

Software Dynamics

оптимизация производительности
программного обеспечения

Ричард Л. Сайтс



Санкт-Петербург · Москва · Минск

2024

КРАТКОЕ СОДЕРЖАНИЕ

Предисловие.....	20
Вступление	21
Благодарности.....	26
Об авторе.....	27
От издательства.....	28

Часть I. Измерение

Глава 1. Моя программа слишком медленная	31
Глава 2. Измерение производительности процессора	45
Глава 3. Измерение скорости доступа к памяти.....	64
Глава 4. Взаимодействие процессора и памяти.....	87
Глава 5. Измерение скорости диска/SSD.....	100
Глава 6. Измерение скорости сетей	130
Глава 7. Взаимодействие с дисковой и сетевой базами данных.....	162

Часть II. Наблюдение

Глава 8. Логирование	187
Глава 9. Совокупные показатели	195
Глава 10. Информационные панели	212
Глава 11. Другие инструменты.....	225
Глава 12. Трассировщики.....	254
Глава 13. Принципы проектирования инструментов наблюдения	273

Часть III. Трассировщик для режимов ядра и пользователя

Глава 14. KUtrace: цели, проектирование, реализация.....	285
Глава 15. KUtrace: патчи ядра Linux.....	294
Глава 16. KUtrace: загружаемый модуль Linux.....	307
Глава 17. KUtrace: управление средой режима пользователя.....	315
Глава 18. KUtrace: постобработка.....	319
Глава 19. KUtrace: показ динамики ПО.....	329

Часть IV. Анализ

Глава 20. Что нужно искать.....	345
Глава 21. Слишком много кода.....	348
Глава 22. Медленное выполнение.....	357
Глава 23. Ожидание процессора.....	368
Глава 24. Задержки памяти.....	379
Глава 25. Ожидание диска.....	388
Глава 26. Ожидание сети.....	403
Глава 27. Ожидание блокировок.....	424
Глава 28. Ожидание таймера.....	447
Глава 29. Ожидание очереди.....	451
Глава 30. Итоги.....	477
Приложение А. Тестовые серверы.....	482
Приложение Б. Записи трассировки.....	485
Глоссарий.....	491
Ссылки.....	500
Источники фотографий.....	510

ОГЛАВЛЕНИЕ

Предисловие.....	20
Вступление.....	21
Исходный код.....	25
Полноцветные иллюстрации.....	25
Благодарности.....	26
Об авторе.....	27
От издательства.....	28

Часть I. Измерение

Глава 1. Моя программа слишком медленная.....	31
1.1. Контекст дата-центра.....	31
1.2. Оборудование дата-центра.....	34
1.3. Программное обеспечение дата-центра.....	35
1.4. Длинный хвост задержки.....	37
1.5. Схема мышления.....	39
1.6. Оценка порядка величины.....	40
1.7. Почему транзакции медленно выполняются.....	41
1.8. Пять фундаментальных ресурсов.....	43
1.9. Резюме.....	43
Глава 2. Измерение производительности процессора.....	45
2.1. Как мы к этому пришли.....	46
2.2. Где мы сейчас.....	50

2.3. Измерение задержки инструкции add	51
2.4. Сбой линейного участка программы.....	53
2.5. Простой цикл, проблемы из-за накладных издержек цикла и оптимизирующего компилятора.....	54
2.6. Проблемы из-за неиспользуемой переменной.....	57
2.7. Улучшение цикла.....	57
2.8. Зависимые переменные	59
2.9. Фактическая задержка выполнения.....	59
2.10. Другие нюансы.....	60
2.11. Резюме	61
Упражнения.....	62
Глава 3. Измерение скорости доступа к памяти.....	64
3.1. Тайминг памяти.....	64
3.2. О памяти.....	65
3.3. Организация кэша	68
3.4. Выравнивание данных.....	71
3.5. Организация буфера ассоциативной трансляции	72
3.6. Измерения	73
3.7. Измерение размера кэш-линии	74
3.8. Задача: предвыборка $N + 1$	76
3.9. Зависимые загрузки	77
3.10. Непроизвольная динамическая память с произвольным доступом	78
3.11. Измерение общего размера каждого уровня кэша	80
3.12. Измерение ассоциативности кэша каждого уровня	83
3.13. Временные издержки буфера трансляции	83
3.14. Недоиспользование кэша.....	84
3.15. Резюме	84
Упражнения.....	85
Глава 4. Взаимодействие процессора и памяти.....	87
4.1. Взаимодействие с кэшем	88
4.2. Динамика простого матричного умножения.....	89
4.3. Приблизительные оценки	90

4.4. Инициализация, перекрестная проверка и наблюдение	91
4.5. Первые результаты	92
4.6. Ускоренное матричное умножение, метод транспонирования.....	95
4.7. Ускоренное матричное умножение, метод подблоков	97
4.8. Вычисления с учетом кэша.....	98
4.9. Резюме.....	99
Упражнение	99
Глава 5. Измерение скорости диска/SSD	100
5.1. О жестких дисках	101
5.2. Про SSD.....	103
5.3. Программный доступ к диску и буферизация	106
5.4. Анализ скорости чтения диска	109
5.5. Приблизительный расчет	113
5.6. Анализ скорости записи на диск.....	115
5.7. Результаты.....	115
5.8. Считывание с диска.....	116
5.9. Запись на диск.....	120
5.10. Считывание с SSD.....	124
5.11. Запись на SSD.....	125
5.12. Множество передач данных	126
5.13. Резюме	127
Упражнения.....	128
Глава 6. Измерение скорости сетей	130
6.1. Об Ethernet.....	132
6.2. О хабах, коммутаторах и маршрутизаторах.....	135
6.3. О TCP/IP.....	136
6.4. О пакетах.....	137
6.5. Об удаленных вызовах процедур (RPC)	138
6.6. Слоп	140
6.7. Анализ сетевого трафика.....	142
6.8. Пример определения RPC сообщения	144
6.9. Пример структуры логирования.....	148

6.10. Пример клиент-серверной системы, использующей RPC	150
6.11. Пример серверной программы.....	150
6.12. Спин-блокировка.....	151
6.13. Пример клиентской программы.....	153
6.14. Измерение одного RPC от клиента к серверу	155
6.15. Постобработка логов RPC.....	157
6.16. Наблюдения.....	157
6.17. Резюме	159
Упражнения.....	160
Глава 7. Взаимодействие с дисковой и сетевой базами данных	162
7.1. Синхронизация времени	162
7.2. Несколько клиентов.....	169
7.3. Спинлоки	170
7.4. Эксперимент 1.....	170
7.5. База данных на диске.....	174
7.6. Эксперимент 2.....	174
7.7. Эксперимент 3.....	179
7.8. Логирование.....	182
7.9. Прояснение вариативности в задержке транзакций.....	183
7.10. Резюме	183
Упражнения.....	184

Часть II. Наблюдение

Глава 8. Логирование.....	187
8.1. Инструменты наблюдения.....	187
8.2. Логирование.....	188
8.3. Базовое логирование.....	188
8.4. Расширенное логирование.....	189
8.5. Временные метки	190
8.6. RPC ID.....	191
8.7. Форматы журналов	193
8.8. Управление журналами	194
8.9. Резюме.....	194

Глава 9. Совокупные показатели	195
9.1. Однородная и пиковая частота событий	196
9.2. Интервалы измерений.....	197
9.3. Временные линии	198
9.4. Дальнейшее обобщение временных линий.....	199
9.5. Временные шкалы гистограмм.....	202
9.6. Агрегирование измерений отдельных событий.....	204
9.7. Паттерны изменения значений во времени.....	206
9.8. Интервалы обновления.....	207
9.9. Образцы транзакций.....	209
9.10. Резюме	211
Глава 10. Информационные панели	212
10.1. Образец сервиса	212
10.2. Образцы панелей мониторинга	214
10.3. Основная панель мониторинга	215
10.4. Панели мониторинга отдельных экземпляров.....	220
10.5. Панели мониторинга серверов.....	221
10.6. Проверка исправности	222
10.7. Резюме	223
Упражнение	224
Глава 11. Другие инструменты	225
11.1. Виды инструментов наблюдения	225
11.2. Данные для наблюдения	228
11.3. Команда top.....	229
11.4. Файловые системы /proc и /sys.....	230
11.5. Команда time.....	231
11.6. Команда perf.....	231
11.7. Программа orfprofile – профилировщик процессора	233
11.8. Команда strace – системные вызовы	237
11.9. Команда ltrace – трассировка библиотечных вызовов	240
11.10. Фреймворк ftrace – трассировка процессора.....	240
11.11. Программа mtrace – выделение/освобождение памяти	244
11.12. Команда blktrace – трассировка диска.....	246

11.13. Команда tcpdump и Wireshark, трассировка сети	249
11.14. Инструмент locktrace, блокировки критических разделов.....	251
11.15. Предлагаемая нагрузка, исходящие вызовы и задержка транзакций ..	251
11.16. Резюме	253
Упражнения.....	253
Глава 12. Трассировщики	254
12.1. Преимущества трассировки.....	254
12.2. Недостатки трассировки	255
12.3. Три основных вопроса	256
12.4. Пример: первый трассировщик счетчика команд	260
12.5. Пример: количество вызовов и время выполнения функции.....	262
12.6. Практическая задача: трассировка функций Gmail по отдельности	266
12.7. Резюме	272
Глава 13. Принципы проектирования инструментов наблюдения.....	273
13.1. Что наблюдать.....	273
13.2. Как часто и как долго.....	274
13.3. Каковы допустимые издержки.....	275
13.4. Последствия проектирования	277
13.5. Практическая задача: бакеты гистограмм	277
13.6. Проектирование отображения данных	280
13.7. Резюме	281
Часть III. Трассировщик для режимов ядра и пользователя	
Глава 14. KUtrace: цели, проектирование, реализация.....	285
14.1. Обзор.....	285
14.2. Цели.....	286
14.3. Проектирование	287
14.4. Реализация	290
14.5. Патчи и модуль ядра.....	291
14.6. Управляющая программа.....	292
14.7. Постобработка	292
14.8. Заметка о безопасности	293
14.9. Резюме	293

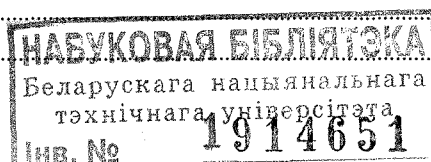
Глава 15. KUtrace: патчи ядра Linux	294
15.1. Структуры данных буфера трассировки	295
15.2. Сырой формат блока трейса	296
15.3. Записи трейсов	297
15.4. Записи трейса для IPC	299
15.5. Временные метки	300
15.6. Номера событий	301
15.7. Вложенные записи трейсов	301
15.8. Код	302
15.9. Трассировка пакетов	303
15.10. Патчи AMD/Intel x86-64	304
15.11. Резюме	306
Упражнение	306
Глава 16. KUtrace: загружаемый модуль Linux	307
16.1. Структуры данных интерфейса ядра	307
16.2. Загрузка/выгрузка модуля	308
16.3. Инициализация и управление трассировкой	309
16.4. Реализация вызовов трассировки	309
16.5. Insert1	310
16.6. InsertN	313
16.7. Переключение на новый блок трейса	313
16.8. Резюме	314
Глава 17. KUtrace: управление средой режима пользователя	315
17.1. Управление трассировкой	315
17.2. Отдельная программа kutraces_control	316
17.3. Внутренняя библиотека kutraces_lib	317
17.4. Интерфейс управления для загружаемого модуля	317
17.5. Резюме	318
Глава 18. KUtrace: постобработка	319
18.1. Детали постобработки	319
18.2. Программа rawtoevent	320
18.3. Программа eventtospan	322
18.4. Программа spantotrim	324

18.5. Программа <code>spantospan</code>	325
18.6. Программы <code>samptoname_k</code> и <code>samptoname_u</code>	325
18.7. Программа <code>makeself</code>	326
18.8. Формат JSON в KUTrace.....	326
Метаданные JSON.....	326
События JSON.....	327
18.9. Резюме.....	328
Глава 19. KUTrace: показ динамики ПО.....	329
19.1. Обзор.....	329
19.2. Область 1, управление.....	330
19.3. Область 2, ось Y.....	332
19.4. Область 3, временные линии.....	333
19.5. Область 4, легенда IPC.....	340
19.6. Область 5, ось X.....	340
19.7. Область 6, сохранение/восстановление.....	341
19.8. Вспомогательное управление.....	341
19.9. Резюме.....	342
Часть IV. Анализ	
Глава 20. Что нужно искать.....	345
20.1. Обзор.....	345
Глава 21. Слишком много кода.....	348
21.1. Обзор.....	348
21.2. Программа.....	349
21.3. Загадка.....	349
21.4. Изучение и анализ.....	350
21.5. Загадка разгадана.....	355
21.6. Резюме.....	356
Глава 22. Медленное выполнение.....	357
22.1. Обзор.....	357
22.2. Программа.....	358
22.3. Загадка.....	358
22.4. Программа-антагонист, работающая с плавающей запятой.....	361

22.5. Программа-антагонист, активно потребляющая память	364
22.6. Загадка разгадана	365
22.7. Резюме	366
Глава 23. Ожидание процессора	368
23.1. Программа	368
23.2. Загадка	369
23.3. Изучение и анализ	369
23.4. Загадка 2	371
23.5. Загадка 2 разгадана	373
23.6. Дополнительная загадка	376
23.7. Резюме	378
Упражнения	378
Глава 24. Задержки памяти	379
24.1. Программа	379
24.2. Загадка	380
24.3. Изучение и анализ	380
24.4. Загадка 2: доступ к таблице страниц	385
24.5. Загадка 2 разгадана	386
24.6. Резюме	386
Упражнения	387
Глава 25. Ожидание диска	388
25.1. Программа	388
25.2. Загадка	389
25.3. Изучение и анализ	389
25.4. Считывание 40 Мбайт	392
25.5. Считывание последовательных блоков по 4 Кбайт	393
25.6. Считывание произвольных блоков по 4 Кбайт	395
25.7. Запись и синхронизация 40 Мбайт на SSD	398
25.8. Считывание 40 Мбайт с SSD	398
25.9. Одновременное обращение двух программ к двум файлам	399
25.10. Загадки разгаданы	401
25.11. Резюме	401
Упражнение	402

Глава 26. Ожидание сети.....	403
26.1. Обзор.....	404
26.2. Программы.....	405
26.3. Эксперимент 1.....	405
26.4. Загадка эксперимента 1.....	406
26.5. Эксперимент 1. Анализ и рассуждение.....	409
26.6. Эксперимент 1. А как насчет времени между RPC?.....	413
26.7. Эксперимент 2.....	415
26.8. Эксперимент 3.....	415
26.9. Эксперимент 4.....	417
26.10. Загадки разгаданы.....	419
26.11. Дополнительная аномалия.....	420
26.12. Резюме.....	423
Глава 27. Ожидание блокировок.....	424
27.1. Обзор.....	424
27.2. Программа.....	429
27.3. Эксперимент 1. Длительное удержание блокировки.....	433
Простая блокировка.....	433
Перегрузка блокировки.....	434
27.4. Загадки эксперимента 1.....	435
27.5. Изучение и анализ эксперимента 1.....	436
Перехват блокировки.....	437
«Голодание» потока.....	438
27.6. Эксперимент 2. Устранение перехвата блокировки.....	438
27.7. Эксперимент 3. Уменьшение соперничества за счет нескольких блокировок.....	440
27.8. Эксперимент 4. Уменьшение соперничества за счет сокращения работы во время блокировки.....	442
27.9. Эксперимент 5. Уменьшение соперничества с помощью механизма RCU для информационной панели.....	444
27.10. Резюме.....	446
Глава 28. Ожидание таймера.....	447
28.1. Периодическая работа.....	447
28.2. Тайм-ауты.....	448

28.3. Квантование времени выполнения	449
28.4. Встроенные задержки выполнения	449
28.5. Резюме	450
Глава 29. Ожидание очереди	451
29.1. Обзор	451
29.2. Распределение запросов	454
29.3. Структура очереди	455
29.4. Рабочие задачи	456
29.5. Основная задача	456
29.6. Операция Dequeue	457
29.7. Операция Enqueue	457
29.8. Класс Spinlock	457
29.9. Процедура Work	458
29.10. Простые примеры	458
29.11. Что могло пойти не так?	460
29.12. Частота процессора	461
29.13. Сложные примеры	462
29.14. Ожидание ядер процессора: журнал RPC	463
29.15. Ожидание ядер процессора: Kutrace	464
29.16. Поток PlainSpinLock	467
29.17. Первопричина	469
29.18. PlainSpinLock исправлен: наблюдаемость	470
29.19. Балансировка нагрузки	471
29.20. Глубина очереди: наблюдение	472
29.21. Холостое выполнение в конце	472
29.22. Еще один изъясн	473
29.23. Перепроверка	473
29.24. Резюме	474
Упражнения	475
Глава 30. Итоги	477
30.1. Чему вы научились	477
30.2. О чем мы не говорили	479
30.3. Дальнейшие шаги	480
30.4. Резюме (всей книги)	481



Приложение А. Тестовые серверы.....	482
А.1. Оборудование тестовых серверов.....	482
А.2. Подключение серверов.....	484
Приложение Б. Записи трассировки.....	485
Б.1. Записи трейса фиксированной длины.....	486
Б.2. Записи трейсов переменной длины.....	486
Б.3. Номера событий.....	488
События, вставляемые патчами KUttrace режима ядра.....	488
События, вставляемые кодом режима пользователя.....	489
События, вставляемые кодом постобработки.....	490
Глоссарий.....	491
Ссылки.....	500
Источники фотографий.....	510