

С. Л. Бабичев, К. А. Коньков

РАСПРЕДЕЛЕННЫЕ СИСТЕМЫ

УЧЕБНОЕ ПОСОБИЕ ДЛЯ ВУЗОВ

Рекомендовано Учебно-методическим отделом высшего образования в качестве учебного пособия для студентов высших учебных заведений, обучающихся по ИТ-направлениям

**Книга доступна
на образовательной платформе «Юрайт» urait.ru,
а также в мобильном приложении «Юрайт.Библиотека»**

Москва ■ Юрайт ■ 2020

Оглавление

Предисловие	11
-------------------	----

Часть 1 ТЕОРЕТИЧЕСКАЯ ЧАСТЬ

Глава 1. Вводная глава	15
1.1. Введение. Понятие распределенной системы	15
1.1.1. Определение	15
1.1.2. Особенности распределенных систем	17
1.1.3. Целесообразность построения распределенных систем	17
1.1.4. Примеры и применения	18
1.2. Параллельные и распределенные системы	19
1.3. Архитектурные особенности	24
1.3.1. Сервисы, роли и архитектурные стили	26
1.3.2. Клиент-сервер	26
1.3.3. Одноранговые сети	28
1.3.4. Сервисно-ориентированная архитектура	29
1.4. Дизайн масштабируемых распределенных систем	30
1.4.1. Масштабируемость	31
1.4.2. Особенности проектирования распределенных систем	33
Глава 2. Модели	35
2.1. Введение в моделирование и понятие модели	35
2.2. Модель распределенного исполнения	36
2.2.1. Общее описание	36
2.2.2. Модель коммуникационного канала	36
2.2.3. Событийное описание	37
2.2.4. Упорядочивание событий	37
2.3. Отношение причинного предшествования	40
2.4. Логическое время. Отметки времени Лампорта	41
2.4.1. Реализация логических часов	42
2.4.2. Скалярное время	43
2.4.3. Векторное время	44
2.4.4. Алгоритмы реализации векторных часов	46
2.5. Синхронное и асинхронное исполнение	47
2.5.1. Введение	47
2.5.2. Эмуляции синхронных систем асинхронными и наоборот	49
2.5.3. Эмуляции	52

2.6. Модели отказов	52
2.6.1. Отказы процессов.....	53
2.6.2. Отказы коммуникационных каналов.....	54
2.6.3. Иерархия моделей неисправности.....	54
2.7. Свойства распределенных алгоритмов	55
2.8. Глобальное состояние	57
2.8.1. Распределенная сборка мусора.....	59
2.8.2. Распределенное обнаружение тупиков.....	59
2.8.3. Распределенное обнаружение завершения	59
2.8.4. Фиксация глобального состояния	61
Глава 3. Коммуникационная подсистема	63
3.1. Введение и состав коммуникационной подсистемы	63
3.1.1. Введение	63
3.1.2. Состав коммуникационной подсистемы.....	63
3.2. Сети и сетевые технологии	66
3.2.1. Введение.....	66
3.2.2. Типы сетей	66
3.2.3. Ключевые проблемы использования сетей в распределенных системах	68
3.2.4. Принципы построения сетей.....	70
3.3. Маршрутизация и алгоритмы на графах.....	77
3.3.1. Введение.....	77
3.3.2. Графы	78
3.3.3. Алгоритмы маршрутизации	78
3.4. Межпроцессный обмен	81
3.4.1. Особенности обмена сообщениями	81
3.4.2. Адресация.....	86
3.4.3. Пример системы обмена сообщениями — пакет MPI	87
3.4.4. Широковещательные и многоадресные рассылки	90
3.4.5. IP-multicast	91
3.5. Удаленные вызовы	92
3.5.1. Введение.....	92
3.5.2. Протокол «запрос-ответ» (request-reply)	92
3.5.3. Удаленный вызов процедуры.....	93
3.5.4. Пример удаленного вызова — веб-сервисы.....	98
3.6. Косвенные (indirect) коммуникации	99
3.6.1. Очереди сообщений.....	101
3.6.2. Групповые коммуникации	104
3.7. Координация и согласие в групповых коммуникациях	109
3.7.1. Базовые многоадресные рассылки	110
3.7.2. Надежная многоадресная рассылка	110
3.7.3. Упорядоченные многоадресные рассылки.....	112
3.7.4. Открытые группы и виртуальная синхронность	113
Глава 4. Синхронизация	117
4.1. Введение	117

4.2. Алгоритмы синхронизации часов	119
4.2.1. Алгоритм Кристиана.....	119
4.2.2. Алгоритм Беркли	120
4.2.3. Усредняющие алгоритмы	120
4.3. Алгоритмы выбора.....	120
4.4. Распределенное взаимноеисключение	121
4.5. Консенсус.....	124
4.5.1. Введение.....	124
4.5.2. Модель системы и основные определения.....	125
4.5.3. Согласие в системах с отказами	128
4.5.4. Отсутствие детерминированного решения в асинхронных системах	131
4.5.5. Способы обхода результата FLP impossibility.....	132
4.5.6. Алгоритм PAXOS	134
4.6. Распределенные транзакции	134
4.6.1. Введение.....	134
4.6.2. Свойства ACID.....	135
4.6.3. Типы транзакций	136
4.6.4. Компоненты архитектуры, необходимые для поддержки распределенных транзакций.....	137
4.6.5. Управление параллельным выполнением транзакций	142
4.6.6. Метод временных меток	147
4.6.7. Контроль параллельного исполнения транзакций в некоторых популярных распределенных системах	149
4.6.8. Протоколы распределенного завершения	150
4.6.9. Восстановление после отказов	152
4.6.10. Создание контрольных точек и откат путем восстановления.....	154
Глава 5. Репликация и консистентность	157
5.1. Введение	157
5.2. Модель и архитектура управления реплицированными данными	158
5.3. Пассивная и активная репликации	161
5.3.1. Пассивная репликация.....	161
5.3.2. Активная репликация	163
5.4. Отказоустойчивость сервиса репликации.....	163
5.5. Модели консистентности	164
5.5.1. Модели непротиворечивости, ориентированные на данные.....	165
5.5.2. Модели непротиворечивости, ориентированные на клиента. Потенциальная согласованность	169
5.5.3. Протоколы кворума	172
5.6. Размещение и обновление реплик	172
5.6.1. Размещение реплик	173
5.6.2. Распространение обновлений	173
Глава 6. Безопасность	176
6.1. Введение и модель.....	176

6.1.1. Модель	176
6.1.2. Требования к безопасности	178
6.1.3. Злоумышленники, угрозы, атаки	178
6.1.4. Политика безопасности и механизмы	179
6.2. Безопасность. Криптография	180
6.3. Аутентификация	182
6.4. Авторизация	184
6.4.1. Общие вопросы контроля доступа	184
6.4.2. Модели контроля доступа	185
6.4.3. Дискреционное управление доступом	186
6.4.4. Межсетевые экраны (брандмауэры)	191
6.5. Аудит	192
Глава 7. Системы хранения данных	194
7.1. Введение	194
7.2. Краткий обзор современных подходов к построению систем распределенного хранения данных	194
7.2.1. Программно-определяемые хранилища	194
7.2.2. Механизм хранения данных на уровне объектов	196
7.2.3. Архитектурные особенности	197
7.2.4. Механизм регулируемой избыточности	197
7.2.5. Механизм георепликации	198
7.2.6. Проблемы теоремы CAP	199
7.3. Распределенные кластерные файловые системы	201
7.4. Пиринговые системы	205
7.4.1. Одноранговые (пиринговые) сети. Введение	205
7.4.2. Эволюция пиринговых сетей	206
7.4.3. Масштабируемость P2P-сетей	207
7.4.4. BitTorrent	207
7.4.5. DHT — распределенные хеш-таблицы	208
7.4.6. Безопасность P2P-сетей	210
7.4.7. Некоторые меры защиты P2P-сетей	211

Часть 2 ПРАКТИЧЕСКАЯ ЧАСТЬ

Глава 8. Введение	215
8.1. Организация раздела с практической частью	218
8.2. Что будет изучаться?	219
8.3. Что требуется для прохождения практического курса	220
8.4. Главные цели практической части	221
Глава 9. Архитектура распределенных систем	222
9.1. Моделирование распределенных систем. Процессы и распределенные процессы. Сообщения. Сеть. Связи. Топология	222
Вопросы к теме	224

9.2. Моделирование распределенных процессов	224
9.2.1. Модель асинхронной распределенной системы и ее симулятор	225
9.2.2. Рабочие функции модели	228
9.2.3. Архитектура и состав симулятора	235
Задание к теме	238
9.3. Сериализация	238
Задания к теме	243
9.4. Моделирование распределенных процессов — продолжение	243
9.4.1. Контекст процесса, контекст алгоритма	243
Вопросы и задания к теме	249
9.5. Взаимодействие распределенных процессов	249
9.5.1. Именованное взаимодействие	251
9.5.2. Связь между реальным процессом и модельным	252
Вопросы к теме	253
9.6. Архитектура клиент-сервер. Роли процессов	253
9.6.1. Ролевые игры в распределенных системах	254
9.6.2. Синхронные и асинхронные запросы	263
9.6.3. Сессионное взаимодействие распределенных процессов	266
Задание к теме	267
9.6.4. Обслуживание сервером клиента. Контекст клиента на сервере	267
Вопросы к теме	271
9.7. Механизмы взаимодействия: RPC, клиент, сервер	271
9.7.1. Определение вызываемой процедуры по идентификатору	275
9.7.2. Клиенты и клиентские вызовы серверных процедур	277
9.7.3. Структура исполнителя	278
9.7.4. Пример контекстного взаимодействия. Remote Mutex	280
9.7.5. Работа с серверными ресурсами	284
9.7.6. Передача аргументов	286
9.7.7. Объекты как аргументы	288
9.7.8. Фабрика объектов	293
Задания к теме	296
Глава 10. Алгоритмы. Практика	297
10.1. Физическое время	297
10.1.1. Синхронизация времени, синхронизация данных, временные метки, приоритеты и роли узлов	297
10.1.2. Алгоритм Кристиана	297
10.1.3. Алгоритм Кристиана с использованием синхронного взаимодействия процессов	306
10.1.4. Алгоритм Кристиана с использованием RPC	308
Вопросы и задания к теме	309
10.2. Алгоритм Беркли	309
Вопросы к теме	311
10.3. Логическое время	311

10.3.1. Отношение причинного предшествования (happens before) Лампорта.....	311
10.3.2. Использование логического времени	313
10.3.3. Скалярные и векторные часы	314
10.4. Сохранность и синхронность	317
10.4.1. Синхронные и асинхронные примитивы обмена сообщениями	317
10.4.2. Буферизируемые и небуферизируемые примитивы обмена сообщениями	318
Вопросы к теме	320
10.4.3. Синхронное и асинхронное исполнение.....	320
10.4.4. Синхронизаторы	322
Вопросы и задания к теме	325
10.4.5. Основные свойства алгоритмов	326
10.5. Сети. Топология физических сетей.....	326
10.5.1. Топология сетей	326
10.5.2. Транспортный уровень	332
Задание к теме.....	335
10.6. Маршрутизация. Алгоритмы на графах	335
Задание к теме.....	337
10.6.1. Сети с полной информацией о связях	337
10.6.2. Алгоритм Дейкстры	339
Задание к теме.....	343
10.6.3. Сети с динамически изменяемыми связями. Алгоритм остовного дерева	343
10.6.4. Построение связующих деревьев	345
Вопросы к теме	350
10.6.5. Алгоритм заливки (FLOOD).....	350
Задания к теме	353
10.6.6. Физическая и логическая маршрутизация. Наложение логической сети на физическую	353
10.6.7. Алгоритм передачи сообщения по маршруту	356
Вопросы и задания к теме	359
10.6.8. Построение таблиц маршрутизации: алгоритм Флойда — Уоршалла.....	359
Вопросы и задания к теме	362
10.6.9. Алгоритм Туэга	362
Вопросы и задания по теме	367
10.6.10. Алгоритм Чанди — Мисры.....	367
Вопросы и задания к теме	369
10.6.11. Алгоритм NETCHANGE.....	370
Вопросы и задания к теме	375
10.6.12. Операции broadcast и convergecast	375
Вопросы и задания к теме	377
10.6.13. Сочетания алгоритмов заливки и CONVERGECAST.....	377
Вопросы и задания к теме	380
10.6.14. Именованное. Служба именования. Поиск	380

Задание к теме	381
10.6.15. Поиск в ширину	381
Задания к теме	384
10.6.16. Причинно-упорядоченная рассылка	385
10.6.17. Полностью упорядоченная рассылка	385
10.7. Алгоритмы голосования	386
10.7.1. Алгоритм забияки	387
Задания к теме	395
10.7.2. Кольцевой алгоритм Ле-Ланна — Чанга — Робертса	396
Вопросы и задания к теме	397
10.7.3. Голосование в сетях с произвольной топологией	397
Вопросы и задания к теме	400
10.8. Глобальное состояние	401
10.8.1. Алгоритмы обнаружения завершения	401
Задания к теме	407
10.9. Синхронизация. Алгоритмы взаимного исключения	407
10.9.1. Централизованный алгоритм	407
Вопросы и задания к теме	408
10.9.2. Распределенный алгоритм	409
Задание к теме	409
10.10. Алгоритмы консенсуса. Алгоритм PAXOS	409
Вопросы и задания к теме	416
10.11. Распределенные транзакции	417
Вопросы к теме	418
Глава 11. Безопасность	419
11.1. Виды атак	419
Вопросы к теме	421
11.2. Принцип КААА	421
11.3. Криптография	422
11.3.1. Симметричные алгоритмы	422
11.3.2. Несимметричные алгоритмы	427
11.3.3. Хеширование	431
11.3.4. Криптографически стойкий генератор случайных чисел	435
11.4. Аутентификация	438
11.4.1. Алгоритмы взаимной аутентификации	438
11.4.2. Взаимная аутентификация с помощью центра аутентификации	442
11.4.3. Алгоритм Диффи — Хеллмана	444
Вопросы и задания к теме	447
11.4.4. Цифровая подпись сообщений	447
11.5. Аудит	449
Вопросы и задания к теме	450
11.6. Небольшая практическая задача	450
Вопросы и задания к теме	454
Глава 12. Распределенное хранение	455

12.1. Репликация	455
12.2. Синхронизация больших объектов	455
12.2.1. Надежны ли современные вычислительные системы?	456
12.2.2. Синхронизация больших объектов. Алгоритм	457
12.2.3. Синхронизация больших объектов: более продвинутый алгоритм	459
Задания к теме	460
12.3. Избыточное хранение	460
Задания к теме	464
12.4. Дедупликация	465
12.4.1. Дедупликация: реализация	467
12.4.2. Технические вопросы реализации дедуплицированного хранилища	467
Глава 13. Распределенные вычисления	470
13.1. Немного истории	470
13.2. Влияние аппаратной архитектуры сети на производительность	471
13.2.1. Gigabit Ethernet	472
13.2.2. Myrinet	472
13.2.3. InfiniBand	473
13.3. Влияние решаемой задачи на производительность	473
13.4. MPI	475
Вопросы к теме	478
13.5. Распараллеливание задач по вычислительным ресурсам. Сильно связанные задачи и слабо связанные задачи	478
13.6. Сравнение моделей распределенных вычислений. Анализ достоинств и недостатков моделей	480
13.7. Структура организации метакомпьютинга	480
13.8. Общая структура функционирования	484
13.9. Проблемный и системный компоненты метасистемы	484
Глава 14. Проектирование метасистемы	486
14.1. Сервер: декомпозиция задачи	487
14.2. Сервер: структура множеств	488
14.3. Особенности реализации	490
14.3.1. Выбор целевых платформ и языков программирования	490
14.3.2. Слой абстракции от вычислительной системы	491
14.4. Распределитель ресурсов	494
14.5. RPC API со стороны метаклиента	494
14.6. API метасервера (MetaServerAPI)	496
14.7. Проблемный компонент на сервере: пример решения задачи оптимизации методом полного перебора значений	498
14.8. Проблемный компонент на клиенте	502
Задание к теме	504
Литература	505
Новые издания по дисциплине	507