

Postgresql: нормализация, индексы, транзакции

онлайн-воркшоп с Кириллом Храпковым

День 1:
Нормализация

МОЙ ОПЫТ

04.2011 - 01.2018 - web-разработчик компании eXteam
01.2018 - 04.2020 - ведущий разработчик компании Cifronomika
05.2020 - 06.2022 - ведущий разработчик компании JetRockets
11.2022 - н.в. - ведущий разработчик компании Trucker

Навыки:

бэкенд: Ruby on Rails, PHP

фронтэнд: javascript, jQuery, coffeescript, html, css2/css3, haml, sass (кроссбраузерная, адаптивная, резиновая верстка)

БД: MySQL, Postgresql, sqlight, elasticsearch

DSS: redis, rabbitmq



Кирилл Храпков

Цель воркшопа

Нас с вами ждет 3 встречи. И цели у нас такие:

1. Нормализация
2. Индексы
3. Транзакции

Как будем двигаться по теме?

- теория + практика

Когда и куда задавать вопросы?

- в процессе рассказа -

можно поднимать руку в реакциях и подключаться голосом

День 1:

- познакомимся с термином “нормализация” и “денормализация”
- разберемся с тем, что это такое, для чего нужна и как правильно ее “готовить”
- поделимся личным опытом

Нормализация - это метод проектирования базы данных с целью уменьшения избыточности данных и улучшения их целостности путем организации в таблицы и установления связей между ними

дата	товар	цена
01.09.2023	молоко	100р
01.09.2023	хлеб	30р
01.09.2023	цветы	1500р
02.09.2023	такси	250р

дата	товар	цена	покупатель
01.09.2023	молоко	100р	Храпкова Татьяна
01.09.2023	хлеб	30р	Храпкова Татьяна
01.09.2023	цветы	1500р	Храпкова Татьяна
01.09.2023	цветы	2000р	Иванова Мария
01.09.2023	цветы	2000р	Петрова Анна
02.09.2023	такси	250р	Храпкова Татьяна

Нормализация

товары	
id	Наименование
1	молоко
2	хлеб
3	цветы
4	такси

покупатели	
id	ФИО
1	Храпкова Татьяна
2	Иванова Мария
3	Петрова Анна

валюта	
id	наименование
1	Рубль
2	Евро
3	Доллар
4	Юань

траты				
дата	стоимость	id валюты	id товара	id покупателя
01.09.2023	100	1	1	1
01.09.2023	30	1	2	1
01.09.2023	1500	1	3	1
01.09.2023	2000	1	3	2
01.09.2023	2000	1	3	3
02.09.2023	250	1	4	1

Зачем нормализовать базу данных?

1. уменьшить объем и вес базы данных
2. повысить удобство управления данными
3. упростить поиск и повысить производительность
4. устранить аномалии

Нормальная форма - требование, предъявляемое к структуре таблиц в теории реляционных баз данных для устранения из базы избыточных функциональных зависимостей между атрибутами

Нормальные формы:

1. Ненормализованная форма или нулевая нормальная форма (UNF)
2. Первая нормальная форма (1NF)
3. Вторая нормальная форма (2NF)
4. Третья нормальная форма (3NF)
5. Нормальная форма Бойса-Кодда (BCNF)
6. Четвертая нормальная форма (4NF)
7. Пятая нормальная форма (5NF)
8. Доменно-ключевая нормальная форма (DKNF)
9. Шестая нормальная форма (6NF)

Ненормализованная форма или нулевая нормальная форма (UNF)

Отношение находится в 0НФ, если порядок его строк и столбцов не имеет значения

Пример нарушения 0НФ:

Порядковый номер	А	Б
1	Храпкова	Татьяна
2	Иванова	Мария
3	Петрова	Анна

Давайте приведем эту таблицу к 0НФ:

surname	name
Храпкова	Татьяна
Иванова	Мария
Петрова	Анна

Первая нормальная форма

Отношение находится в 1НФ, если:

- в таблице нет дублирующих строк
- в каждой ячейке таблицы хранится атомарное значение (одно не составное значение)
- в столбце хранятся данные одного типа
- отсутствуют массивы и списки в любом виде

Пример нарушения 1НФ:

ФИО матери	дети
Храпкова Татьяна	сыновья: Артем, Семен
Иванова Мария	дочь: Ксения
Петрова Анна	сын Сергей, дочь Вероника

Давайте приведем эту таблицу к 1НФ:

ФИО матери	имя ребенка	пол ребенка
Храпкова Татьяна	Артем	м
Храпкова Татьяна	Семен	м
Иванова Мария	Ксения	ж
Петрова Анна	Сергей	м
Петрова Анна	Вероника	ж

Вторая нормальная форма

Отношение находится во 2НФ, если:

- оно находится в 1НФ
- таблица имеет Первичный Ключ (ПК)
- все неключевые столбцы таблицы должны зависеть от полного ключа (в случае если он составной)

Пример нормализации 2НФ:

модель	фирма	цена	скидка
SE	Apple	40000	5%
13	Apple	70000	5%
14	Apple	90000	5%
14 Pro	Apple	110000	5%
Galaxy S23	Samsung	105000	10%

Пожалуйста приведем эту таблицу ко 2НФ:

модель	id фирмы	цена
SE	1	40000
13	1	70000
14	1	90000
14 Pro	1	110000
Galaxy S23	2	105000

фирмы	
id	название
1	Apple
2	Samsung

скидки	
id фирмы	значение
1	5%
2	10%

Третья нормальная форма

Отношение находится в 3НФ, если:

- оно находится во 2НФ
- каждый неключевой столбец не зависит от других неключевых столбцов

Пример нормализации 3НФ:

футболисты			
ФИО	команда	город	гражданство
К.Роналду	Реал	Мадрид	Португалия
Л.Месси	Барселона	Барселона	Аргентина
К.Мбапе	ПСЖ	Париж	Франция

Давайте приведем эту таблицу к 3НФ:

футболисты		
ФИО	гражданство	id команды
К.Роналду	Португалия	1
Л.Месси	Аргентина	2
К.Мбапе	Франция	3

команды		
id	Название	город
1	Реал	Мадрид
2	Барселона	Барселона
3	ПСЖ	Париж

Нормальная форма Бойса-Кодда (BCNF)

Отношение находится в BCNF, если:

- оно находится в 3НФ
- ключевые атрибуты составного ключа не зависят от неключевых атрибутов

Пример нарушения BCNF:

Бронирования кортов			
Номер корта	Время начала	Время окончания	Тариф
1	09:30	10:30	«Корт 1 для членов клуба»
1	11:00	12:00	«Корт 1 для членов клуба»
1	14:00	15:30	«Корт 1 для не членов клуба»
2	10:00	11:30	«Корт 2 для не членов клуба»
2	11:30	13:30	«Корт 2 для не членов клуба»
2	15:00	16:30	«Корт 2 для членов клуба»

Давайте приведем эту таблицу к BCNF:

Тарифы		
Тариф	Номер корта	Для членов клуба
«Корт 1 для членов клуба»	1	Да
«Корт 1 для не членов клуба»	1	Нет
«Корт 2 для членов клуба»	2	Да
«Корт 2 для не членов клуба»	2	Нет

Бронирования		
Тариф	Время начала	Время окончания
«Корт 1 для членов клуба»	09:30	10:30
«Корт 1 для членов клуба»	11:00	12:00
«Корт 1 для не членов клуба»	14:00	15:30
«Корт 2 для не членов клуба»	10:00	11:30
«Корт 2 для не членов клуба»	11:30	13:00
«Корт 2 для членов клуба»	15:00	16:30

Четвертая нормальная форма

Отношение находится в 4НФ, если:

- оно находится в 3НФ
- в таблицах должны отсутствовать нетривиальные многозначные зависимости

Пример нарушения 4НФ:

Программисты		
ФИО	Язык	Хобби
Храпков К.А.	Ruby	Футбол
Храпков К.А.	Python	Хоккей
Богданов С.В.	Ruby	Теннис
Александров Б.А.	Python	Бокс

Программист ->-> Язык

Программист ->-> Хобби

Давайте приведем эту таблицу к 4НФ:

Хобби программистов		Языки, на которых пишут программисты	
ФИО	Хобби	ФИО	Язык
Храпков К.А.	Футбол	Храпков К.А.	Ruby
Храпков К.А.	Хоккей	Храпков К.А.	Python
Богданов С.В.	Теннис	Богданов С.В.	Ruby
Александров Б.А.	Бокс	Александров Б.А.	Python

Пятая нормальная форма

Отношение находится в 5НФ, если:

- оно находится в 4НФ
- в таблице каждая нетривиальная зависимость соединения определяется потенциальным ключом этой таблицы

Пример нарушения 5НФ:

Сотрудник	Организация	Направление
Федоров С.А.	Красный крест	Хирургия
Орлов Г.Ф.	Красный крест	Терапия
Орлов Г.Ф.	Здоровье	Кардиология
Митяев С.В.	Врачи без границ	Терапия
Федоров С.А.	Врачи без границ	Хирургия
Федоров С.А.	Мать и дитя	Хирургия

Давайте приведем эту таблицу к 5НФ:

Сотрудник	Организация	Организация	Направление
Федоров С.А.	Красный крест	Красный крест	Хирургия
Орлов Г.Ф.	Красный крест	Красный крест	Терапия
Орлов Г.Ф.	Здоровье	Здоровье	Кардиология
Митяев С.В.	Врачи без границ	Врачи без границ	Терапия
Федоров С.А.	Врачи без границ	Врачи без границ	Хирургия
Федоров С.А.	Мать и дитя	Мать и дитя	Хирургия

Сотрудник	Направление
Федоров С.А.	Хирургия
Орлов Г.Ф.	Терапия
Орлов Г.Ф.	Кардиология
Митяев С.В.	Терапия

Доменно-ключевая нормальная форма (DKNF)

Отношение находится в DKNF, если:

- оно находится в 5НФ
- каждое наложенное ограничение на таблицу должно являться логическим следствием ограничений доменов и ограничений ключей

Шестая нормальная форма

Отношение находится в 6НФ, если:

- оно находится в 5НФ
- таблица удовлетворяет всем нетривиальным зависимостям соединения

Пример нарушения 6НФ:

Работники			
<u>Таб.№</u>	<u>Время</u>	<u>Должность</u>	<u>Домашний адрес</u>
6575	01-01-2000:10-02-2003	слесарь	ул.Ленина,10
6575	11-02-2003:15-06-2006	слесарь	ул.Советская,22
6575	16-06-2006:05-03-2009	бригадир	ул.Советская,22

Давайте приведем эту таблицу к 6НФ:

Должности работников		
<u>Таб.№</u>	<u>Время</u>	Должность
6575	01-01-2000:10-02-2003	слесарь
6575	16-06-2006:05-03-2009	бригадир

Домашние адреса работников		
<u>Таб.№</u>	<u>Время</u>	Домашний адрес
6575	01-01-2000:10-02-2003	ул.Ленина,10
6575	11-02-2003:15-06-2006	ул.Советская,22

Денормализация - намеренное приведение структуры базы данных в состояние, не удовлетворяющее требованиям нормализации

ВЫВОДЫ

1. что ценного сегодня узнали?
2. какие методы будете использовать для решения своих задач?
3. закрепления материала + дз для практиков

ВОПРОСЫ

Оставьте отзыв!