

БНТУ

Научная библиотека



* 8 0 1 2 5 5 0 6 6 *

Масштабируемые данные

Лучшие шаблоны
высоконагруженных архитектур

Питхейн Стренгхольт



 ПИТЕР®

Санкт-Петербург · Москва · Минск

2022

Краткое содержание

Предисловие	13
Введение	15
Благодарности	20
От издательства	21
Глава 1. Проблемы при управлении данными.....	22
Глава 2. Введение в масштабируемую архитектуру: организация данных в масштабе.....	40
Глава 3. Управление большими объемами данных: архитектура хранилищ данных только для чтения.....	76
Глава 4. Управление сервисами и API: архитектура API	115
Глава 5. Управление событиями и ответами: потоковая архитектура.....	152
Глава 6. Соединение точек.....	188
Глава 7. Управление данными и их безопасность.....	223
Глава 8. Превращение данных в ценность	260
Глава 9. Управление основными корпоративными данными	293
Глава 10. Демократизация данных с помощью метаданных.....	311
Глава 11. Заключение	336
Глоссарий.....	346
Об авторе	366
Об обложке	367

Оглавление

Предисловие	13
Введение	15
Для кого предназначена эта книга	17
Что я узнаю?	18
Структура книги	18
Условные обозначения	19
Благодарности	20
От издательства	21
Глава 1. Проблемы при управлении данными.....	22
Управление данными	23
Аналитика фрагментирует ландшафт данных.....	26
Скорость доставки программного обеспечения меняется	28
Сети становятся быстрее	29
Приоритет — вопросы конфиденциальности и безопасности	29
Необходимо интегрировать операционные системы и системы обработки транзакций	30
Для монетизации данных требуется экосистемная архитектура	31
Предприятия обременены устаревшими архитектурами данных.....	32
Корпоративное хранилище данных и бизнес-аналитика	32
Озеро данных	35
Централизованное представление.....	38
Итоги главы	38
Масштабируемая архитектура	39
Глава 2. Введение в масштабируемую архитектуру: организация данных в масштабе	40
Общепризнанные отправные точки.....	41
У каждого приложения есть база данных.....	41
Приложения специфичны и обладают уникальным контекстом.....	41

Золотой источник.....	42
Дилеммы интеграции данных не избежать.....	42
Приложения играют роли поставщиков и потребителей данных.....	43
Основные теоретические соображения.....	44
Принципы объектно-ориентированного программирования.....	44
Предметно-ориентированное проектирование.....	46
Бизнес-архитектура.....	49
Шаблоны связи и интеграции.....	57
Двухточечная связь.....	57
Разрозненные хранилища.....	58
Звездообразная модель.....	59
Масштабируемая архитектура.....	60
Золотые источники и хранилища данных предметной области.....	60
Контракты на поставку данных и соглашения о совместном их использовании.....	63
Устранение разрозненного подхода.....	64
Предметно-ориентированное проектирование в масштабе предприятия.....	65
Оптимизация данных для чтения.....	68
Уровень данных как целостная картина.....	69
Метаданные и целевая операционная модель.....	73
Итоги главы.....	74

Глава 3. Управление большими объемами данных:

архитектура хранилищ данных только для чтения.....	76
Знакомство с архитектурой RDS.....	76
Разделение ответственности команд и запросов.....	77
Что такое CQRS.....	77
CQRS в масштабе.....	80
Службы и компоненты хранилища данных только для чтения.....	85
Метаданные.....	85
Качество данных.....	89
Уровни RDS.....	90
Получение данных.....	92
Интеграция готовых коммерческих решений.....	95
Извлечение данных из внешних API и SaaS.....	96
Служба исторических данных.....	97
Измерения только для добавления.....	100
Варианты проектирования.....	101

Репликация данных	103
Уровень доступа	104
Служба обработки файлов	106
Служба уведомлений о доставке	106
Служба удаления персональной информации.....	107
Распределенная оркестрация.....	107
Интеллектуальные службы потребления	108
Заполнение RDS по запросу	111
Рекомендации по использованию RDS	112
Итоги главы	113
Глава 4. Управление сервисами и API: архитектура API	115
Знакомство с архитектурой API	115
Что такое сервис-ориентированная архитектура.....	116
Интеграция корпоративных приложений	120
Оркестрация сервисов	122
Хореография сервисов.....	125
Публичные и частные сервисы	126
Модели сервисов и канонические модели данных.....	127
Сходства между SOA и архитектурой корпоративного хранилища данных	128
Современный взгляд на SOA.....	130
API-шлюз	130
Модель ответственности	132
Новая роль ESB	134
Контракты на обслуживание.....	135
Обнаружение сервисов.....	135
Микросервисы.....	136
Роль API-шлюза в микросервисах.....	138
Функции	139
Сервисная сетка	140
Границы микросервисов.....	142
Микросервисы в эталонной архитектуре API	142
Коммуникация между экосистемами.....	143
Каналы связи на основе API.....	145
GraphQL.....	147
Метаданные	148

Использование RDS для чтения в реальном времени и активного чтения	149
Итоги главы	150
Глава 5. Управление событиями и ответами: потоковая архитектура	152
Знакомство с потоковой архитектурой	152
Асинхронная модель событий имеет значение	153
Как выглядят событийно-ориентированные архитектуры.....	154
Топология посредника.....	155
Топология брокера	156
Стили обработки событий	157
Введение в Apache Kafka.....	158
Распределенные события.....	161
Возможности Apache Kafka.....	162
Потоковая архитектура	163
Производители событий	163
Потребители событий	166
Платформа событий	168
Источники событий и команд.....	169
Модель управления	172
Бизнес-потоки.....	173
Шаблоны потребления потоковой передачи.....	176
Передача состояния с помощью событий	178
Роли RDS	179
Использование потоковой передачи для заполнения RDS.....	180
Элементы управления и политики для управления доменами	180
Потоковая передача как операционный конвейер	181
Гарантии и согласованность.....	182
Уровень согласованности	182
Семантики обработки «хотя бы один раз», «точно один раз» и «не больше одного раза».....	183
Порядок сообщений	183
Очередь недоставленных сообщений	183
Потоковое взаимодействие.....	184
Метаданные для моделей управления и самообслуживания	185
Итоги главы	186

Глава 6. Соединение точек	188
Кратко об архитектурах.....	188
Архитектура RDS.....	189
Архитектура API	190
Архитектура потоковой передачи	190
Усиливающие шаблоны	191
Стандарты корпоративной совместимости	193
Конечные точки устойчивых данных	193
Контракты на доставку данных.....	196
Доступные и адресуемые данные.....	198
Принципы пересечения сетей.....	198
Стандарты корпоративных данных.....	204
Принципы оптимизации потребления.....	205
Возможность обнаружения метаданных.....	208
Семантическая согласованность	212
Предоставление соответствующих метаданных.....	216
Происхождение и перемещение данных.....	217
Эталонная архитектура	219
Итоги главы	221
Глава 7. Управление данными и их безопасность	223
Управление данными	223
Организация: роли в управлении данными.....	225
Процессы: деятельность по управлению данными.....	228
Люди: доверительные и этические, социальные и экономические соображения	229
Технологии: золотой источник, владение приложениями и их администрирование.....	230
Данные: золотые источники, золотые наборы данных и классификации	232
Безопасность данных	239
Текущий разрозненный подход.....	240
Единая защита данных для архитектур.....	241
Поставщики удостоверений.....	243
Эталонная архитектура безопасности и подход к контексту данных	244
Процесс безопасности	246
Практическое руководство	250
Архитектура RDS.....	250

Архитектура API	252
Потоковая архитектура	256
Интеллектуальный механизм обучения	258
Итоги главы	259
Глава 8. Превращение данных в ценность	260
Модели потребления.....	261
Прямое использование хранилищ данных только для чтения.....	261
Хранилища данных предметной области	262
Целевая операционная модель.....	264
Специалисты по данным как целевая группа пользователей.....	265
Бизнес-требования.....	267
Нефункциональные требования	267
Построение конвейера данных и модели данных.....	269
Распространение интегрированных данных	277
Возможности бизнес-аналитики	278
Возможности самообслуживания	280
Возможности аналитики	283
Стандартная инфраструктура для автоматизированного развертывания.....	284
Модели без сохранения состояния.....	285
Предварительно настроенные рабочие места.....	285
Стандартизация шаблонов интеграции моделей.....	286
Автоматизация.....	286
Метаданные модели.....	287
Эталонная архитектура продвинутой аналитики	288
Итоги главы	292
Глава 9. Управление основными корпоративными данными	293
Демистификация управления мастер-данными	294
Стили управления основными данными	294
Эталонная архитектура MDM	297
Разработка решения для управления основными данными	299
Распространение MDM.....	300
Основные идентифицирующие номера	300
Справочные и основные данные	302
Определение области видимости корпоративных данных	302
MDM и качество данных как услуга	305

Курируемые данные.....	305
Обмен метаданными	306
Интегрированные представления.....	307
Повторно используемые компоненты и логика интеграции	307
Повторная публикация данных	308
Связь с управлением данными	309
Итоги главы	309
Глава 10. Демократизация данных с помощью метаданных.....	311
Управление метаданными.....	312
Модель метаданных предприятия	313
Граф корпоративных знаний.....	321
Архитектурные подходы к управлению метаданными	325
Совместимость метаданных.....	326
Хранилища метаданных	328
Площадка для быстрого доступа к авторизованным данным.....	331
Итоги главы	334
Глава 11. Заключение	336
Модель доставки	337
Полностью децентрализованный подход.....	338
Частично децентрализованный подход	339
Структурирование команд.....	339
Стратегия InnerSource.....	340
Культура	341
Выбор технологий.....	342
Упадок традиционной архитектуры предприятия	343
Чертежи и схемы.....	344
Современные навыки	344
Контроль и управление	344
Послесловие	345
Глоссарий	346
Об авторе	366
Об обложке	367