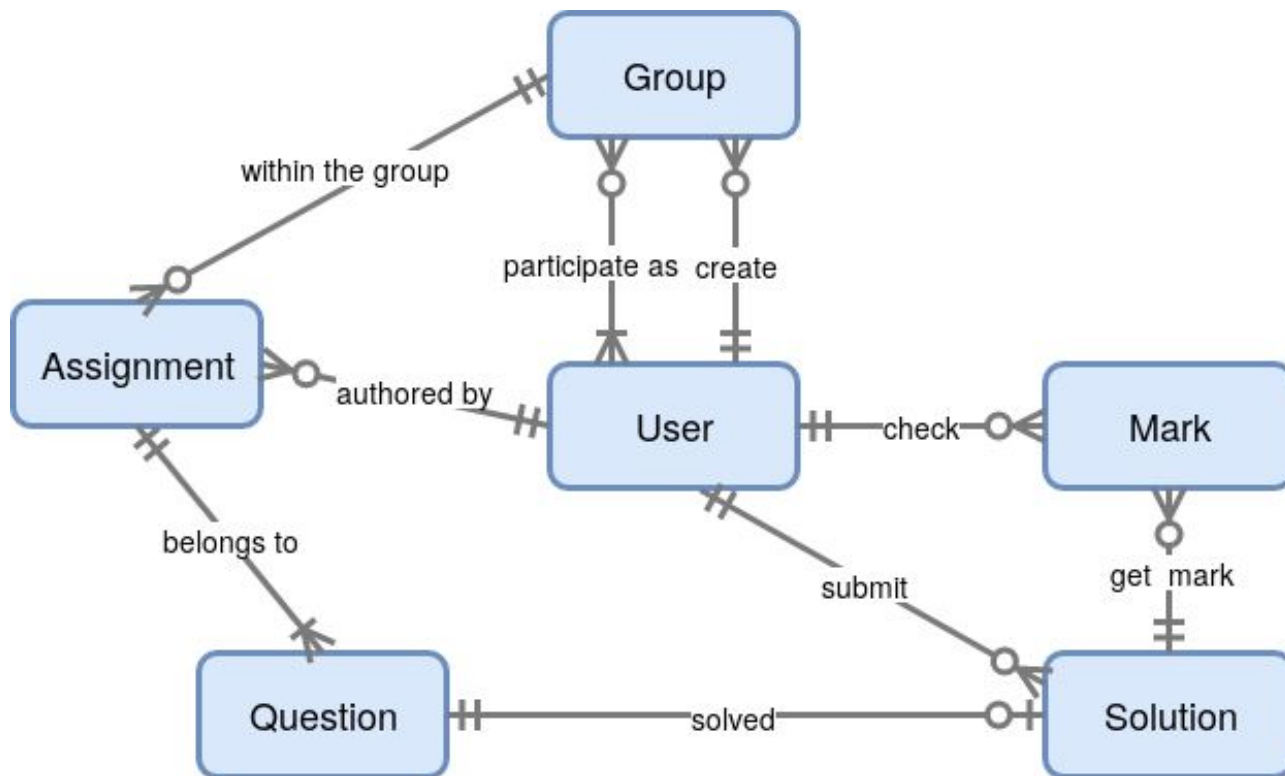
Группа I



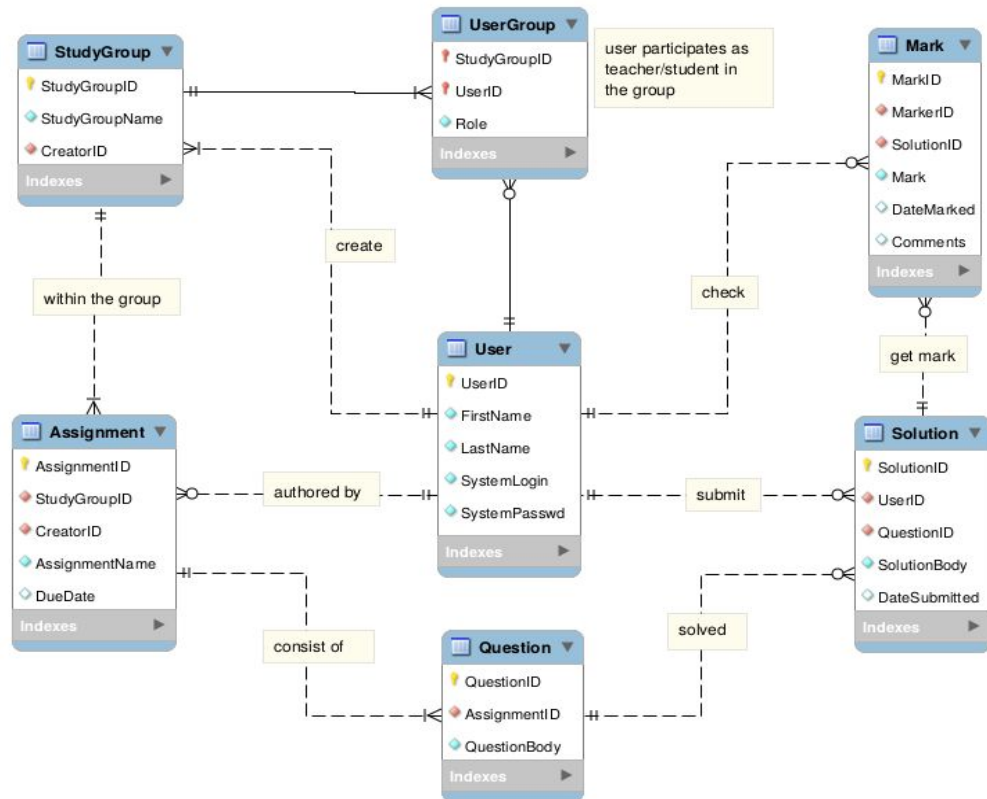
Состав группы

- Голдобина Ольга (Подготовка запросов)
- Демин Георгий (Подготовка запросов)
- Лебедь Федор (Загрузка данных)
- Петренко Дарья (Создание схемы)
- Ташевцев Артём (Подготовка запросов)

Концептуальная модель



Логическая модель - BCNF



Время загрузки на сервер

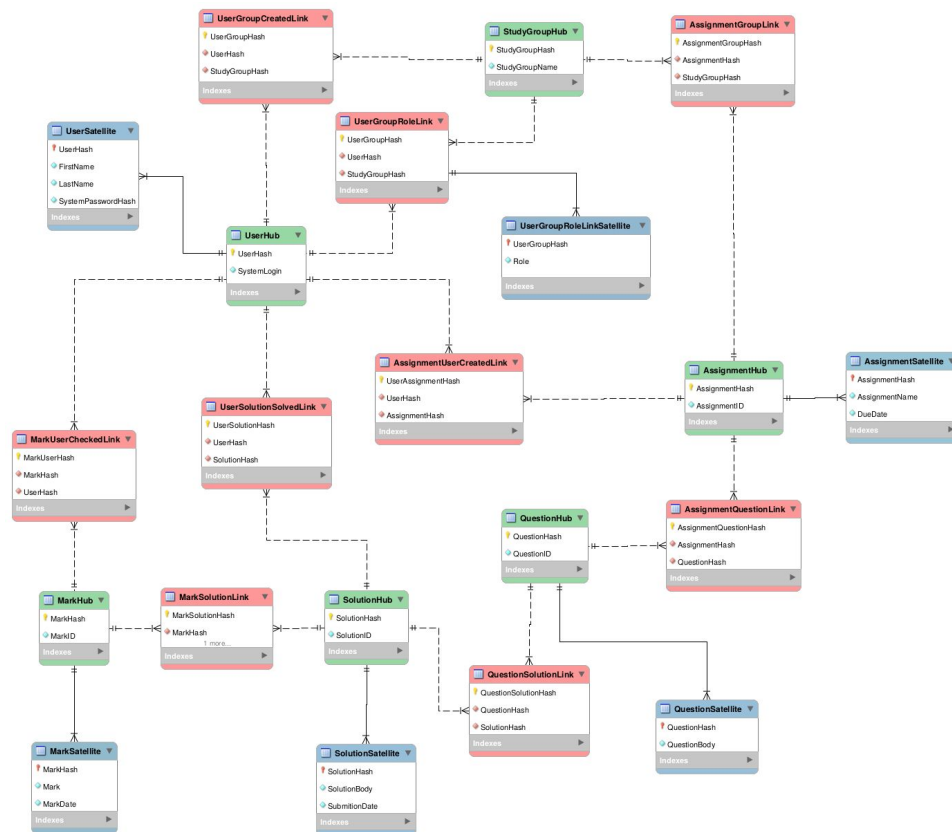
Название таблицы	Время загрузки
users	1m 24s
groups	10s
user-group	2m 26s
assignment	24s
question	57s
solution	2m 27s
mark	1m 30s
Total	7m 38s

Протокол запуска запросов от бизнеса

Формулировка запроса	число возвращаемых строк	Время выполнения
1. Топ 5 студентов за последние 30 дней по средней оценка за любой один учебный курс (т.е. лидеры могут оказаться с разных учебных курсов)	5	0.084с (8 runs)
2. Топ 5 преподавателей за прошлый календарный месяц по количеству проверенных решений по всем учебным курсам	5	0.0151с (8 runs)
3. Доля сданных решений из всех возможных по заданиям с уже прошедшими сроками по каждому учебному курсу	301	0.021с (8 runs)
4. Структура студентов по учебным курсам (т.е. учебные курсы с числом студентов, отсортированные по убыванию)	12	0.0189с (8 runs)

Логическая модель - Data Vault

■ - Link
■ - Hub
■ - Satellite



Во всех хабах, ссылках и сателлитах также содержатся мета-данные:

load timestamp - время загрузки записи

record source - откуда загрузили

Протокол запуска запросов от бизнеса

Формулировка запроса	Число возвращаемых строк	Время выполнения
1. Топ 5 студентов за последние 30 дней по средней оценка за любой один учебный курс (т.е. лидеры могут оказаться с разных учебных курсов)	5	0.040 s
2. Топ 5 преподавателей за прошлый календарный месяц по количеству проверенных решений по всем учебным курсам	5	0.026
3. Доля сданных решений из всех возможных по заданиям с уже прошедшими сроками по каждому учебному курсу	301	0.0090
4. Структура студентов по учебным курсам (т.е. учебные курсы с числом студентов, отсортированные по убыванию)	12	0.030

Выводы

- 1) Модели Data Vault и BCNF для данной модели и выборки имеют сравнимые скорости работы
- 2) Модель BCNF практичнее с точки зрения составления запросов
- 3) Загрузка данных — узкое место системы

Заключение

В рамках 8 задания было выполнено:

1. Реализованы модели BCNF и Data Vault
2. Сгенерированы и загружены данные
3. Составлены запросы
4. Произведены временные замеры