

Проектирование БД

Подходы в проектировании структур данных

Восходящий подход

Сбор информации об объектах решаемой задачи в рамках одной таблицы и последующая декомпозиция ее на несколько взаимосвязанных таблиц на основе процедуры нормализации отношений.

Нисходящий подход

Формулирование знаний о системе и требований к обработке данных, получение с помощью CASE-системы готовой схемы БД.

Избыточность

Избыточное дублирование данных создает проблемы при попытке удаления, добавления или редактирования данных в таблице.

| Сотрудник | Телефон | Кабинет |
|-------------|---------|---------|
| Иванов И.М. | 3721 | 109 |
| Петров М.И. | 4328 | 111 |
| Сидоров Н.П | 4328 | 111 |
| Егоров В.В. | 4328 | 111 |

Аномалии модификации

Изменение значения одного данного влечёт за собой просмотр всей таблицы и соответствующее изменение некоторых других записей таблицы.

| Сотрудник | Телефон | Кабинет |
|--------------|-----------------|---------|
| Иванов И.М. | 3721 | 109 |
| Петров М.И. | 4328 | 111 |
| Сидоров Н.П. | 4328 | 111 |
| Егоров В.В. | 4328 | 111 |



Аномалии удаления

При удалении какого-либо данного из таблицы может пропасть и другая информация, которая не связана напрямую с удаляемыми данными.

| Сотрудник | Телефон | Кабинет |
|------------------------|-----------------|----------------|
| Иванов И.М. | 3721 | 109 |
| Петров М.И. | 4328 | 111 |
| Сидоров Н.П | 4328 | 111 |
| Егоров В.В. | 4328 | 111 |

Аномалии добавления

Информацию нельзя поместить в таблицу до тех пор, пока она неполная, либо вставка новой записи требует дополнительного просмотра таблицы.

1)

| Сотрудник | Телефон | Кабинет |
|-------------|---------|---------|
| Иванов И.М. | 3721 | 109 |
| Петров М.И. | 4328 | 111 |
| Сидоров Н.П | 4328 | 111 |
| Егоров В.В. | 4328 | 111 |
| | 4332 | 110 |

2)

| Сотрудник | Телефон | Кабинет |
|--------------|---------|---------|
| Иванов И.М. | 3721✓ | 109✓ |
| Петров М.И. | 4328 | 111 |
| Сидоров Н.П | 4328 | 111 |
| Егоров В.В. | 4328 | 111 |
| Стёпкин Е.Н. | 3721 | 109 |

Аномалии в базах данных

Аномалия – ситуацию в таблицах БД, которая приводит к *противоречиям в БД* либо *существенно усложняет обработку данных*.

Виды аномалий:

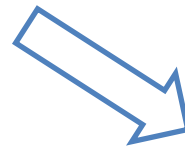
- Аномалии модификации
- Аномалии удаления
- Аномалии добавления

Решение

| Сотрудник | Телефон | Кабинет |
|-------------|---------|---------|
| Иванов И.М. | 3721 | 109 |
| Петров М.И. | 4328 | 111 |
| Сидоров Н.П | 4328 | 111 |
| Егоров В.В. | 4328 | 111 |



| Телефон | Кабинет |
|---------|---------|
| 3721 | 109 |
| 4328 | 111 |



| Сотрудник | Кабинет |
|--------------|---------|
| Иванов И.М. | 109 |
| Петров М.И. | 111 |
| Сидоров Н.Г. | 111 |
| Егоров В.В. | 111 |

Декомпозиция

Декомпозиция (разбиение) таблицы – процесс деления таблицы на несколько таблиц с целью устранения избыточности данных.

Принцип нормализации

| Номер судна | Название | Рейс | Дата погруз. | Порт погруз. | Дата прибыт. | Порт прибыт. | Капитан | Вид судна | Тоннаж |
|-------------|-----------|------|--------------|--------------|--------------|--------------|---------|-----------|--------|
| 526 | Победа | 9201 | 31.05.18 | SFO | 06.06.18 | HNL | Емелин | Сухогруз | 500 |
| 603 | Гордый | 9202 | 05.05.18 | OAK | 19.06.18 | OSA | Крылов | Ролкер | 1000 |
| 531 | Жемчужина | 9203 | 20.06.18 | LAX | 10.07.18 | PAP | Мухин | Универсал | 1500 |
| 526 | Победа | 9204 | 20.08.18 | SFO | 27.08.18 | HNL | Емелин | Сухогруз | 500 |

| Номер судна | Название | Капитан | Вид судна | Тоннаж |
|-------------|-----------|---------|-----------|--------|
| 526 | Победа | Емелин | Сухогруз | 500 |
| 603 | Гордый | Крылов | Ролкер | 1000 |
| 531 | Жемчужина | Мухин | Универсал | 1500 |

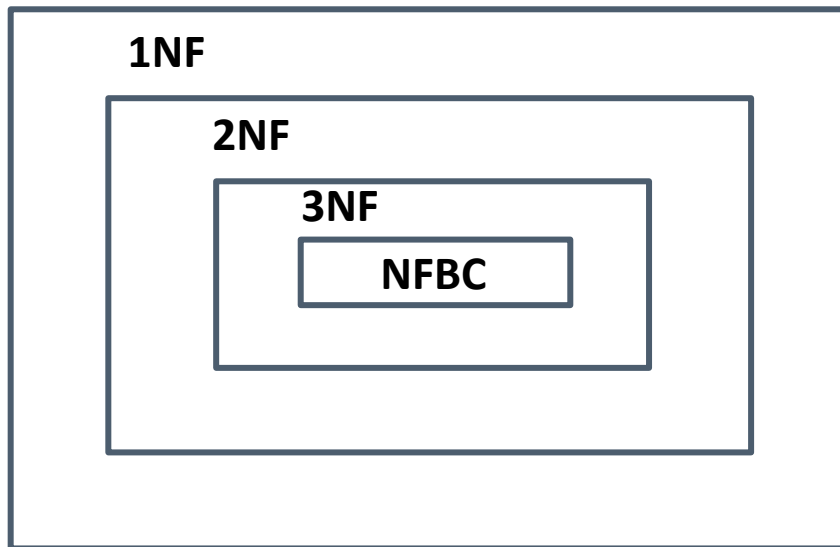
| Номер судна | Название | Капитан | Вид судна |
|-------------|-----------|---------|-----------|
| 526 | Победа | Емелин | Сухогруз |
| 603 | Гордый | Крылов | Ролкер |
| 531 | Жемчужина | Мухин | Универсал |

| Рейс | Дата погруз. | Порт погрузки. | Дата прибыт. | Порт прибыт. | Номер судна |
|------|--------------|----------------|--------------|--------------|-------------|
| 9201 | 31.05.18 | SFO | 06.06.18 | HNL | 526 |
| 9202 | 05.05.18 | OAK | 19.06.18 | OSA | 603 |
| 9203 | 20.06.18 | LAX | 10.07.18 | PAP | 531 |
| 9204 | 20.08.18 | SFO | 27.08.18 | HNL | 526 |

Назначение нормализации

- Исключить избыточность;
- Устранить некоторые аномалии;
- Разработать в проекте «качественное», интуитивно понятное и легко расширяемое представлением реального мира;
- Упростить задание ограничений целостности.

Нормальные формы данных (НФ)



Нормальные формы

Нормальная форма

- свойство отношения (таблицы), характеризующее его с точки зрения избыточности, потенциально приводящей к логически ошибочным результатам выборки или изменения данных;
- совокупность требований, которым должно удовлетворять отношение.

Нормальные формы:

- Первая нормальная форма (1НФ);
- Вторая нормальная форма (2НФ);
- Третья нормальная форма (3НФ);
- Нормальная форма Бойса-Кодда (БКНФ);
- Четвёртая нормальная форма (4НФ);
- Пятая нормальная форма (5НФ).

Требования первой нормальной формы (1NF)

Чтобы база данных находилась в 1 нормальной форме, необходимо чтобы ее таблицы соблюдали следующие реляционные принципы:

- В таблице не должно быть дублирующих строк
- В каждой ячейке таблицы хранится атомарное значение (одно не составное значение)
- В столбце хранятся данные одного типа

Таблица сотрудников в ненормализованном виде

| Сотрудник | Контакт |
|--------------|---|
| Иванов И.И. | 123-456-789, 987-654-321 |
| Сергеев С.С. | Рабочий телефон 555-666-777, Домашний телефон 777-888-999 |
| John Smith | 123-456-789 |
| John Smith | 123-456-789 |

Таблица сотрудников в первой нормальной форме

| Сотрудник | Телефон | Тип телефона |
|--------------|-------------|------------------|
| Иванов И.И. | 123-456-789 | |
| Иванов И.И. | 987-654-321 | |
| Сергеев С.С. | 555-666-777 | Рабочий телефон |
| Сергеев С.С. | 777-888-999 | Домашний телефон |
| John Smith | 123-456-789 | |

Требования второй нормальной формы (2NF)

Чтобы база данных находилась во второй нормальной форме (2NF), необходимо чтобы ее таблицы удовлетворяли следующим требованиям:

- Таблица должна находиться в первой нормальной форме
- Таблица должна иметь ключ
- Все неключевые столбцы таблицы должны зависеть от полного ключа (*в случае если он составной*)

Таблица сотрудников в первой нормальной форме

| Название проекта | Участник | Должность | Срок проекта (мес.) |
|--------------------------|--------------|-------------|---------------------|
| Внедрение приложения | Иванов И.И. | Программист | 8 |
| Внедрение приложения | Сергеев С.С. | Бухгалтер | 8 |
| Внедрение приложения | John Smith | Менеджер | 8 |
| Открытие нового магазина | Сергеев С.С. | Бухгалтер | 12 |
| Открытие нового магазина | John Smith | Менеджер | 12 |

| Идентификатор проекта | Название проекта | Срок проекта (мес.) |
|-----------------------|--------------------------|---------------------|
| 1 | Внедрение приложения | 8 |
| 2 | Открытие нового магазина | 12 |

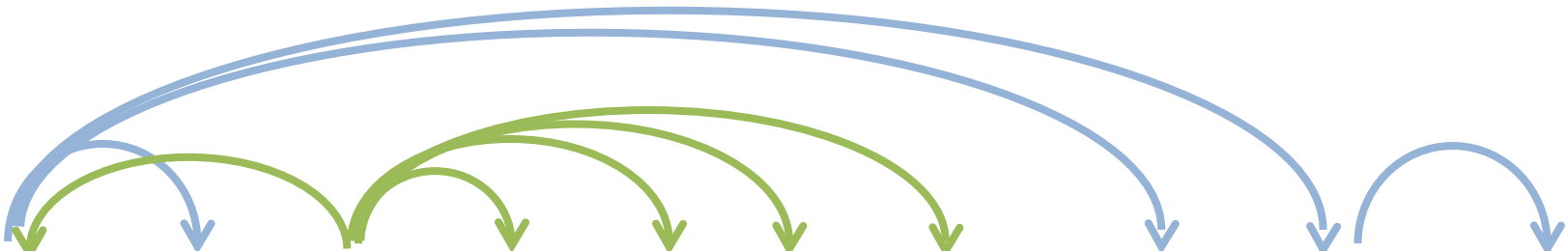
Проекты

| Идентификатор участника | Участник | Должность |
|-------------------------|--------------|-------------|
| 1 | Иванов И.И. | Программист |
| 2 | Сергеев С.С. | Бухгалтер |
| 3 | John Smith | Менеджер |

Участники

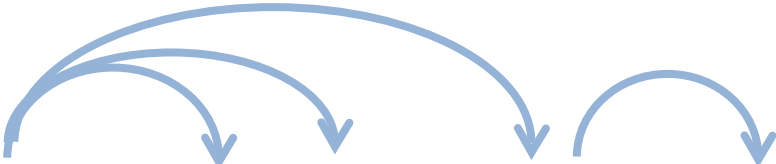
| Идентификатор проекта | Идентификатор участника |
|-----------------------|-------------------------|
| 1 | 1 |
| 1 | 2 |
| 1 | 3 |
| 2 | 2 |
| 2 | 3 |

Связь
проектов и
участников



| Номер судна | Название | Рейс | Дата погруз. | Порт погруз. | Дата прибыт. | Порт прибыт. | Капитан | Вид судна | Тоннаж |
|-------------|-----------|------|--------------|--------------|--------------|--------------|---------|-----------|--------|
| 526 | Победа | 9201 | 31.05.18 | SFO | 06.06.18 | HNL | Емелин | Сухогруз | 500 |
| 603 | Гордый | 9202 | 05.05.18 | OAK | 19.06.18 | OSA | Крылов | Ролкер | 1000 |
| 531 | Жемчужина | 9203 | 20.06.18 | LAX | 10.07.18 | PAP | Мухин | Универсал | 1500 |
| 526 | Победа | 9204 | 20.08.18 | SFO | 27.08.18 | HNL | Емелин | Сухогруз | 500 |

Вторая нормальная форма



| Номер судна | Название | Капитан | Вид судна | Тоннаж |
|-------------|-----------|---------|-----------|--------|
| 526 | Победа | Емелин | Сухогруз | 500 |
| 603 | Гордый | Крылов | Ролкер | 1000 |
| 531 | Жемчужина | Мухин | Универсал | 1500 |

Суда

| Рейс | Дата погруз. | Порт погруз. | Дата прибыт. | Порт прибыт. | Номер судна |
|------|--------------|--------------|--------------|--------------|-------------|
| 9201 | 31.05.18 | SFO | 06.06.18 | HNL | 526 |
| 9202 | 05.05.18 | OAK | 19.06.18 | OSA | 603 |
| 9203 | 20.06.18 | LAX | 10.07.18 | PAP | 531 |
| 9204 | 20.08.18 | SFO | 27.08.18 | HNL | 526 |

Рейсы

Третья нормальная форма

ЗНФ – все неключевые атрибуты таблицы взаимно независимы и полностью зависят от первичного ключа.

Третья нормальная форма

| Номер судна | Название | Капитан | Вид судна |
|-------------|-----------|---------|-----------|
| 526 | Победа | Емелин | Сухогруз |
| 603 | Гордый | Крылов | Ролкер |
| 531 | Жемчужина | Мухин | Универсал |

Суда

| Вид судна | Тоннаж |
|-----------|--------|
| Сухогруз | 500 |
| Ролкер | 1000 |
| Универсал | 1500 |

Грузоподъемность

| Рейс | Дата погруз. | Порт погруз. | Дата прибыт. | Порт прибыт. | Номер судна |
|------|--------------|--------------|--------------|--------------|-------------|
| 9201 | 31.05.18 | SFO | 06.06.18 | HNL | 526 |
| 9202 | 05.05.18 | OAK | 19.06.18 | OSA | 603 |
| 9203 | 20.06.18 | LAX | 10.07.18 | PAP | 531 |
| 9204 | 20.08.18 | SFO | 27.08.18 | HNL | 526 |

Рейсы

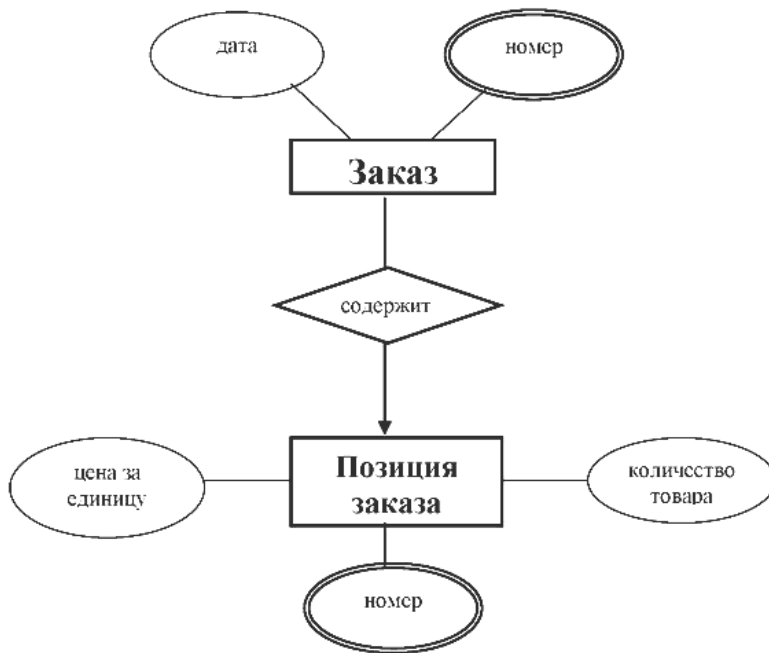
ER-модель

(модель «сущность-связь»)

Структура данных отображается графически, в виде диаграммы сущностей и связей, состоящей из:

- Сущностей
- Атрибутов
- Связей

Разные нотации



Нотация Чена

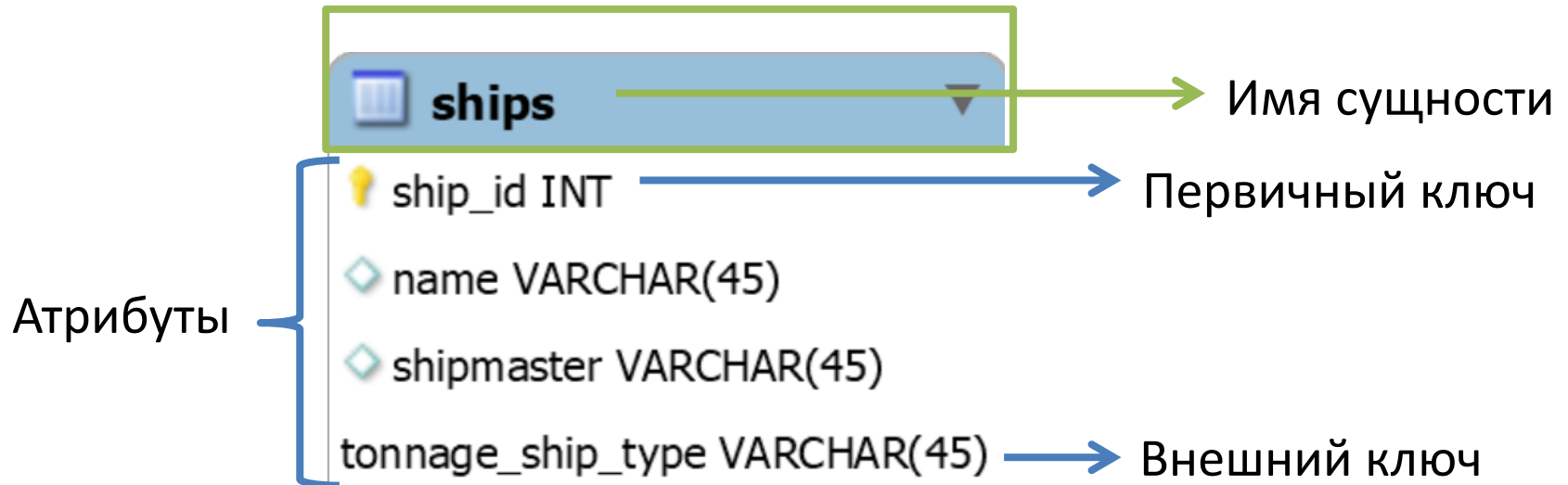


Crow's Foot
(«воронья лапка»)

Сущности и атрибуты

- **Сущность** - это реальный или представляемый *объект, информация о котором должна сохраняться* и быть доступна.
- **Атрибутом** сущности является любая *деталь*, которая служит для уточнения, идентификации, классификации, числовой характеристики или выражения состояния сущности

Сущности и атрибуты



Внешний ключ

Внешние ключи служат для связи таблиц базы данных между собой.

Таблица, в которой определен **внешний ключ**, ссылается на таблицу, в которой этот же атрибут является **первичным ключом**.

Связи

Связь – это отношение одной сущности к другой или к самой себе.

Модальность связей:

- Обязательная;
- Необязательная.

Типы связей:

- Один к одному;
- Один ко многим;
- Многие ко многим.

СВЯЗИ

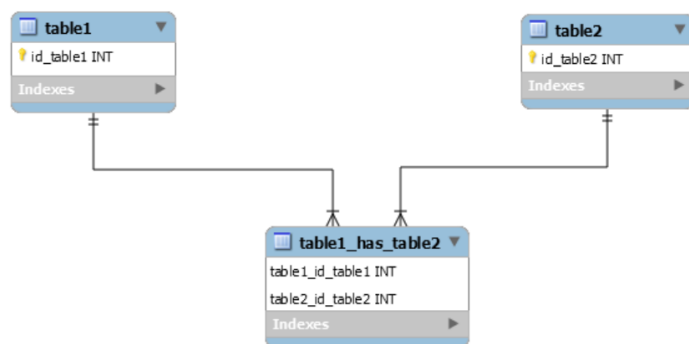
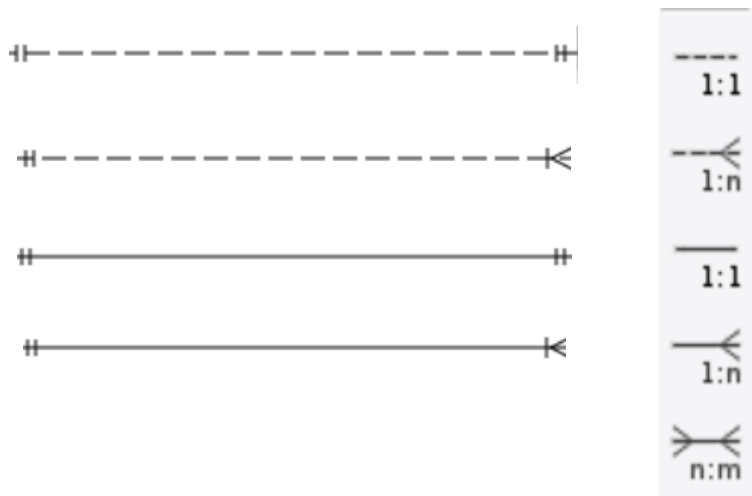


Схема БД для судоходной компании

Суда торговой компании совершают рейсы. О каждом рейсе известна дата и порт погрузки, а также дата и порт прибытия. У каждого судна есть уникальный номер и имя. Кроме того, одним и тем же судном управляет один капитан. Грузоподъёмность судна определяется его типом.

Схема БД для судоходной компании

Суда

- Номер судна
- Название
- Капитан
- Вид судна

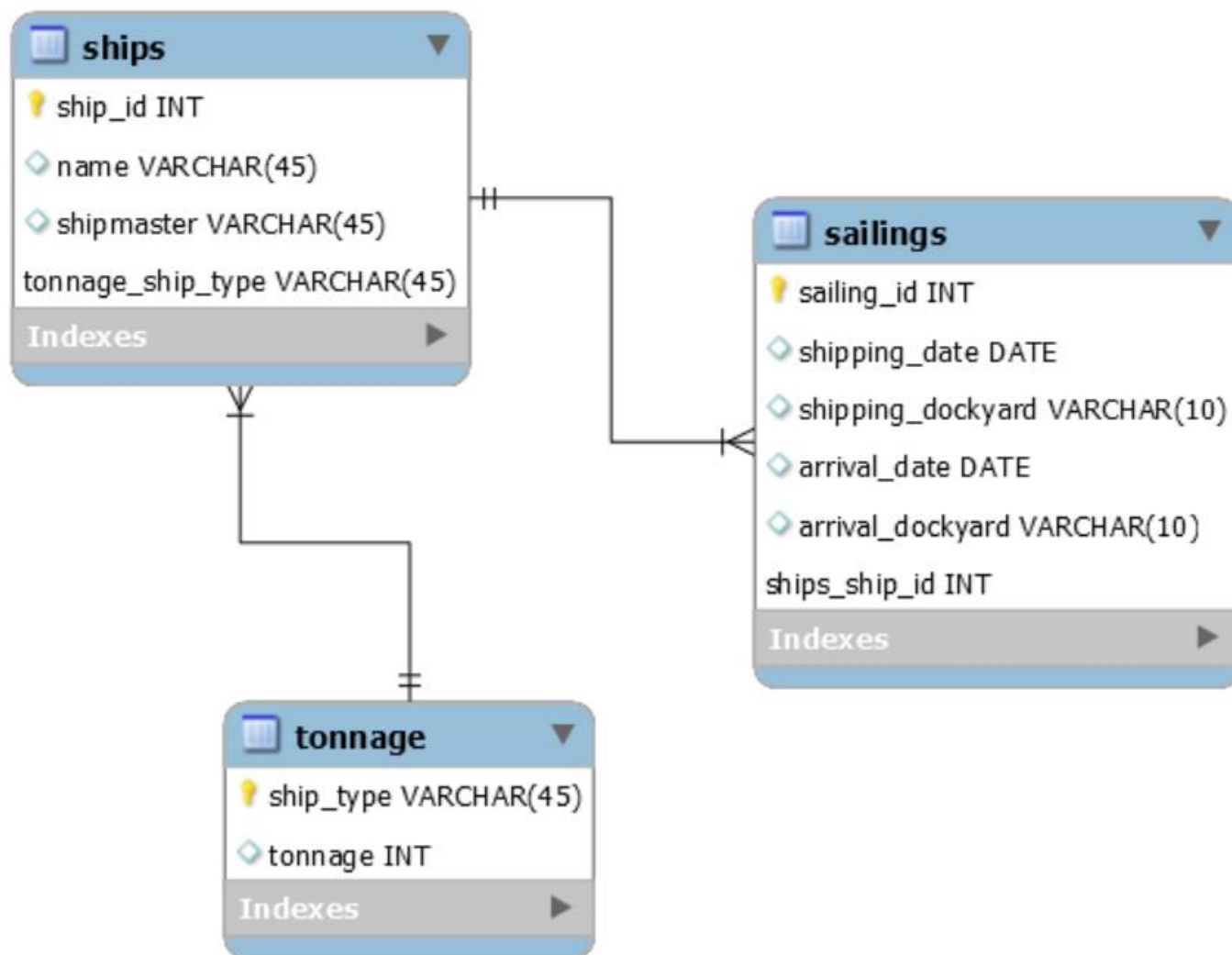
Тоннаж

- Вид судна
- Тоннаж

Рейсы

- Рейс
- Дата погруз.
- Порт погруз.
- Дата прибыт.
- Порт прибыт.
- Номер судна

Схема БД для судоходной компании



Рекомендации по разработке структур

- «Каждой сущности - отдельную таблицу»;
- Использование ключей;
- Приведение таблиц в 3 НФ;
- Удаление всех ссылок на удаляемую таблицу;
- Связи должны быть бинарными;
- Исключение из таблиц связей характеристик сущностей.
- Контроль правильности ссылок на существующие объекты.