

Оглавление

Предисловие Марка Берджеса	19
Предисловие авторов	22
Условные обозначения	27
Использование примеров кода	27
Благодарности	29

Часть I. Введение

Глава 1. Вступление	35
Подход системного администратора к управлению сервисами	35
Подход Google к управлению сервисами: техника обеспечения надежности сайтов	37
Принципы SRE	40
Уделяем особое внимание инженерным задачам	40
Добиваемся максимальной скорости внедрения изменений без потери качества обслуживания	41
Мониторинг	42
Реагирование на критические ситуации	43
Управление изменениями	44
Прогнозирование нагрузки и планирование производительности	44
Материально-техническое обеспечение	45
Эффективность и производительность	45
Конец начала	46
Глава 2. Среда промышленной эксплуатации Google с точки зрения SRE	47
Оборудование	47
Системное ПО, которое «организует» оборудование	49
Управление машинами	49
Хранилище	51
Сеть	53

Другое системное ПО	54
Сервис блокировок.....	54
Мониторинг и оповещение.....	54
Наша инфраструктура ПО.....	55
Наша среда разработки.....	55
Shakespeare: пример сервиса.....	56
Жизненный цикл запроса	57
Организация задач и данных	59

Часть II. Принципы

Глава 3. Приручаем риски	64
Управление рисками	64
Измерение рисков, связанных с сервисом.....	65
Рискоустойчивость сервисов	67
Определение рискоустойчивости пользовательских сервисов.....	68
Определение рискоустойчивости инфраструктурных сервисов	71
Обоснование критерия суммарного уровня ошибок (бюджета ошибок).....	73
Формируем бюджет ошибок.....	74
Преимущества.....	75
Глава 4. Целевой уровень качества обслуживания	77
Терминология для уровня качества обслуживания.....	78
Показатели	78
Целевые показатели	79
Соглашения.....	80
Показатели на практике.....	81
Что важно для вас и ваших пользователей	81
Сбор показателей.....	82
Агрегирование.....	83
Стандартизация показателей.....	84
Целевые показатели на практике.....	85
Определение целей.....	85
Выбор целевых показателей	86
SLO как инструмент управления	87
SLO формируют ожидания	88
Соглашения на практике	89
Глава 5. Избавляемся от рутины	90
Что такое рутина.....	90
Почему лучше иметь меньше рутинной работы.....	92
Что такое инженерная работа	93
Всегда ли рутина вредна	94
Итоги главы.....	95

Глава 6. Мониторинг распределенных систем	96
Определения	96
Зачем нужен мониторинг	97
Ставим для мониторинга реальные задачи	98
Симптомы и причины	100
Методы черного и белого ящика	100
Четыре золотых сигнала	101
Позаботимся о своем «хвосте» (или производительность измеряемая и наблюдаемая)	103
Выбор подходящего уровня детализации для измерений	104
Максимально просто, но не проще	104
Сводим все принципы воедино	105
Долгосрочное наблюдение	107
Bigtable SRE: когда оповещений слишком много	107
Gmail: предсказуемые ответы человека	108
Долгосрочная перспектива	109
Итоги главы	109
Глава 7. Эволюция автоматизации в Google	111
Польза автоматизации	111
Постоянство	112
Платформа	112
Быстрое восстановление	113
Быстродействие	113
Экономия времени	114
Польза автоматизации для Google SRE	114
Применение автоматизации	115
Применение автоматизации в Google SRE	116
Уровни автоматизации	117
Исключаем себя из процесса: автоматизируем все!	119
Облегчаем жизнь: автоматизируем процесс запуска кластера	121
Выявление несоответствий с помощью Prodtest	122
Идемпотентное разрешение несоответствий	124
Движение к специализации	126
Запуск кластера, ориентированный на сервисы	127
Vog: появление компьютера размером с дом	128
Основное качество — надежность	130
Рекомендации	132
Глава 8. Технологии выпуска ПО	133
Роль релиз-инженера	134
Основные положения	134
Модель самообслуживания	134
Высокая скорость	135

«Герметичные» сборки.....	135
Обязательные политики и процедуры	136
Непрерывная сборка и развертывание	136
Сборка.....	136
Ветвление.....	137
Тестирование.....	137
Пакеты	138
Rapid	138
Развертывание	140
Управление конфигурацией.....	140
Итоги главы.....	142
Это работает не только для Google.....	142
Управляйте релизами с самого начала.....	143
Глава 9. Простота	144
Стабильность или гибкость?.....	144
«Скучность» как добродетель.....	145
Не отдам свой код!	146
Показатель «Отрицательные строки кода».....	146
Минимальные API.....	147
Модульность	147
Простота релизов.....	148
Итоги главы.....	148

Часть III. Практики

Мониторинг.....	150
Реагирование в критических ситуациях	150
Постмортем и анализ основных причин	152
Тестирование	152
Планирование пропускной способности	152
Разработка.....	153
Продукт.....	153
Другие источники от Google SRE.....	154
Глава 10. Оповещения на основании данных временных рядов	155
Укрепление позиций Borgmon	156
Инструментарий для приложений.....	157
Сбор экспортированных данных.....	158
Память временных рядов как хранилище данных.....	159
Метки и векторы	160
Вычисление правил	162
Оповещение.....	167
Разбиваем топологию системы мониторинга на части	168

Мониторинг методом черного ящика.....	170
Обслуживаем конфигурацию	170
Десять лет спустя	172
Глава 11. Быть на связи	173
Введение	173
Жизнь дежурного инженера.....	174
Сбалансированная организация дежурств.....	175
Баланс количества.....	175
Баланс качества	176
Компенсация.....	177
Чувство безопасности	177
Избегание неуместной нагрузки	179
Операционная перегрузка	179
Коварный враг: операционная недоработка	181
Итоги главы.....	181
Глава 12. Эффективная диагностика и решение проблем	182
Теория	183
Практика.....	185
Отчет об ошибках.....	185
Первичная обработка и сортировка	186
Обследование	187
Диагноз	189
Подтверждение диагноза.....	192
Волшебная сила отрицательных результатов	193
Лечение.....	195
Пример: анализ реальной ситуации.....	196
Как облегчить решение проблем.....	200
Итоги главы.....	200
Глава 13. Реагирование в критических ситуациях	201
Что делать, когда система сломалась.....	201
Авария, вызванная тестированием.....	202
Подробности	202
Развитие ситуации	202
Подведение итогов	203
Авария, вызванная изменениями конфигурации	203
Подробности	203
Развитие ситуации	204
Подведение итогов	204
Авария, вызванная процессом	206
Подробности.....	206
Развитие ситуации	206
Подведение итогов	207

У всех проблем есть решение	208
Учитесь на прошлом опыте. И не повторяйте его.....	208
Храните историю перебоев в работе	208
Задавайте себе сложные и даже неправдоподобные вопросы: «А что если?..».....	209
Поощряйте упреждающее тестирование	209
Итоги главы.....	209
Глава 14. Управление в критических ситуациях	211
Неуправляемый инцидент	211
Анатомия неуправляемого инцидента.....	212
Излишняя сосредоточенность на технической стороне проблемы.....	212
Низкий уровень взаимодействия.....	213
«Вольный стрелок»	213
Элементы процесса управления в критических ситуациях	213
Рекурсивное разделение обязанностей.....	213
Выделенный центр управления	214
Обновляемый документ о состоянии инцидента	215
Четкая и оперативная передача полномочий.....	215
Управляемый инцидент	215
Когда следует сообщать об инциденте.....	217
Подведем итоги	217
Глава 15. Культура постмортема: учимся на ошибках	219
Философия постмортема от Google.....	220
Сотрудничайте и делитесь знаниями.....	222
Внедрение культуры постмортема	223
Итоги главы: непрерывные улучшения.....	225
Глава 16. Контроль неисправностей	227
Escalator	228
Outalator	228
Агрегирование.....	230
Маркировка.....	230
Анализ.....	231
Отчетность и общение.....	232
Неочевидная польза.....	232
Глава 17. Тестирование надежности систем	234
Виды тестирования ПО.....	236
Традиционное тестирование.....	236
Модульное тестирование	237
Интеграционное тестирование.....	237
Системное тестирование	238
Тестирование в промышленном окружении.....	239
Тестирование конфигураций.....	240
Нагрузочное тестирование	241

Окружения сборки и тестирования проекта.....	243
Масштабирование тестирования.....	246
Тестирование масштабируемых инструментов.....	247
Тестируем катастрофы.....	249
В погоне за скоростью.....	250
Передача в промышленную эксплуатацию.....	252
Ожидание сбоя тестов.....	253
Интеграция.....	255
Зондирование системы в промышленной эксплуатации.....	257
Итоги главы.....	260
Глава 18. Разработка ПО службой SRE	261
Почему так важна разработка ПО внутри службы SRE.....	262
Пример Auhop: история проекта и предметная область.....	263
Традиционное планирование производительности.....	263
Врожденная неустойчивость.....	264
Трудоемкость и неточность.....	265
Наше решение: планирование производительности, основанное на целях.....	266
Планирование производительности, основанное на целях.....	266
Предпосылки для целей.....	267
Знакомимся с Auhop.....	268
От требований до реализации: достижения и полученные уроки.....	271
Повышаем осведомленность и способствуем внедрению.....	273
Команда в развитии.....	276
Культивирование разработки ПО в службе SRE.....	277
Успешное создание культуры разработки ПО службой SRE: набор персонала и время разработки.....	278
Движемся к успеху.....	279
Итоги главы.....	281
Глава 19. Балансировка нагрузки на уровне фронтенда	282
Мощность — это не ответ.....	282
Балансировка нагрузки с использованием DNS.....	284
Балансировка нагрузки с использованием виртуального IP-адреса.....	287
Глава 20. Балансировка нагрузки в дата-центре	290
Идеальный случай.....	291
Выявляем плохие задачи: управление потоками и «хромые утки».....	293
Простой подход к контролю работоспособности задач: управление потоками.....	293
Усовершенствованный подход к контролю работоспособности задач: состояние «хромой утки».....	293
Ограничение пула соединений с помощью подмножеств.....	295
Выбираем правильное подмножество.....	295
Алгоритм выбора подмножества: случайное подмножество.....	296
Алгоритм определения подмножества: предопределенное подмножество.....	298

Политики балансировки нагрузки	301
Простой Round Robin	301
Least-Loaded Round Robin	304
Weighted Round Robin	306
Глава 21. Справляемся с перегрузками	308
Ловушки понятия «запросов в секунду»	309
Лимиты потребителей	310
Регулирование на стороне клиента	311
Критичность	313
Сигналы загрузки	314
Обработка ошибок, связанных с перегрузкой	315
Решаем выполнить повторную попытку	316
Нагрузка от соединений	319
Итоги главы	320
Глава 22. Справляемся с каскадными сбоями	322
Причины каскадных сбоев и способы их избежать	323
Перегруженность сервера	323
Истощение ресурсов	324
Процессор	325
Память	326
Потоки	326
Дескрипторы файлов	326
Зависимости между ресурсами	327
Недоступность сервисов	327
Предотвращаем перегруженность сервера	328
Управление очередями	330
Сегментация нагрузки и мягкая деградация	330
Повторные попытки	332
Задержка и дедлайны	335
Холодный запуск и холодное кэширование	338
Срабатывание условий каскадного сбоя	341
«Гибель» процесса	341
Обновления процессов	341
Новые релизы	341
Органический рост	342
Запланированные изменения, исчерпание ресурсов или отключения	342
Тестирование на предмет каскадных сбоев	343
Тестируйте до возникновения сбоев и после	343
Тестируйте популярные клиенты	345
Тестируйте некритические бэкенды	345
Мгновенное реагирование на каскадные сбои	345
Увеличьте количество ресурсов	346
Прекратите выполнять проверки на сбой/«гибель»	346

Перезапускайте серверы.....	346
Отбрасывайте трафик	346
Входите в деградированные режимы.....	347
Избавляйтесь от пакетной нагрузки	347
Избавляйтесь от плохого трафика.....	348
Итоги главы.....	348
Глава 23. Разрешение конфликтов: консенсус в распределенных системах и обеспечение надежности	350
Мотивация к использованию консенсуса: сбой координации распределенных систем.....	353
Пример 1: возникновение двух и более ведущих в схеме с ведущим и ведомыми.....	354
Пример 2: восстановление после сбоя требует вмешательства человека.....	354
Пример 3: некорректные алгоритмы членства в группе	355
Как работает распределенный консенсус	355
Шаблоны системной архитектуры для распределенного консенсуса	357
Надежные машины с реплицированным состоянием.....	358
Надежные реплицированные хранилища данных и конфигураций.....	359
Общедоступная обработка, применяющая алгоритм выбора лидера.....	359
Распределенная координация и блокировка сервисов.....	360
Надежный распределенный механизм помещения в очередь и отправки сообщений	361
Производительность систем с распределенным консенсусом	363
Multi-Paxos: детализированный поток сообщений.....	364
Масштабирование нагрузки, связанной с операциями чтения	366
Кворумная аренда.....	366
Производительность распределенного консенсуса и задержка сети.....	367
Размышляем о производительности: Fast Paxos	369
Стабильные лидеры	369
Пакетная обработка.....	370
Доступ к диску	370
Развертывание распределенных систем, основанных на консенсусе	372
Количество реплик.....	372
Расположение реплик	374
Производительность и балансировка нагрузки	375
Наблюдение за распределенными системами, основанными на консенсусе	381
Итоги главы.....	383
Глава 24. Cron: планирование и расписание в распределенных системах	384
Cron	385
Введение.....	385
Надежность	385
Задачи cron и идемпотентность.....	386

Масштабирование Cron	386
Расширенная инфраструктура.....	387
Расширенные требования	387
Создание cron в Google.....	388
Отслеживание состояния задач cron.....	389
Использование Paxos.....	389
Роли лидера и последователя	390
Сохраняем состояние	394
Запуск большого экземпляра Cron.....	395
Итоги главы.....	396
Глава 25. Конвейеры обработки данных	397
Происхождение шаблона проектирования «Конвейер».....	397
Основной эффект, который большие данные оказывают на шаблон простого конвейера	398
Сложности, связанные с шаблоном циклического конвейера	398
Проблемы, вызванные неравномерным распределением работы	399
Недостатки циклических конвейеров в распределенных средах	400
Наблюдение за проблемами в циклических конвейерах	401
Проблемы «шумной толпы»	402
Интерференция нагрузки	402
Знакомство с Google Workflow.....	404
Workflow как шаблон «Модель — представление — контроллер».....	404
Этапы выполнения в Workflow	406
Гарантии корректности Workflow	406
Гарантируем непрерывность бизнеса.....	408
Итоги главы и завершающие ремарки	410
Глава 26. Сохранность данных: как пишется, так и читается	411
Прямые требования к сохранности данных.....	412
Выбор стратегии для обеспечения отличной сохранности данных.....	413
Резервные копии или архивы?.....	415
Требования к облачному окружению на перспективу.....	416
Целевые значения показателей сохранности и доступности данных для SRE.....	417
Сохранность данных — это средство, доступность данных — цель.....	417
Создаем систему восстановления, а не систему резервного копирования.....	418
Типы сбоев, которые ведут к потере данных.....	418
Сложности поддержания высокой сохранности данных.....	420
Как служба SRE справляется с трудностями обеспечения сохранности данных	422
Двадцать четыре комбинации типов сбоев с точки зрения сохранности данных.....	422
Первый уровень: мягкое удаление.....	423

Второй уровень: резервные копии и связанные с ними методы восстановления.....	426
Основной уровень: репликация.....	429
От терабайта к экзабайту — это не просто «больше».....	429
Третий уровень: раннее обнаружение.....	430
Знаем, что восстановление данных работает.....	434
Примеры.....	436
Gmail — февраль 2011 года: восстановление с помощью GTape.....	436
Google Music — март 2012 года: определение бесконтрольного удаления данных.....	438
Общие принципы SRE, применяемые для сохранности данных.....	444
Мышление новичка.....	444
Доверяй, но проверяй.....	444
Надежда — плохая стратегия.....	444
Глубокая защита.....	444
Итоги главы.....	445
Глава 27. Надежный масштабируемый выпуск продукта.....	447
Управление запусками.....	449
Роль инженера — координатора запуска.....	450
Настройка процесса запуска.....	450
Чек-лист для запуска.....	451
Сходимость и простота.....	452
Запускаем неожиданное.....	453
Разрабатываем чек-лист для запуска.....	453
Архитектура и зависимости.....	454
Интеграция.....	454
Планирование производительности.....	455
Типы сбоев.....	455
Поведение клиента.....	456
Процессы и автоматизация.....	456
Процесс разработки.....	457
Внешние зависимости.....	457
Планирование отправок.....	458
Избранные приемы выполнения надежных запусков.....	459
Постепенные отправки и отправки в несколько этапов.....	459
Фреймворки флагов функциональности.....	460
Справляемся со злоупотреблениями клиента.....	462
Поведение при перегрузке и нагрузочные тесты.....	463
Развитие службы LCE.....	464
Эволюция чек-листа LCE.....	465
Проблемы, которые инженеры LCE не решили.....	466
Итоги главы.....	467

Часть IV. Управление

Глава 28. Ускоренное обучение SR-инженеров для работы на дежурствах и не только	472
Вы наняли новых SR-инженеров, что теперь?	472
Первый тренинг: структурировать хаос	475
Кумулятивные и методичные пути обучения	476
Целевая, а не черновая работа над проектом	478
Подготовка выдающихся реверс-инженеров, умеющих импровизировать	479
Обратное проектирование: узнаем, как работают системы	480
Статистика и сравнение: последователи научного метода под давлением	480
Импровизация: когда случается неожиданное	481
Связываем все воедино: обратное проектирование сервиса, находящегося в промышленной эксплуатации	481
Пять приемов для вдохновения дежурных работников	482
Охотники за сбоями: чтение отчетов о происшествиях и обмен ими	483
Катастрофа: ролевая игра	484
Ломаем и чиним по-настоящему	485
Документация как обучение	486
Периодически проводите теневые дежурства	487
На дежурстве и не только: обряд перехода и практика постоянного обучения	488
Итоги главы	489
Глава 29. Справляемся с отвлекающими факторами и прерываниями	490
Управляем операционной нагрузкой	491
Факторы, определяющие обработку поступающих запросов	492
Неидеальные машины	492
Состояние когнитивного потока	493
Делаем хорошо что-то одно	494
Серьезно, скажите, что мне делать	496
Уменьшение количества прерываний	498
Глава 30. Добавляем в команду нового SR-инженера, чтобы предотвратить операционную перегрузку	500
Фаза 1: изучаем сервис и рабочее окружение	501
Определите главные источники стресса	502
Определите очаги возгорания	502
Фаза 2: делимся контекстом	503
Напишите для команды хороший постморт	503
Сортируйте перегрузки в соответствии с их типами	504
Фаза 3: навязываем изменения	504
Начните с основ	505
Получите помощь в исправлении ошибок	505
Обосновывайте свои аргументы	505
Задавайте наводящие вопросы	506
Итоги главы	507

Глава 31. Общение и взаимодействие в службе SRE	508
Общение: производственные совещания	509
Повестка дня	510
Посещение	512
Взаимодействия внутри службы SRE	513
Структура команды	514
Приемы эффективной работы	514
Пример сотрудничества SR-инженеров: Viceroy	515
Пришествие Viceroy	515
Сложности	518
Рекомендации	519
Взаимодействие за пределами службы SRE	520
Пример: переход с DFP на F1	521
Итоги главы	524
Глава 32. Развитие модели вовлеченности SR-инженеров	525
Вовлеченность SR-инженеров: что, как и почему	525
Модель PRR	526
Модель вовлечения SR-инженеров	526
Альтернативная поддержка	527
Проверка готовности продукта: простая модель PRR	528
Вовлечение	529
Анализ	529
Улучшения и рефакторинг	530
Обучение	531
Освоение	531
Непрерывное улучшение	531
Развиваем простую модель PRR: раннее вовлечение	532
Кандидаты для раннего вовлечения	533
Преимущества модели раннего вовлечения	533
Развитие процесса разработки сервисов: фреймворки и платформа SRE	535
Усвоенные уроки	535
Внешние факторы, влияющие на SRE	536
На пути к структурному решению: фреймворки	536
Преимущества новой модели обслуживания и управления	538
Итоги главы	540

Часть V. Выводы

Глава 33. Полезные уроки из других отраслей	543
Познакомьтесь с нашими участниками	544
Готовность к катастрофам и их тестирование	546
Постоянная концентрация на безопасности	546
Внимание к деталям	547
Мгновенная производительность	547

Симуляции и прогоны.....	547
Тренировка и сертификация.....	548
Концентрация на сборе детальных требований и проектировании.....	548
Всесторонняя и глубокая защита.....	549
Культура написания постмортемов.....	549
Автоматизация и снижение служебной нагрузки.....	551
Структурированное и рациональное принятие решений.....	553
Итоги главы.....	554
Глава 34. Заключение.....	556

Приложения

Приложение А. Таблица доступности.....	560
Приложение Б. Практические рекомендации для сервисов в промышленной эксплуатации.....	561
Не бойтесь неудач.....	561
Постепенное выведение на рынок.....	562
Определяйте целевые значения с точки зрения пользователя.....	562
Бюджет ошибок.....	563
Мониторинг.....	563
Постмортемы.....	564
Планирование производительности.....	564
Перегрузки и отказы.....	565
Команды SRE.....	565
Приложение В. Пример документа о происшествиях.....	567
Приложение Г. Пример постмортема.....	569
Полученные уроки.....	571
Вспомогательная информация.....	573
Приложение Д. Список действий для координации запуска.....	574
Приложение Е. Пример протокола рабочего совещания.....	577
Об авторах.....	579
Библиография.....	580