

# Оглавление

<b>Предисловие</b> .....	14
Кому стоит прочесть эту книгу .....	16
Что рассматривается в издании .....	17
Структура книги.....	18
Ссылки и дополнительная литература.....	19
Благодарности .....	19

## **Часть I. Основы информационных систем**

<b>Глава 1.</b> Надежные, масштабируемые и удобные в сопровождении приложения.....	23
1.1. Подходы к работе над информационными системами .....	24
1.2. Надежность .....	27
Аппаратные сбои.....	28
Программные ошибки.....	29
Человеческий фактор.....	30
Насколько важна надежность .....	31
1.3. Масштабируемость .....	32
Описание нагрузки .....	32
Описание производительности .....	36
Как справиться с нагрузкой .....	42
1.4. Удобство сопровождения .....	43
Удобство эксплуатации.....	44
Простота: регулируем сложность .....	45
Возможность развития: облегчаем внесение изменений.....	47
1.5. Резюме.....	47
1.6. Библиография .....	48

<b>Глава 2. Модели данных и языки запросов</b> .....	53
2.1. Реляционная модель в сравнении с документоориентированной моделью .....	54
Рождение NoSQL .....	55
Объектно-реляционное несоответствие.....	56
Связи «многие-к-одному» и «многие-ко-многим» .....	59
Повторяется ли история в случае документоориентированных баз данных.....	62
Реляционные и документоориентированные базы данных сегодня .....	65
2.2. Языки запросов для данных .....	70
Декларативные запросы в Интернете .....	71
Выполнение запросов с помощью MapReduce .....	73
2.3. Графоподобные модели данных .....	76
Графы свойств .....	78
Язык запросов Cypher.....	79
Графовые запросы в SQL.....	81
Хранилища тройных кортежей и SPARQL .....	83
Фундамент: Datalog .....	88
2.4. Резюме .....	90
2.5. Библиография .....	92
<b>Глава 3. Подсистемы хранения и извлечение данных</b> .....	97
3.1. Базовые структуры данных БД .....	98
Хеш-индексы .....	100
SS-таблицы и LSM-деревья .....	104
B-деревья.....	108
Сравнение B- и LSM-деревьев.....	113
Другие индексные структуры.....	115
3.2. Обработка транзакций или аналитика? .....	120
Складирование данных.....	122
«Звезды» и «снежинки»: схемы для аналитики .....	125
3.3. Столбцовое хранилище .....	127
Сжатие столбцов .....	129
Порядок сортировки в столбцовом хранилище .....	131
Запись в столбцовое хранилище.....	132
Агрегирование: кубы данных и материализованные представления .....	133
3.4. Резюме .....	135
3.5. Библиография .....	136

<b>Глава 4.</b> Кодирование и эволюция.....	143
4.1. Форматы кодирования данных .....	144
Форматы, ориентированные на конкретные языки .....	145
JSON, XML и двоичные типы данных.....	146
Thrift и Protocol Buffers .....	150
Avro.....	154
Достоинства схем .....	160
4.2. Режимы движения данных.....	161
Поток данных через БД .....	162
Поток данных через сервисы: REST и RPC .....	164
Поток данных передачи сообщений.....	170
4.3. Резюме.....	173
4.4. Библиография .....	175

## **Часть II. Распределенные данные**

<b>Глава 5.</b> Репликация .....	185
5.1. Ведущие и ведомые узлы .....	186
Синхронная и асинхронная репликация.....	188
Создание новых ведомых узлов.....	190
Перебои в обслуживании узлов .....	190
Реализация журналов репликации .....	193
5.2. Проблемы задержки репликации .....	196
Читаем свои же записи.....	197
Монотонные чтения.....	200
Согласованное префиксное чтение.....	201
Решения проблемы задержки репликации .....	202
5.3. Репликация с несколькими ведущими узлами .....	203
Сценарии использования репликации с несколькими ведущими узлами .....	204
Обработка конфликтов записи .....	207
Топологии репликации с несколькими ведущими узлами.....	212
5.4. Репликация без ведущего узла .....	214
Запись в базу данных при отказе одного из узлов .....	215
Ограничения согласованности по кворуму.....	218
Нестрогие кворумы и направленная передача .....	221
Обнаружение конкурентных операций записи .....	222
5.5. Резюме.....	230
5.6. Библиография .....	232

---

<b>Глава 6. Секционирование</b> .....	239
6.1. Секционирование и репликация .....	240
6.2. Секционирование данных типа «ключ — значение» .....	241
Секционирование по диапазонам значений ключа .....	242
Секционирование по хешу ключа .....	243
Асимметричные нагрузки и разгрузка горячих точек .....	245
6.3. Секционирование и вторичные индексы .....	246
Секционирование вторичных индексов по документам .....	247
Секционирование вторичных индексов по термам .....	248
6.4. Перебалансировка секций .....	250
Методики перебалансировки .....	250
Эксплуатация: автоматическая или ручная перебалансировка .....	254
6.5. Маршрутизация запросов .....	255
6.6. Резюме .....	259
6.7. Библиография .....	260
<b>Глава 7. Транзакции</b> .....	265
7.1. Неустоявшаяся концепция транзакции .....	267
Смысл аббревиатуры ACID .....	267
Однообъектные и многообъектные операции .....	272
7.2. Слабые уровни изоляции .....	277
Чтение зафиксированных данных .....	279
Изоляция снимков состояния и воспроизводимое чтение .....	282
Предотвращение потери обновлений .....	288
Асимметрия записи и фантомы .....	292
7.3. Сериализуемость .....	298
Последовательное выполнение .....	299
Двухфазная блокировка (2PL) .....	304
Сериализуемая изоляция снимков состояния (SSI) .....	308
7.4. Резюме .....	314
7.5. Библиография .....	316
<b>Глава 8. Проблемы распределенных систем</b> .....	323
8.1. Сбои и частичные отказы .....	324
8.2. Ненадежные сети .....	328
Сетевые сбои на практике .....	329
Обнаружение сбоев .....	331
Время ожидания и неограниченные задержки .....	332
Асинхронные и синхронные сети .....	335

8.3. Ненадежные часы .....	338
Монотонные часы и часы истинного времени .....	339
Синхронизация часов и их точность .....	341
Ненадежность синхронизированных часов .....	343
Паузы при выполнении процессов .....	348
8.4. Знание, истина и ложь .....	353
Истина определяется большинством .....	353
Византийские сбои .....	357
Модели системы на практике .....	360
8.5. Резюме.....	364
8.6. Библиография .....	366
<b>Глава 9. Согласованность и консенсус.....</b>	<b>375</b>
9.1. Гарантии согласованности.....	376
9.2. Линеаризуемость.....	378
Что делает систему линеаризуемой .....	379
Опора на линеаризуемость.....	384
Реализация линеаризуемых систем .....	387
Цена линеаризуемости .....	390
9.3. Гарантии упорядоченности.....	394
Порядок и причинность .....	395
Упорядоченность по порядковым номерам .....	399
Рассылка общей последовательности .....	404
9.4. Распределенные транзакции и консенсус .....	409
Атомарная и двухфазная фиксация (2PC) .....	411
Распределенные транзакции на практике.....	417
Отказоустойчивый консенсус.....	422
Сервисы членства и координации.....	428
9.5. Резюме.....	432
9.6. Библиография .....	435
<b>Часть III. Производные данные</b>	
III.1. Системы записи и производные данные .....	446
III.2. Обзор глав .....	447
<b>Глава 10. Пакетная обработка .....</b>	<b>449</b>
10.1. Пакетная обработка средствами Unix.....	451
Простой анализ журнала .....	452
Философия Unix .....	454

---

10.2. MapReduce и распределенные файловые системы .....	458
Выполнение задач в MapReduce .....	460
Объединение и группировка на этапе сжатия.....	464
Объединения на этапе сопоставления .....	470
Выходные данные пакетных потоков .....	473
Сравнение Hadoop и распределенных баз данных .....	478
10.3. За пределами MapReduce.....	482
Материализация промежуточного состояния .....	483
Графы и итеративная обработка .....	488
API и языки высокого уровня.....	491
10.4. Резюме .....	494
10.5. Библиография .....	496
<b>Глава 11. Поточковая обработка.....</b>	<b>505</b>
11.1. Передача потоков событий .....	506
Системы обмена сообщениями .....	507
Секционирование журналов .....	513
11.2. Базы данных и потоки .....	519
Синхронизация систем .....	519
Перехват изменений данных .....	521
Источники событий .....	525
Состояние, потоки и неизменяемость .....	528
11.3. Обработка потоков .....	533
Применение обработки потоков .....	534
Рассуждения о времени .....	538
Объединения потоков.....	543
Отказоустойчивость .....	547
11.4. Резюме .....	551
11.5. Библиография .....	553
<b>Глава 12. Будущее информационных систем .....</b>	<b>563</b>
12.1. Интеграция данных .....	564
Объединение специализированных инструментов путем сбора информации.....	565
Пакетная и потоковая обработка.....	569
12.2. Отделение от баз данных .....	574
Объединение технологий хранения данных.....	575
Проектирование приложений на основе потока данных .....	580
Наблюдение за производными состояниями .....	586

12.3. Стремление к корректности .....	594
Сквозные аргументы в базе данных.....	595
Принудительные ограничения .....	600
Своевременность и целостность .....	604
Доверяй, но проверяй .....	609
12.4. Делать что должно .....	614
Предсказательная аналитика.....	615
Конфиденциальность и отслеживание .....	619
12.5. Резюме .....	627
12.6. Библиография .....	628