

NoSQL

Um Guia Conciso para o Mundo Emergente
da Persistência Poliglota

Essencial

Pramod J. Sadalage
Martin Fowler

Novatec

Authorized translation from the English language edition, entitled *NOSQL DISTILLED: A BRIEF GUIDE TO THE EMERGING WORLD OF POLYGLOT PERSISTENCE*, 1st Edition, 0321826620 by SADALAGE, PRAMOD J.; FOWLER, MARTIN, published by Pearson Education, Inc, publishing as Addison-Wesley Professional, Copyright © 2013 Pearson Education, Inc.

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage retrieval system, without permission from Pearson Education, Inc. PORTUGUESE language edition published by NOVATEC EDITORA LTDA., Copyright © 2013.

Tradução autorizada da edição em língua inglesa, intitulada *NOSQL DISTILLED: A BRIEF GUIDE TO THE EMERGING WORLD OF POLYGLOT PERSISTENCE*, 1st Edition, 0321826620 por SADALAGE, PRAMOD J.; FOWLER, MARTIN, publicada pela Pearson Education, Inc como Addison-Wesley Professional, Copyright © 2013 Pearson Education, Inc.

Todos os direitos reservados. Nenhuma parte deste livro deve ser reproduzida ou transmitida em qualquer formato ou por qualquer meio, eletrônico ou mecânico, incluindo fotocópia, gravação ou qualquer sistema de recuperação de informações, sem permissão da Pearson Education, Inc. Edição em língua PORTUGUESA publicada pela Novatec Editora Ltda, Copyright © 2013.

© Novatec Editora Ltda. 2013.

Todos os direitos reservados e protegidos pela Lei 9610 de 19/02/1998. É proibida a reprodução desta obra, mesmo parcial, por qualquer processo, sem prévia autorização, por escrito, do autor e da Editora.

Editor: Rubens Prates

Tradução: Acauan Fernandes

Revisão técnica: BrodTec.com / Aurelio Jargas

Revisão gramatical: Cristiane Bernardi

Editoração eletrônica: Carolina Kuwabata

ISBN: 978-85-7522-338-3

Histórico de impressões:

Junho/2013

Primeira edição

Novatec Editora Ltda.

Rua Luís Antônio dos Santos 110

02460-000 – São Paulo, SP – Brasil

Tel.: +55 11 2959-6529

Fax: +55 11 2950-8869

E-mail: novatec@novatec.com.br

Site: www.novatec.com.br

Twitter: twitter.com/novateceditora

Facebook: facebook.com/novatec

LinkedIn: linkedin.com/in/novatec

MP20130611

Sumário

Prefácio	13
Parte I ■ Compreender	23
Capítulo 1 ■ Por que NoSQL?.....	25
1.1 O valor dos bancos de dados relacionais	25
1.1.1 Chegando aos dados persistentes	25
1.1.2 Concorrência	26
1.1.3 Integração.....	27
1.1.4 Um modelo padrão (na sua maior parte).....	27
1.2 Incompatibilidade de impedância	28
1.3 Bancos de dados de integração e de aplicativos	30
1.4 Ataque dos clusters.....	32
1.5 Surgimento do NoSQL.....	34
1.6 Pontos chave	38
Capítulo 2 ■ Modelos de dados agregados	41
2.1 Agregados	42
2.1.1 Exemplo de relações e agregados.....	43
2.1.2 Consequências da orientação a agregados.....	48
2.2 Modelos de dados de chave-valor e de documentos	50
2.3 Armazenamentos de famílias de colunas	52
2.4 Resumindo os bancos de dados orientados a agregados	55
2.5 Leituras complementares	56
2.6 Pontos chave	56
Capítulo 3 ■ Mais detalhes sobre modelos de dados.....	57
3.1 Relacionamentos	57
3.2 Bancos de dados de grafos.....	59
3.3 Bancos de dados sem esquema	61
3.4 Visões materializadas (Materialized views)	64

3.5 Modelando para o acesso aos dados.....	66
3.6 Pontos chave	72
Capítulo 4 ■ Modelos de distribuição	73
4.1 Um único servidor	74
4.2 Fragmentação	74
4.3 Replicação mestre-escravo	77
4.4 Replicação ponto a ponto (p2p)	79
4.5 Combinando fragmentação e replicação	81
4.6 Pontos chave	82
Capítulo 5 ■ Consistência	83
5.1 Consistência de atualização	83
5.2 Consistência de leitura	86
5.3 Relaxando a consistência	90
5.3.1 Teorema CAP	91
5.4 Relaxando a durabilidade	96
5.5 Quóruns.....	97
5.6 Leituras complementares.....	99
5.7 Pontos chave	100
Capítulo 6 ■ Marcadores de versões.....	101
6.1 Transações comerciais e de sistema	101
6.2 Marcadores de versões em múltiplos nodos	104
6.3 Pontos chave	106
Capítulo 7 ■ Map-Reduce (Mapear-Reduzir)	107
7.1 Map-reduce básico.....	108
7.2 Particionando e combinando	110
7.3 Criando cálculos map-reduce.....	113
7.3.1 Um exemplo de map-reduce em duas etapas.....	114
7.3.2 Map-reduce incremental	118
7.4 Leituras complementares	119
7.5 Pontos chave	120
Parte II ■ Implementar.....	121
Capítulo 8 ■ Bancos de dados de chave-valor	123
8.1 Depósitos de chave-valor.....	123
8.2 Recursos dos depósitos de chave-valor	126

8.2.1 Consistência	126
8.2.2 Transações	127
8.2.3 Recursos de consultas.....	127
8.2.4 Estrutura de dados.....	129
8.2.5 Escalabilidade.....	130
8.3 Casos apropriados para uso.....	131
8.3.1 Armazenando informações de sessão.....	131
8.3.2 Perfis de usuários, preferências	131
8.3.3 Dados de carrinhos de compras	131
8.4 Quando não utilizar	132
8.4.1 Relacionamentos entre dados	132
8.4.2 Transações com múltiplas operações	132
8.4.3 Consulta por dados	132
8.4.4 Operações por conjuntos	132
Capítulo 9 ■ Bancos de dados de documentos	133
9.1 O que é um banco de dados de documentos?.....	134
9.2 Características.....	135
9.2.1 Consistência	136
9.2.2 Transações	137
9.2.3 Disponibilidade.....	138
9.2.4 Recursos de consulta.....	140
9.2.5 Escalabilidade	142
9.3 Casos de uso apropriados	144
9.3.1 Registro de eventos (log)	144
9.3.2 Sistema de Gerenciamento de Conteúdo (CMS), plataformas de blog ..	145
9.3.3 Análises web ou em tempo real (analytics)	145
9.3.4 Aplicativos de comércio eletrônico	145
9.4 Quando não utilizar	145
9.4.1 Transações complexas que abranjam diferentes operações	145
9.4.2 Consultas em estruturas agregadas variáveis	146
Capítulo 10 ■ Armazenamento em famílias de colunas	147
10.1 O que é um depósito de dados de família de colunas?	147
10.2 Características.....	148
10.2.1 Consistência.....	151
10.2.2 Transações.....	153
10.2.3 Disponibilidade	154
10.2.4 Recursos de consulta.....	154
10.2.5 Escalabilidade	158
10.3 Casos de uso apropriados	158

10.3.1 Registro de eventos (log)	158
10.3.2 Sistemas de gerenciamento de conteúdo (CMS), plataformas de blog	159
10.3.3 Contadores	159
10.3.4 Expirando o uso.....	159
10.4 Quando não utilizar	160
Capítulo 11 ■ Bancos de dados de grafos	161
11.1 O que é um banco de dados de grafos?.....	161
11.2 Recursos.....	163
11.2.1 Consistência	165
11.2.2 Transações	165
11.2.3 Disponibilidade	166
11.2.4 Recursos de consulta	167
11.2.5 Escalabilidade.....	172
11.3 Casos de uso apropriados.....	173
11.3.1 Dados conectados	173
11.3.2 Roteamento, envio e serviços baseados em localização.....	174
11.3.3 Mecanismos de recomendação	174
11.4 Quando não utilizar	175
Capítulo 12 ■ Migrações de esquema.....	177
12.1 Alterações no esquema	177
12.2 Alterações de esquema em SGBDRs	177
12.2.1 Migrações para projetos sem restrições anteriores	178
12.2.2 Migrações em projetos legados	181
12.3 Alterações no esquema em um armazenamento de dados NoSQL.....	183
12.3.1 Migração incremental	185
12.3.2 Migrações em bancos de dados de grafos	187
12.3.3 Alterando a estrutura do agregado	188
12.4 Leituras complementares	189
12.5 Pontos chave	189
Capítulo 13 ■ Persistência poliglota.....	191
13.1 Necessidades diferentes de armazenamento de dados	191
13.2 Uso de armazenamentos de dados políglotas	193
13.3 Uso de serviços em vez do acesso direto ao depósitos de dados	194
13.4 Expandindo para melhorar a funcionalidade	196
13.5 Escolhendo a tecnologia certa	197
13.6 Preocupações empresariais com a persistência poliglota	197
13.7 Complexidade de instalação.....	199
13.8 Pontos chave	199

Capítulo 14 ■ Além do NoSQL201

14.1 Sistemas de arquivos	201
14.2 Event sourcing.....	202
14.3 Imagem de memória	205
14.4 Controle de versões.....	206
14.5 Bancos de dados XML	207
14.6 Bancos de dados de objetos.....	208
14.7 Ponto chave	208

Capítulo 15 ■ Escolhendo o seu banco de dados209

15.1 Produtividade do programador.....	209
15.2 Desempenho no acesso aos dados	212
15.3 Permanecendo com o padrão	213
15.4 Restringindo suas apostas	214
15.5 Pontos chave	215
15.6 Considerações finais	216