

Погружение в СУБД. Сезон 2017

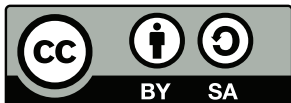
Проектирование схемы. Часть II

Дмитрий Барашев

Computer Science Center

Санкт-Петербург 2017

Эти материалы распространяются под лицензией
Creative Commons "Atribution - ShareAlike 4.0"



МОЖНО ИСПОЛЬЗОВАТЬ
с указанием авторства • с сохранением условий

Сверстано в Папирии



онлайн редактор для \LaTeX и Markdown • совместное
редактирование в реальном времени • интеграция с Git
репозиториями • графики

подсветка синтаксиса • автодополнение • проверка
орфографии • предпросмотр математических формул •
галерея шаблонов

Погружение в СУБД

Погружение в СУБД

Ключи

Погружение в СУБД

Ключи

Сущности и связи

Погружение в СУБД

Ключи

Сущности и связи

Внешние ключи

Погружение в СУБД

Ключи

Сущности и связи

Внешние ключи

Критерии качества схемы

Значения ключа уникальны

СУБД гарантирует

Conference	City
-------------------	-------------

VLDB'11	Seattle, US
---------	-------------

SIGMOD'14	Snowbird, US
-----------	--------------

SIGMOD'15	Melbourne, AU
-----------	---------------

Значения ключа уникальны

СУБД гарантирует

Conference	City
VLDB'11	Seattle, US
SIGMOD'14	Snowbird, US
SIGMOD'15	Melbourne, AU
VLDB'11	London, UK

Значения ключа уникальны

СУБД гарантирует

Conference	City
VLDB'11	Seattle, US
SIGMOD'14	Snowbird, US
SIGMOD'15	Melbourne, AU
VLDB'11	London, UK

Слишком узкий ключ

Uni	Researcher	Conference
Stanford	Jennifer	VLDB'11
MIT	Samuel	VLDB'11
ETH	Donald	SIGMOD'14
Stanford	Jennifer	SIGMOD'15

Слишком широкий ключ

A	B	C
a1	b1	c1

--

Слишком широкий ключ

A	B	C
a1	b1	c1
a1	b1	c2

Слишком широкий ключ

<hr/>		
A	B	C
a1	b1	c1
a1	b1	c2
a1	b3	c3
<hr/>		

Слишком широкий ключ

Conference	City	Date
VLDB'11	Seattle, US	Aug 29
SIGMOD'14	Snowbird, US	Jun 22
SIGMOD'15	Melbourne, AU	May 31

Слишком широкий ключ

Conference	City	Date
VLDB'11	Seattle, US	Aug 29
SIGMOD'14	Snowbird, US	Jun 22
SIGMOD'15	Melbourne, AU	May 31
SIGMOD'15	Melbourne, AU	Jun 31

Слишком широкий ключ

Conference	City	Date
VLDB'11	Seattle, US	Aug 29
SIGMOD'14	Snowbird, US	Jun 22
SIGMOD'15	Melbourne, AU	May 31
SIGMOD'15	Melbourne, AU	Jun 31
SIGMOD'15	Sydney, AU	May 31

Суррогатного ключ

Uni	Researcher	Conference
Stanford	Jennifer	VLDB'11
MIT	Samuel	VLDB'11
ETH	Donald	SIGMOD'14
Stanford	Jennifer	SIGMOD'15

Хранимые объекты

Researcher

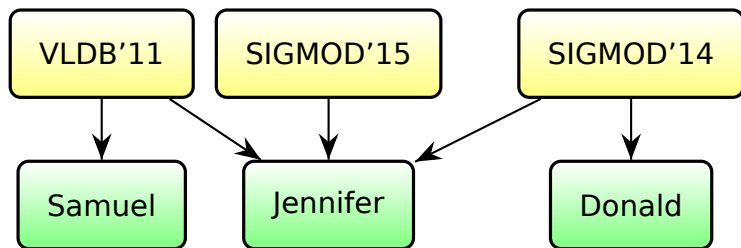
Uni	Name	Email
Stanford	Jennifer	jennifer@stanford.edu
MIT	Samuel	samuel@mit.edu
ETH	Donald	donald@ethz.ch
Waterloo	Gordon	gordon@uwaterloo.ca

Conference

City	Name
Seattle, US	VLDB'11
Snowbird, US	SIGMOD'14
Melbourne, AU	SIGMOD'15
Seattle, US	SIGIR'06

Ссылки между объектами

```
class Conference {  
    Researcher[] participants;  
}
```



Объект идентифицируется ключом

Researcher

Uni	Name	Email
Stanford	Jennifer	jennifer@stanford.edu
MIT	Samuel	samuel@mit.edu
ETH	Donald	donald@ethz.ch
Waterloo	Gordon	gordon@uwaterloo.ca

Conference

City	Name
Seattle, US	VLDB'11
Snowbird, US	SIGMOD'14
Melbourne, AU	SIGMOD'15
Seattle, US	SIGIR'06

Как же сделать список участников?

наивное решение

Researcher

Name

Jennifer

Samuel

Donald

Gordon

Conference

Name	Participants
------	--------------

VLDB'11	[Jennifer, Samuel]
---------	--------------------

SIGMOD'14	[Donald]
-----------	----------

SIGMOD'15	[Jennifer]
-----------	------------

SIGIR'06	[Gordon]
----------	----------

Как же сделать список участников?

наивное решение

Researcher

Name

Jennifer

Samuel

Donald

Gordon

Conference

Name

Participants

VLDB'11 [Jennifer, Samuel]

SIGMOD'14 [Donald, Alice]

SIGMOD'15 [Jennifer]

SIGIR'06 [Gordon]

Как же сделать список участников?

работоспособное решение

Participant

Researcher	Conference
Jennifer	VLDB'11
Samuel	VLDB'11
Donald	SIGMOD'14
Gordon	SIGIR'06
Jennifer	SIGMOD'15

Conference

Name	City
VLDB'11	Seattle, US
SIGMOD'14	Snowbird, US
SIGMOD'15	Melbourne, AU
SIGIR'06	Seattle, US

Чем это лучше?

что мешает Alice появиться в Participant?

Participant

Researcher	Conference
Jennifer	VLDB'11
Samuel	VLDB'11
Donald	SIGMOD'14
Gordon	SIGIR'06
Jennifer	SIGMOD'15
Alice	SIGMOD'14

Conference

Name	City
VLDB'11	Seattle, US
SIGMOD'14	Snowbird, US
SIGMOD'15	Melbourne, AU
SIGIR'06	Seattle, US

Внешний ключ¹

- ▶ Связывает два столбца из двух таблиц:

$$T_1.Откуда \rightarrow T_2.Куда$$

- ▶ Столбец *Куда* должен быть ключом в своей таблице T_2
- ▶ Внешний ключ даёт гарантию, что в столбце *Откуда* не может появиться значения, не имеющегося в столбце *Куда*

¹FOREIGN KEY

Применение внешнего ключа

Participant

Researcher	Conference
FK: Researcher.name	FK: Conference.name
Jennifer	VLDB'11
Samuel	VLDB'11
Donald	SIGMOD'14
Gordon	SIGIR'06
Jennifer	SIGMOD'15
Alice	SIGMOD'14

Researcher

Uni	Name
Stanford	Jennifer
MIT	Samuel
ETH	Donald
Waterloo	Gordon

Conference

Name	City
VLDB'11	Seattle, US
SIGMOD'14	Snowbird, US
SIGMOD'15	Melbourne, AU
SIGIR'06	Seattle, US

Связь 1:N

University

Name	Address
Key	
ETH	Switzerland
MIT	Massachusetts, US
Waterloo	Ontario, Canada
Stanford	California, US

Researcher

Uni	Name
FK: University.Name	Key
Stanford	Jennifer
MIT	Samuel
ETH	Donald
Waterloo	Gordon

Связь 1:N

University

Name	Address
Key	
ETH	Switzerland
MIT	Massachusetts, US
Waterloo	Ontario, Canada
Stanford	California, US

Researcher

Uni	Name
FK: University.Name	Key
Stanford	Jennifer
MIT	Samuel
ETH	Donald
Waterloo	Gordon
Stanford	Jeffrey
Stanford	Hector

Связь 1:N в отдельной таблице

University

Name	Address
Key	
ETH	Switzerland
MIT	Massachusetts, US
Waterloo	Ontario, Canada
Stanford	California, US

Researcher

Name	...
Key	
Jennifer	...
Samuel	...
Donald	...
Gordon	...

Employment

Uni	Researcher
FK: University.Name	FK: Researcher.Name
Stanford	Jennifer
MIT	Samuel
ETH	Donald
Waterloo	Gordon

Связь 1:N в отдельной таблице

University

Name	Address
Key	
ETH	Switzerland
MIT	Massachusetts, US
Waterloo	Ontario, Canada
Stanford	California, US

Researcher

Name	...
Key	
Jennifer	...
Samuel	...
Donald	...
Gordon	...

Employment

Uni	Researcher
FK: University.Name	FK: Researcher.Name
Key	
Stanford	Jennifer
MIT	Samuel
ETH	Donald
Waterloo	Gordon

Было

Uni	Address	Name	Email	Conference	City
Stanford	California, US	Jennifer	jennifer@stanford.edu	VLDB'11	Seattle, US
ETH	Switzerland	Donald	donald@ethz.ch	SIGMOD'14	Snowbird, US

Стало

Детерминант в ФЗ

$$X \rightarrow Y$$

Тривиальная ФЗ

$$X \rightarrow X$$

Нормальная форма Бойса-Кодда

В любой нетривиальной
функциональной
зависимости
детерминантом является ключ.

Грабли №1: зависимость от подключа

Uni	Researcher	Conference	City
— Key —			
Stanford	Jennifer	VLDB'11	Seattle, US
MIT	Samuel	VLDB'11	Seattle, US
ETH	Donald	SIGMOD'14	Snowbird, US
Stanford	Jennifer	SIGMOD'15	Melbourne, AU

Габли №1: зависимость от подключа

Uni	Researcher	Conference	City
— Key —			
Stanford	Jennifer	VLDB'11	Seattle, US
MIT	Samuel	VLDB'11	Seattle, US
ETH	Donald	SIGMOD'14	Snowbird, US
Stanford	Jennifer	SIGMOD'15	Melbourne, AU

Researcher, Conference → Uni

Researcher, Conference → City

Габли №1: зависимость от подключа

Uni	Researcher	Conference	City
— Key —			
Stanford	Jennifer	VLDB'11	Seattle, US
MIT	Samuel	VLDB'11	Seattle, US
ETH	Donald	SIGMOD'14	Snowbird, US
Stanford	Jennifer	SIGMOD'15	Melbourne, AU

Researcher, Conference → Uni

Researcher, Conference → City

Conference → City

Researcher → Uni

Грабли №2: транзитивные зависимости

Researcher

Uni	Address	Researcher Key
Stanford	California, US	Jennifer
Stanford	California, US	Jeffrey
Stanford	California, US	Hector

Габли №2: транзитивные зависимости

Researcher

Uni	Address	Researcher Key
Stanford	California, US	Jennifer
Stanford	California, US	Jeffrey
Stanford	California, US	Hector

Researcher \rightarrow Uni

Researcher \rightarrow Address

Габри №2: транзитивные зависимости

Researcher

Uni	Address	Researcher Key
Stanford	California, US	Jennifer
Stanford	California, US	Jeffrey
Stanford	California, US	Hector

Researcher \rightarrow Uni

Researcher \rightarrow Address

Uni \rightarrow Address

Габли №3: ключей может быть много

Researcher

Uni	Researcher Key №1	Email
Stanford	Jennifer	jen@stanford.edu
MIT	Samuel	samuel@mit.edu
ETH	Donald	donald@ethz.ch
Stanford	Jens	jen@stanford.edu

Габли №3: ключей может быть много

Researcher

Uni	Researcher Key №1	Email Key №2
Stanford	Jennifer	jen@stanford.edu
MIT	Samuel	samuel@mit.edu
ETH	Donald	donald@ethz.ch
Stanford	Jens	jen@stanford.edu

В таблице может быть не более одного первичного ключа² но сколько угодно потенциальных ключей³

²ключевое слово PRIMARY KEY в SQL

³candidate key, ключевое слово UNIQUE в SQL

Грабли №4: связь 1:N вместо M:N

Researcher

Uni	Researcher	Conference
FK: University.Name	Key	FK: Conference.Name
Stanford	Jennifer	VLDB'11
MIT	Samuel	VLDB'11
ETH	Donald	SIGMOD'14

Габли №4: связь 1:N вместо M:N

Researcher

Uni	Researcher	Conference
FK: University.Name	Key	FK: Conference.Name
Stanford	Jennifer	VLDB'11
MIT	Samuel	VLDB'11
ETH	Donald	SIGMOD'14
Stanford	Jennifer	SIGMOD'15

Габли №5: связь M:N вместо 1:N

Venue

Id	Name	Address	City
Key №1	Key №2		
1	Westin Hotel	1900 Fifth Avenue	Seattle
2	Melbourne Convention & Exhibition Centre	1 Convention Centre Place, South Wharf	Melbourne

ConferenceVenue

Venueld	Conference
FK: Venue.Id	FK: Conference.Name
1	VLDB'11
2	SIGMOD'15

Габли №5: связь M:N вместо 1:N

Venue

Id	Name	Address	City
Key №1	Key №2		
1	Westin Hotel	1900 Fifth Avenue	Seattle
2	Melbourne Convention & Exhibition Centre	1 Convention Centre Place, South Wharf	Melbourne

ConferenceVenue

Venueld	Conference
FK: Venue.Id	FK: Conference.Name
1	VLDB'11
2	SIGMOD'15
2	VLDB'11

Габли №6: суррогатный ключ

Venue

Id	Name	Address	City
Key №1	Key №2		
1	Westin Hotel	1900 Fifth Avenue	Seattle
2	Melbourne Convention & Exhibition Centre	1 Convention Centre Place, South Wharf	Melbourne

Габли №6: суррогатный ключ

Venue

Id	Name	Address	City
Key №1	Key №2		
100500	Westin Hotel	1900 Fifth Avenue	Seattle
42	Melbourne Convention & Exhibition Centre	1 Convention Centre Place, South Wharf	Melbourne

Грабли №6: суррогатный ключ

- ▶ AUTO_INCREMENT в MySQL
- ▶ SERIAL в PostgreSQL
- ▶ IDENTITY в Microsoft SQL Server и IBM DB2
- ▶ SEQUENCE в Oracle

Грабли №6: суррогатный ключ

- ▶ Удобно использовать для связей
- ▶ Занимает мало места на диске
- ▶ Часто является первичным ключом

Грабли №6: суррогатный ключ

- ▶ Удобно использовать для связей
- ▶ Занимает мало места на диске
- ▶ Часто является первичным ключом
- ▶ Часто только им и ограничиваются
- ▶ Часто делают ложные предположения о его значениях

Габли №6: суррогатный ключ

Researcher

Uni	Researcher	Email	Id
			PRIMARY KEY
Stanford	Jennifer	jennifer@stanford.edu	1
MIT	Samuel	samuel@mit.edu	2
ETH	Donald	donald@ethz.ch	3

Габли №6: суррогатный ключ

Researcher

Uni	Researcher	Email	Id
	UNIQUE	UNIQUE	PRIMARY KEY
Stanford	Jennifer	jennifer@stanford.edu	1
MIT	Samuel	samuel@mit.edu	2
ETH	Donald	donald@ethz.ch	3

Проектируя схему БД

Объявляйте естественные ключи

Проектируя схему БД

Убедитесь, что таблицы находятся в нормальной форме Бойса-Кодда

Проектируя схему БД

Проверяйте, правильные ли
получились связи и объявлены ли
внешние ключи

Проектируя схему БД

Помните, что суррогатные ключи не гарантируют ничего, кроме уникальности значений