

---

# Оглавление

<b>Предисловие к русскому изданию</b> .....	11
<b>Предисловие</b> .....	<b>12</b>
Кому предназначена эта книга .....	13
Зачем мне читать эту книгу? .....	14
Рассматриваемые темы .....	14
Структура книги .....	15
Условные обозначения .....	16
Благодарности .....	17
От издательства .....	17
<b>ЧАСТЬ I. ПОДСИСТЕМА ХРАНЕНИЯ ДАННЫХ</b> .....	<b>18</b>
Сравнение баз данных .....	19
Понимание преимуществ и недостатков .....	22
<b>Глава 1. Введение и обзор</b> .....	24
Архитектура СУБД .....	25
Резидентные и дисковые СУБД .....	27
Колоночные и строчные СУБД .....	30
Файлы данных и индексные файлы .....	34
Буферизация, неизменяемость и упорядочение .....	39
Итоги .....	41
Дополнительная литература .....	42
<b>Глава 2. Введение в B-деревья</b> .....	43
Двоичные деревья поиска .....	43
Дисковые структуры .....	47
Вездесущие B-деревья .....	51
Итоги .....	61
Дополнительная литература .....	61
<b>Глава 3. Форматы файлов</b> .....	62
Актуальность .....	63

---

Двоичное кодирование .....	63
Основные принципы .....	68
Структура страницы .....	69
Слоттированные страницы .....	70
Структура ячеек .....	72
Объединение ячеек в слоттированные страницы .....	73
Управление данными переменного размера .....	75
Управление версиями .....	76
Вычисление контрольной суммы .....	77
Итоги .....	78
Дополнительная литература .....	78
<b>Глава 4. Реализация В-деревьев .....</b>	<b>79</b>
Заголовок страницы .....	79
Двоичный поиск .....	85
Распространение операций разделения и слияния .....	85
Перебалансировка .....	87
Добавление только справа .....	89
Сжатие .....	91
Очистка и обслуживание .....	92
Итоги .....	94
Дополнительная литература .....	95
<b>Глава 5. Обработка транзакций и восстановление .....</b>	<b>96</b>
Организация буферизации данных .....	98
Восстановление .....	106
Управление параллелизмом .....	111
Итоги .....	127
Дополнительная литература .....	127
<b>Глава 6. Варианты В-деревьев .....</b>	<b>129</b>
Копирование при записи .....	129
Абстракции для управления обновлениями .....	131
Ленивые В-деревья .....	132
FD-деревья .....	135
Вw-деревья .....	138
Кэш-независимые В-деревья .....	143

---

Итоги .....	145
Дополнительная литература .....	146
<b>Глава 7. Журналированное хранилище .....</b>	<b>147</b>
LSM-деревья .....	148
Чтение, запись и увеличение пространства .....	162
Подробнее о реализации .....	164
Неупорядоченное LSM-хранилище .....	171
Параллелизм в LSM-деревьях .....	174
Многоуровневое совмещение журналов .....	176
LLAMA и тщательное многоуровневое совмещение .....	180
Итоги .....	182
Дополнительная литература .....	182
<b>ЧАСТЬ I. ЗАКЛЮЧЕНИЕ .....</b>	<b>184</b>
<b>ЧАСТЬ II. РАСПРЕДЕЛЕННЫЕ СИСТЕМЫ .....</b>	<b>186</b>
Основные определения .....	187
<b>Глава 8. Введение и обзор .....</b>	<b>189</b>
Конкурентное выполнение .....	189
Общее состояние в распределенной системе .....	191
Абстракции распределенных систем .....	200
Задача двух генералов .....	206
Невозможность Фишера—Линча—Патерсона .....	207
Синхронность системы .....	209
Модели отказов .....	210
Итоги .....	212
Дополнительная литература .....	213
<b>Глава 9. Обнаружение отказов .....</b>	<b>214</b>
Контрольные пакеты и эхо-запросы .....	215
Детектор отказа с накопленным уровнем подозрительности .....	218
Сплетни и обнаружение отказов .....	219
Обратный взгляд на проблему обнаружения отказов .....	221
Итоги .....	222
Дополнительная литература .....	222

---

<b>Глава 10. Выбор лидера</b> .....	223
Алгоритм забияки .....	224
Аварийное переключение к следующему в очереди .....	226
Оптимизация с кандидатами и обычными узлами .....	227
Алгоритм с приглашениями .....	228
Кольцевой алгоритм .....	229
Итоги .....	230
Дополнительная литература .....	231
<b>Глава 11. Репликация и согласованность</b> .....	232
Обеспечение доступности .....	233
Печально известная теорема CAP .....	233
Общая память .....	236
Упорядочение .....	238
Модели согласованности .....	239
Модели сеансов .....	251
Согласованность в конечном счете .....	252
Настраиваемая согласованность .....	253
Реплики-свидетели .....	255
Строгая согласованность в конечном счете и структуры CRDT .....	256
Итоги .....	259
Дополнительная литература .....	260
<b>Глава 12. Антиэнтропия и распространение</b> .....	261
Исправление при чтении .....	262
Чтение с запросом хэш-суммы .....	264
Передача подсказки .....	264
Деревья Меркла .....	265
Битовая карта векторов версий .....	266
Распространение сплетен .....	268
Итоги .....	273
Дополнительная литература .....	274
<b>Глава 13. Распределенные транзакции</b> .....	275
Обеспечение атомарности операций .....	276
Двухфазная фиксация .....	277

Трехфазная фиксация .....	282
Распределенные транзакции с использованием протокола Calvin .....	284
Распределенные транзакции с использованием протокола Spanner .....	286
Секционирование базы данных .....	289
Распределенные транзакции с использованием библиотеки Percolator .....	290
Исключение координации .....	293
Итоги .....	296
Дополнительная литература .....	297
<b>Глава 14. Консенсус .....</b>	<b>298</b>
Рассылка .....	299
Атомарная рассылка .....	300
Паксос .....	304
Raft .....	320
Византийский консенсус .....	326
Итоги .....	330
Дополнительная литература .....	331
<b>ЧАСТЬ II. ЗАКЛЮЧЕНИЕ .....</b>	<b>333</b>
Дополнительная литература .....	334
<b>Об авторе .....</b>	<b>336</b>
<b>Об обложке .....</b>	<b>336</b>
<b>Приложение А. Библиография .....</b>	<b>www.piter.com</b>