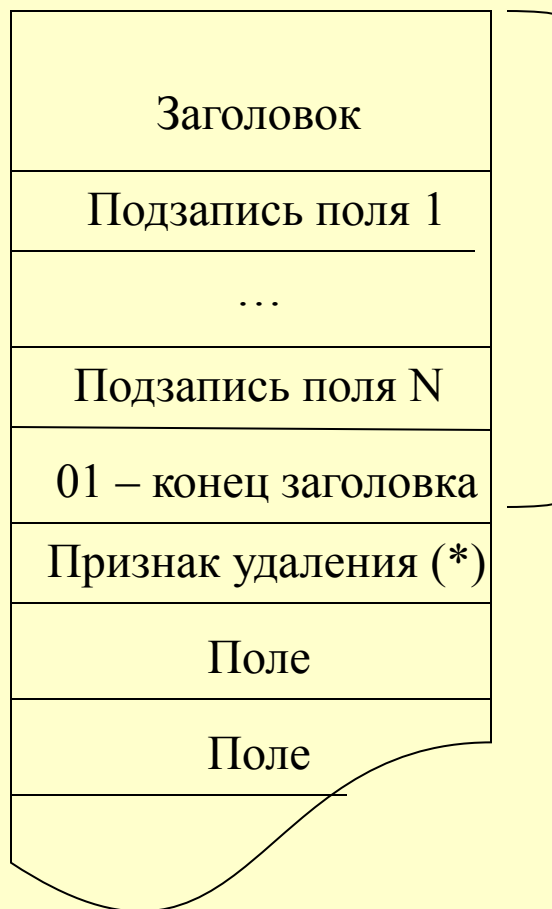


Лекция 20

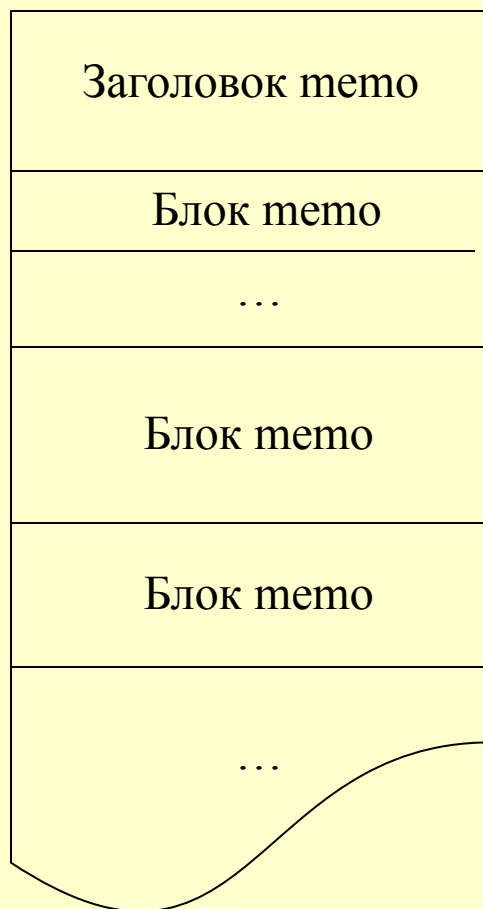
Примеры моделей хранения и организации доступа к БД (dBase, MS SQL Server, Oracle)

Физическая структура данных в dBase

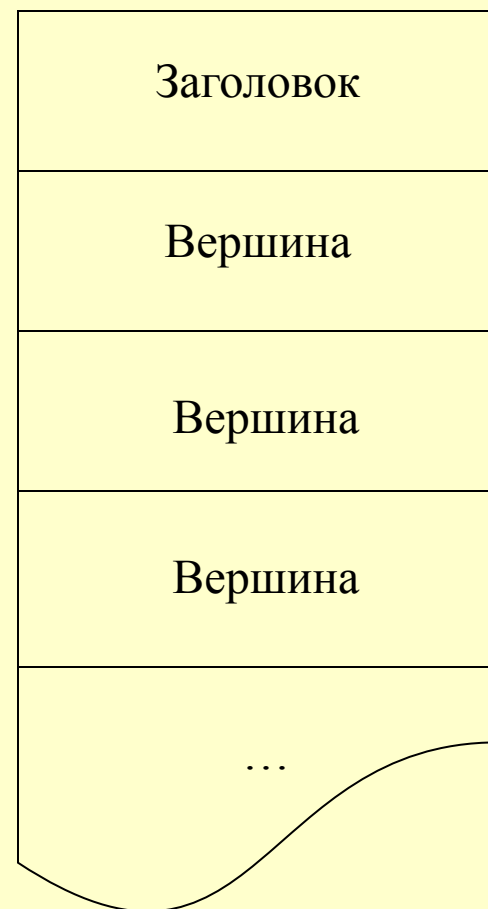
Основной файл базы данных



мето-файл для хранения длинных полей



Индексный файл



Структура заголовка файла данных

Байты	Описание
00	Типы файлов с данными
01-03	Последнее изменение (ГГММДД)
04-07	Число записей в файле
08-09	Положение первой записи с данными
10-11	Длина одной записи с данными (включая признак удаления)
12-27	Зарезервированы
28	1 - есть составной индексный файл (типа .CDX), 0-нет
29-31	Зарезервированы
32-n	Подзаписи полей (для каждого поля - одна подзапись)
N+1	Признак завершения записи заголовка (0x01)

Структура подзаписи поля

Байты	Описание
00-10	Название поля (максимально - 10 символов)
11	Тип данных: С - символьное; N - числовое; L - логическое; M - типа memo; D - дата; F - с плавающей точкой; P - шаблон.
12-15	Расположение поля внутри записи
16	Длина поля (в байтах)
17-31	Зарезервированы

Структура мемо-файла

Структура заголовка мемо-файла

<i>Байты</i>	<i>Описание</i>
00-03	Расположение следующего свободного блока
04-05	Не используются
06-07	Размер блока (число байтов в блоке)
08-511	Не используются

Заголовок блока мемо и текст мемо

00-03	Сигнатура блока (тип данных в блоке): 0 - шаблон (поле типа шаблон) / 1 - текст (поле типа мемо)
04-07	Длина мемо (в байтах)
08-n	Текст мемо (n=длина)

Все мемо-блоки начинаются с четных адресов границ блоков

Структура индексного файла

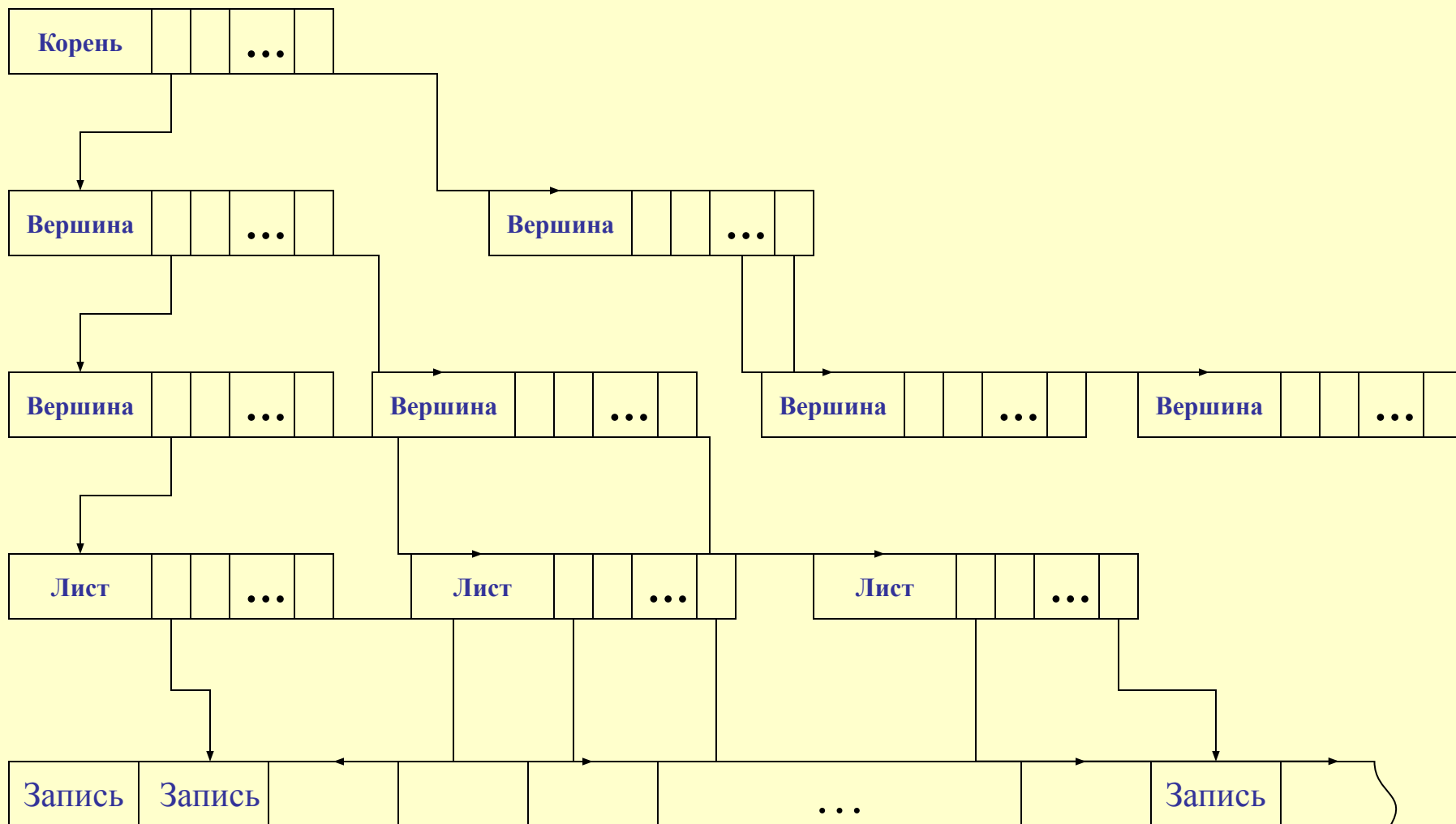
Запись заголовка индексного файла

<i>Байты</i>	<i>Описание</i>
00-03	Указатель на корневую вершину
04-07	Указатель на свободную в списке вершину (-1, если таковая отсутствует)
08-11	Указатель на конец файла (размер файла)
12-13	Длина ключа
14	Особенности индекса : 1 - уникальный индекс; 8 - индекс имеет дополнительный оператор FOR.
15	Сигнатура индекса
16-235	Ключевое выражение (не компилируется; до 220 символов)
236-455	Выражение FOR (не компилируется; до 220 символов, оканчивающееся пустым символом)
456-511	Не используются

Структура записи вершины индекса

Байты	Описание
00-01	Атрибуты вершины: 0 - вершина индекса; 1 - корневая вершина; 2 - лист.
02-03	Количество ключей (0, 1 или больше)
04-07	Указатель на вершину, расположенную непосредственно слева от данной вершины (на том же уровне; -1, если отсутствует)
08-11	Указатель на вершину, расположенную непосредственно справа от данной вершины (на том же уровне; -1, если отсутствует)
12-511	До 500 символов, включающих в себя перечень значений ключа с четырехбайтовым шестнадцатиричным числом (хранящемся в обычном формате слева направо). Если вершина является листом, тогда четыре байта содержат подлинный номер в базе данных в шестнадцатиричном формате.

Структура индекса



Файл данных

Физическая структура данных в MS SQL Server

База данных

Файлы БД

Экстенд

Страница

Журнал транзакций

Файл данных

Однородный

Смешанный

Страница Данных

Индексная Страница

Текстовая Страница

Страница Журнала Транзакций

Карта Распределения Блоков

Карта Свободного Пространства

Индексная Карта Размещения

Типы страниц SQL Server

Страницы, относящиеся к хранению и поиску информации:

- страницы данных;
- индексные страницы;
- текстовые страницы;
- страницы журнала транзакций;

Страницы размещения:

- карты распределения блоков (основная и вторичная);
- карты свободного пространства;
- индексные карты размещения.

Структура страницы



- номер страницы
- идентификатор объекта, которому принадлежит страница;
- индекс и уровень внутри индексного дерева, которому принадлежит страница;
- количество записей на странице;
- общий объем свободного пространства на странице;
- указатель на свободное пространство после последней записи на странице;
- минимальная длина строки на странице;
- объем зарезервированного пространства.

Страницы размещения

Карты распределения экстентов

заголовок + битовый массив
(64000 бит)

1 бит - 1 экстент

1 страница - 64 000 экстентов
(4 Гбайт данных)

глобальная карта распределения
(Global Allocation Map, GAM)

бит = 0, экстент занят

бит = 1, экстент свободен

вторичная глобальная карта
распределения

(Secondary global allocation map,
SGAM)

бит = 1, экстент смешанный

Карты свободного пространства

Page free space (PFS) -
степень заполнения
страниц.

1 PFS - 8000 страниц

1 байт - 1 страница

PFS повторяются
через каждые 8000
страниц

Индексные карты размещения

Index Allocation Map, IAM
- связь между экстентами
и расположенными на них

объектами

объект - одна или более
страниц IAM

IAM объединены друг с
другом в цепочку
двунаправленными
ссылками.

IAM - битовая карта:

Бит = 1, в экстенте есть
страницы, принадлежащие
данному объекту,

Бит = 0 — нет.

Страница данных

Заголовок

Таблица размещения записей (слотов)
1 слот = 4 байта
(смещение на странице и длина записи)

Поля фиксированной длины хранятся до полей переменной длины и занимают свою полную длину.

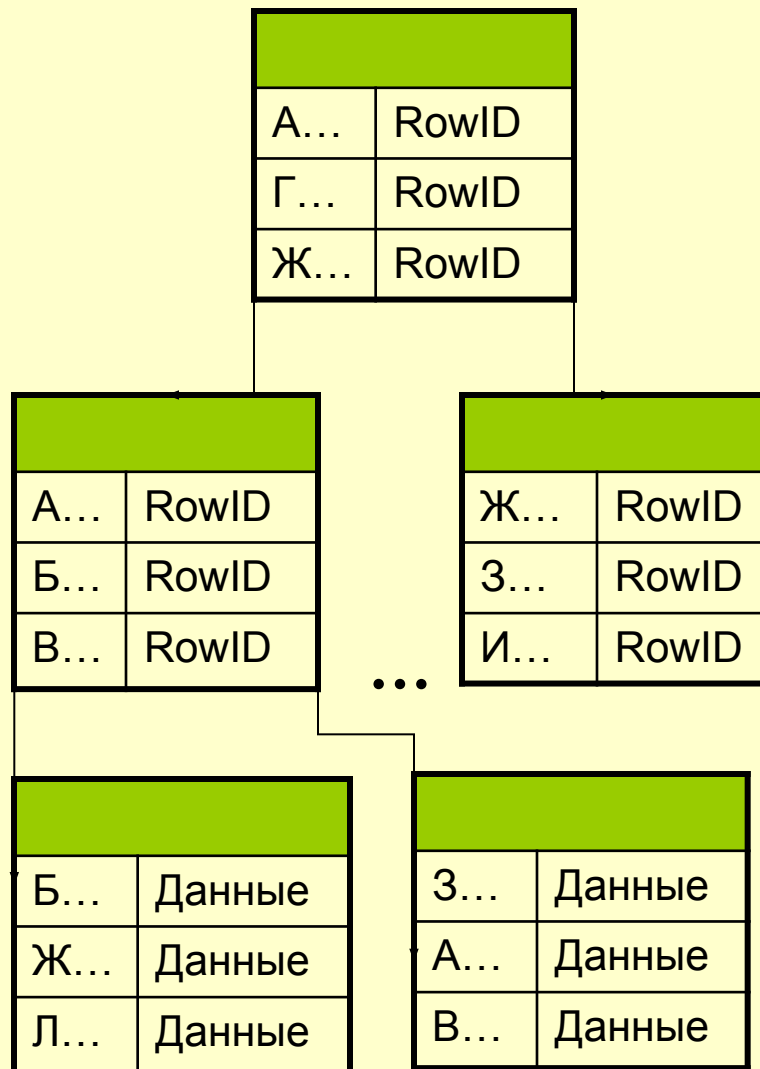
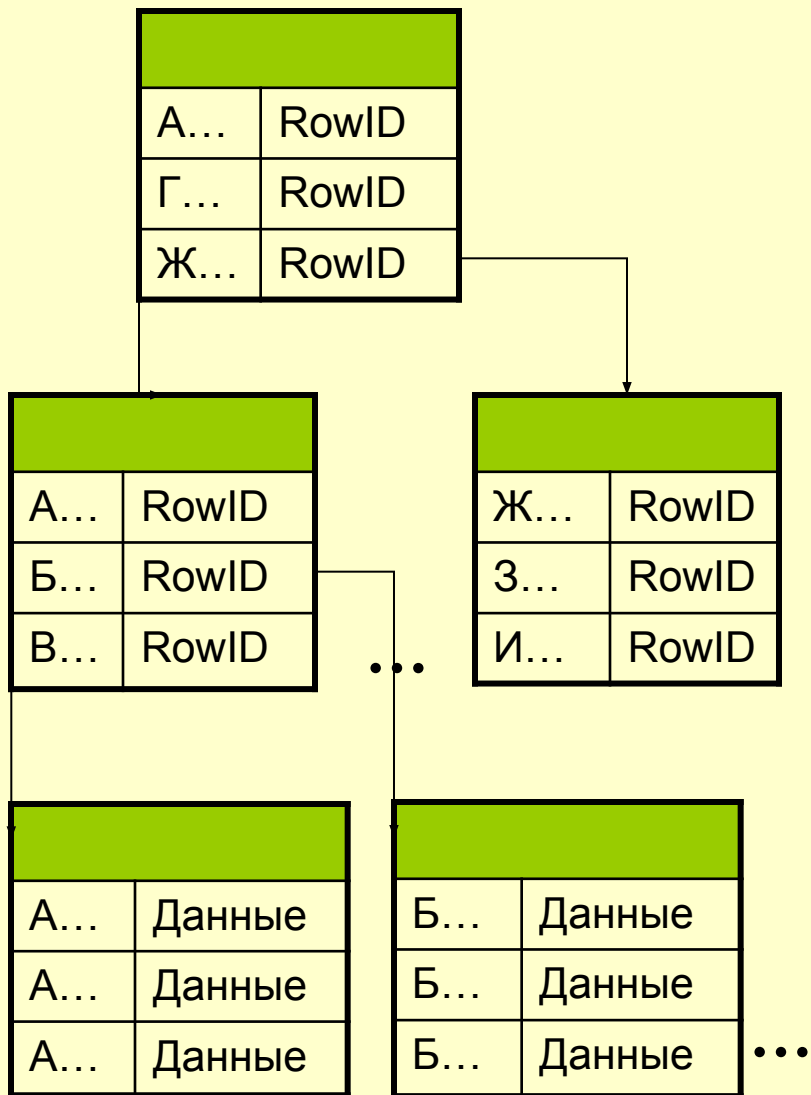
Значение NULL задается специальным флагом.

В каждой записи хранится общая длина записи и текущие длины полей переменной длины.

Индексы

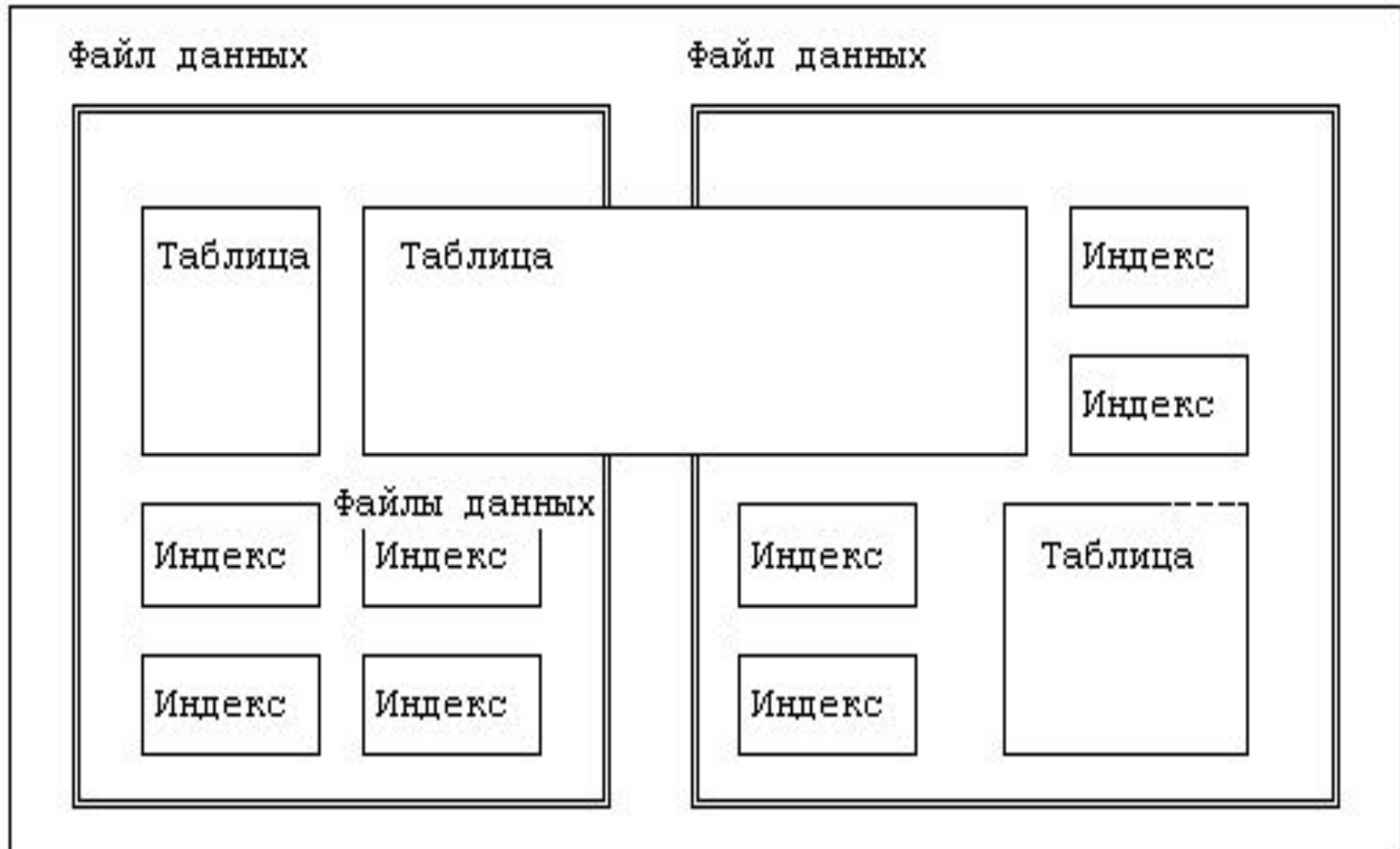
Кластерный индекс

Некластерный индекс



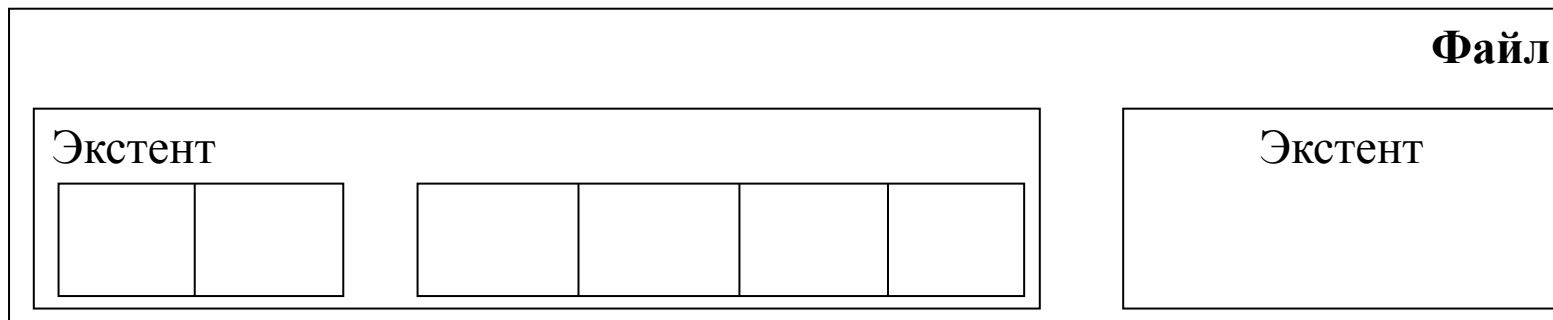
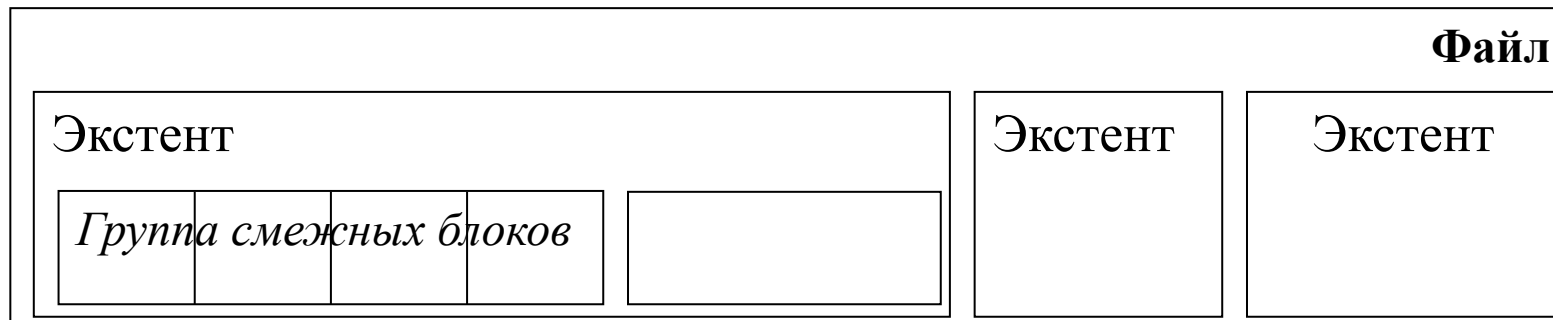
Табличное пространство ORACLE

Табличное пространство (один или несколько файлов данных)



Сегменты, экстенты, файлы и блоки ORACLE

Сегмент



Индексы Oracle

- Таблицы, организованные по индексу
- Индексы кластера
- Индексы с обращенным ключом
- Индексы по убыванию
- Индексы на основе битовых карт
- Индексы по функции
- Прикладные индексы
- Текстовые индексы