

# RDB技術者のための

---

# NoSQL

# ガイド

渡部 徹太郎 [監修・著]

河村 康爾／北沢 匠／佐伯 嘉康／佐藤 直生

原沢 滋／平山 毅／李 昌桓 [著]

**エンタープライズRDBエンジニアに贈る、  
NoSQL活用の最新バイブル登場！**

- ◎バズワード「NoSQL」を分解して実態を解明！
- ◎RDBの課題をいかにNoSQLで解決できるか？  
—HadoopやDWHとの使い分けも丁寧に説明！

- NoSQL 最新バージョン

MongoDB 3.2 / Cassandra 3.0 / Redis 3.0 /  
HBase 1.1 / Neo4j 2.3 / Couchbase Server 4.1

- クラウド

Amazon DynamoDB / Microsoft Azure DocumentDB



想定される  
ユースケースも  
多数掲載！



## エンタープライズRDBエンジニアに贈る、 NoSQL活用の最新バイブル登場！

- ◎バズワード「NoSQL」を分解して実態を解明！
- ◎RDBの課題をいかにNoSQLで解決できるか？  
—HadoopやDWHとの使い分けも丁寧に説明！

- NoSQL 最新バージョン  
MongoDB 3.2 / Cassandra 3.0 / Redis 3.0 /  
HBase 1.1 / Neo4j 2.3 / Couchbase Server 4.1
- クラウド  
Amazon DynamoDB / Microsoft Azure DocumentDB

想定される  
ユースケースも  
多数掲載！

 **香和システム**



RDB技術者のための  
NoSQL  
ガイド

渡部 徹太郎 [監修・著]  
河村 康爾／北沢 匠／佐伯 豊康／佐藤 直生  
原沢 澄／平山 翔／李 昌桓 [著]

■秀和システム

---

## 謝 辞

本書を発刊するに当たって共著していただいた河村 康爾さん、北沢 匠さん、佐伯 嘉康さん、佐藤 直生さん、原沢 滋さん、平山 穀さん、李 昌恒さんに感謝いたします。

また、本書の執筆に際して業務内容を調整していただいた菊地原 拓さん、内容を添削していただいた佐々木 政昭さん、清水 隆介さん、添田 健輔さん、糞庭 秀一郎さんに感謝いたします。

最後に、私が執筆できる時間を作るために、休日に子供の面倒を見てくれた妻に深く感謝いたします。

渡部 徹太郎

---

### 注 意

- 本書は著者が独自に調査した結果を出版したものです。
- 本書は内容に万全を期して作成しましたが、万一ご不審な点や誤り、記載漏れなどお気づきの点がありましたら、出版元まで書面にてご連絡ください。
- 本書の内容に関して運用した結果の影響については、上記にかかわらず責任を負いかねますのであらかじめご了承ください。
- 本書およびソフトウェアの内容に関しては、将来予告なしに変更されることがあります。
- 本書の例に登場する会社名、名前、データは特に明記しない限り、架空のものです。
- 本書の一部または全部は出版元から文書による許諾を得ずに複製することは禁じられています。

### 商 標

本書に記載されている会社名、製品名は各社の商標または登録商標です。

## はじめに

近年、NoSQLが当たり前になってきています。数年前は、NoSQLといえばオープンソースの尖った技術というイメージでしたが、近年は大手ベンダーが商用製品やクラウドサービスを続々と登場させており、エンタープライズでの利用事例も増えています。

今までに出版されたNoSQLの書籍は「最先端技術の解説書」というイメージが強かったと思います。それらに対し本書では、一般的RDBエンジニアを対象として、エンタープライズ環境でNoSQLをどのように活用すべきかを説明します。

本書で伝えたい事は三つあります。

一つ目は、NoSQLというバズワードの実態を説明することです。NoSQLという言葉は、特定のプロダクトを意味する場合もあれば、RDBに搭載されたSQLではないインターフェースを意味する場合もあり、定義があいまいです。そこで本書ではKVS、ドキュメントDB、グラフDB、そしてRDBのNoSQLインターフェースという明確な用語を用いて、NoSQLを分解し、説明していきます。

二つ目は、エンタープライズ視点でのNoSQL活用方法です。エンタープライズでデータベースを使う場合は、機能の評価だけでなく、運用、セキュリティ、サポート体制といった非機能に関する評価が重要ですので、その点をしっかりと説明します。また、企業におけるRDBの課題を解決するという視点で、NoSQLだけではなく、HadoopやDWHといったデータ処理技術との使い分けを明確に説明します。

三つ目は、NoSQLの正確な最新情報を記載している点です。NoSQLは進歩が速いため、昔の書籍やインターネットのブログに載っている情報は、間違っている可能性があります。本書では、紹介するデータベースの最新情報について、それぞれのデータベースのスペシャリストに正確な情報を掲載してもらっています。

2016年1月  
渡部 蔵太郎

## Contents 目次

### 第1章 前提

1-1 この本で伝えたいこと	2
1-1-1 NoSQLというバズワードの実態を理解	2
1-1-2 エンタープライズ視点のNoSQL活用方法	3
1-1-3 NoSQLの最新情報	6
1-2 想定する読者	6
1-3 本書の構成	7

### 第2章 イントロダクション

2-1 RDBだけだと辛くないですか？	10
2-1-1 RDBはデファクトスタンダード	10
2-1-2 Slerのエンジニアにとってのデータベースの経験	11
2-1-3 業務データを扱うだけでは不十分になってきている	12
2-1-4 RDBだけでは立ちいかない	14
2-1-5 RDBが適さない身近なエピソード	16
2-1-6 RDB以外を知ることが重要	24
2-1-7 これまでのまとめ	26
2-2 NoSQLとは	26
2-2-1 NoSQLはバズワード	27
2-2-2 KVS, ドキュメントDB, グラフDBの違い	27
2-2-3 NoSQLがバズワードになった背景	29
2-3 NoSQLにすると嬉しいこと・辛いこと	30
2-3-1 アプリケーション開発者にとって	31
2-3-2 データベース管理者にとって	37
2-3-3 マネージャや経営者にとって	40
2-4 よくあるNoSQLの勘違い	44
2-4-1 「パッチが高速になる」は勘違い	44
2-4-2 「トランザクションが高速になる」は勘違い	45

2-4-3	「ビッグデータを分析できる」は勘違い	46
2-4-4	「非構造データが効率的に扱える」は正確ではない	46
2-4-5	「RDBから置き換えると速くなる」は正確ではない	47
2-4-6	「オープンソースしかない」は昔の話	48
2-4-7	「スキーマがない」は昔の話	49
2-4-8	「SQLが使えない」は昔の話	49

## 第3章 データベースの中のNoSQLの位置づけ

<b>3-1</b>	<b>データベースを分類する2つの軸</b>	<b>52</b>
3-1-1	重視する性能による分類軸	52
3-1-2	性能拡張モデルによる分類軸	54
<b>3-2</b>	<b>データベースの4つのエリア</b>	<b>56</b>
3-2-1	RDB(OLTP)	57
3-2-2	RDB(DWH)	57
3-2-3	Hadoop(HDFS+MapReduce)	60
3-2-4	KVS	61
3-2-5	ドキュメントDB	62
3-2-6	グラフDB	63
<b>3-3</b>	<b>RDB(OLTP)とKVS/DocDBの違い</b>	<b>65</b>
3-3-1	RDB(OLTP)は強い整合性	66
3-3-2	強い整合性を保ったまま性能をスケールするのは困難	67
3-3-3	KVS/DocDBでは3つの工夫で性能をスケールさせている	69
3-3-4	CAPの定理	74
<b>3-4</b>	<b>HadoopとKVS/DocDBの違い</b>	<b>78</b>
3-4-1	Hadoop(HDFS+MapReduce)の動作	78
3-4-2	Hadoopと比較したときのKVS/DocDBの動作	81
3-4-3	4つのデータベースの比較表	82
<b>3-5</b>	<b>4つのエリアを超えて成長するデータベース達</b>	<b>83</b>
3-5-1	応答が速くSQLを扱えるHadoop	83
3-5-2	集計できるKVS/DocDB	84
3-5-3	SQLを扱えるNoSQL	84
3-5-4	JSONを格納するRDB(OLTP)	84
3-5-5	スケールアウトするRDB(DWH)	85

3-5-6	オペレーションも分析もできるRDB(DWH) ……	86
3-5-7	まとめ	86

## 第4章 データモデルごとの NoSQLプロダクト紹介

<b>4-1</b>	データモデルの種類	92
4-1-1	データモデルの説明	92
4-1-2	複雑度比較	98
4-1-3	データ間の関連度とスケーラビリティ比較	99
<b>4-2</b>	データモデル毎のプロダクトの紹介	101
4-2-1	キーバリュー モデルを採用するプロダクト	103
4-2-2	ワイドカラム モデルを採用するプロダクト	105
4-2-3	ドキュメント モデルを採用するプロダクト	107
4-2-4	グラフ モデルを採用するプロダクト	111

## 第5章 NoSQLの代表プロダクト紹介を 読む前に

<b>5-1</b>	紹介するプロダクトの選定基準	114
5-1-1	データモデルの中で広く使われていること	114
5-1-2	国内のサポート体制が整っているもの	115
<b>5-2</b>	プロダクト紹介の観点	115

## 第6章 Redis

<b>6-1</b>	概要	120
<b>6-2</b>	データモデル	121
6-2-1	データ型	122
6-2-2	永続化	127
<b>6-3</b>	API	131
6-3-1	クエリの実行例	131
6-3-2	利用できるクエリ	133
6-3-3	アプリケーションからの通信手段	135

6-3-4	部分的トランザクション	137
<b>6-4 性能拡張</b>		<b>139</b>
6-4-1	Redis Clusterのシャーディング	139
6-4-2	クエリの分散	140
6-4-3	リシャーディング	141
6-4-4	Redis Clusterのレプリケーションによる読み取り負荷分散	142
6-4-5	ハッシュタグを用いたRedis Cluster上での複数キー操作	142
<b>6-5 高可用</b>		<b>143</b>
6-5-1	レプリケーションによる可用性向上	143
6-5-2	フェイルオーバ	144
6-5-3	非同期レプリケーションによるデータのロスト	144
6-5-4	永続化していないマスターのリカバリ時の注意点	145
<b>6-6 運用</b>		<b>146</b>
6-6-1	バックアップ	146
6-6-2	監視	147
6-6-3	稼働統計	147
6-6-4	バージョンアップ	148
<b>6-7 セキュリティ</b>		<b>148</b>
6-7-1	パスワード認証	148
6-7-2	コマンドのリネーム・無効化	149
6-7-3	暗号化	150
<b>6-8 出来ないこと</b>		<b>150</b>
6-8-1	条件検索や集計などの処理が存在しない	150
6-8-2	ロールバック機能が存在しない	150
6-8-3	厳密な一貫性の担保	151
6-8-4	セキュリティ機能に乏しい	151
<b>6-9 主なバージョンと特徴</b>		<b>151</b>
<b>6-10 国内のサポート体制</b>		<b>152</b>
<b>6-11 ライセンス体系</b>		<b>153</b>
<b>6-12 効果的な学習方法</b>		<b>153</b>
<b>6-13 その他</b>		<b>154</b>
6-13-1	Redis Clusterの詳細	154
6-13-2	シャーディングとレプリケーションを組み合わせ	156

**第7章 Cassandra**

<b>7-1 概要</b>	160
7-1-1 Cassandraの特徴	163
7-1-2 Cassandraのニュースケース	164
7-1-3 OSS版と商用版	165
<b>7-2 データモデル</b>	165
7-2-1 Cassandraオブジェクト	167
<b>7-3 API</b>	168
7-3-1 Cassandra Query Language (CQL)	168
7-3-2 Cassandraのドライバ/コネクタ	170
7-3-3 軽量トランザクション	171
7-3-4 パッチ分析	171
7-3-5 外部 Hadoop のサポート	172
7-3-6 データの検索	172
7-3-7 分析と検索に対応したワーカーロードの管理	173
7-3-8 高速な書き込みと読み込み	174
<b>7-4 性能拡張</b>	176
7-4-1 クラスター/アーキテクチャの概要	176
7-4-2 Cassandraのクラスター、データセンタ、ノード	178
7-4-3 データの分散	179
7-4-4 クエリの分散	180
<b>7-5 高可用</b>	180
7-5-1 レプリケーションの基礎	181
7-5-2 マルチデータセンタとクラウドという選択	182
7-5-3 レプリケーション係数とクエリの整合性レベル	184
<b>7-6 運用</b>	186
7-6-1 クエリツール、管理ツール	186
7-6-2 バックアップとリカバリ	187
7-6-3 パフォーマンス管理	189
7-6-4 データの移行	195
<b>7-7 セキュリティ</b>	196
7-7-1 認証	196
7-7-2 権限管理	197
7-7-3 暗号化	197

7-7-4	データの監査	198
<b>7-8</b>	<b>出来ないこと</b>	<b>199</b>
<b>7-9</b>	<b>主なバージョンと特徴</b>	<b>199</b>
<b>7-10</b>	<b>国内のサポート体制</b>	<b>200</b>
<b>7-11</b>	<b>ライセンス体系</b>	<b>201</b>
<b>7-12</b>	<b>効果的な学習方法</b>	<b>201</b>
7-12-1	Cassandraの技術マニュアル	201
7-12-2	Cassandraトレーニング	202
7-12-3	Cassandraの技術情報、不具合情報	203
<b>第8章 HBase</b>		
<b>8-1</b>	<b>概要</b>	<b>206</b>
<b>8-2</b>	<b>データモデル</b>	<b>207</b>
<b>8-3</b>	<b>API</b>	<b>209</b>
8-3-1	テーブルの作成	210
8-3-2	データの格納	212
8-3-3	データの参照	213
8-3-4	データの更新	214
8-3-5	データの削除	215
8-3-6	テーブルの削除	215
8-3-7	APIについての補足	217
8-3-8	部分的トランザクション	217
<b>8-4</b>	<b>性能拡張</b>	<b>219</b>
8-4-1	HBaseクラスタのコンボーネント	219
8-4-2	データの分散とクエリの分散	221
<b>8-5</b>	<b>高可用</b>	<b>222</b>
<b>8-6</b>	<b>運用</b>	<b>224</b>
8-6-1	データのバックアップとリストア	224
8-6-2	監視と稼働統計	226
8-6-3	バージョンアップ	228
<b>8-7</b>	<b>セキュリティ</b>	<b>228</b>
8-7-1	データへのアクセス制御	228

8-7-2	操作記録	229
8-7-3	暗号化	229
<b>8-8</b>	<b>出来ないこと</b>	<b>230</b>
<b>8-9</b>	<b>主なバージョンと特徴</b>	<b>231</b>
<b>8-10</b>	<b>国内のサポート体制</b>	<b>232</b>
<b>8-11</b>	<b>ライセンス体系</b>	<b>232</b>
<b>8-12</b>	<b>効率的な学習方法</b>	<b>232</b>

## 第9章 Amazon DynamoDB

<b>9-1</b>	<b>概要</b>	<b>236</b>
9-1-1	概要	236
9-1-2	特徴	237
<b>9-2</b>	<b>データモデル</b>	<b>238</b>
9-2-1	アトリビュートのデータ型	239
9-2-2	DynamoDB JSON	240
9-2-3	キー	241
9-2-4	インデックス	242
9-2-5	DynamoDB Stream	243
<b>9-3</b>	<b>API</b>	<b>244</b>
9-3-1	APIとCRUD	244
9-3-2	アプリケーションから利用する	247
9-3-3	低レベルAPIと高レベルAPI	247
<b>9-4</b>	<b>性能拡張</b>	<b>250</b>
9-4-1	結果整合性	250
9-4-2	スルーブート、キャッシュティーコニット	251
9-4-3	パーティション	252
<b>9-5</b>	<b>高可用</b>	<b>253</b>
9-5-1	レプリケーション、フェイルオーバー	253
9-5-2	クロスリージョンレプリケーション	254
<b>9-6</b>	<b>運用</b>	<b>255</b>
9-6-1	監視	255
9-6-2	バックアップ	259

<b>9-7 セキュリティ</b> .....	<b>260</b>
9-7-1 セキュリティの考え方 .....	260
9-7-2 通信暗号化 .....	260
9-7-3 アクセスコントロール .....	261
9-7-4 監査 .....	262
<b>9-8 出来ないこと</b> .....	<b>264</b>
<b>9-9 国内のサポート体制</b> .....	<b>264</b>
<b>9-10 効果的な学習方法</b> .....	<b>265</b>
<b>9-11 その他</b> .....	<b>266</b>
9-11-1 バージョンと利用料 .....	266

## 第10章 MongoDB

<b>10-1 概要</b> .....	<b>270</b>
10-1-1 MongoDBの主な特徴 .....	271
<b>10-2 データモデル</b> .....	<b>272</b>
10-2-1 格納するデータの階層 .....	272
10-2-2 格納できるデータ型 .....	274
10-2-3 JSONのスキーマの事前チェック(ドキュメントバリデーション) .....	274
<b>10-3 API</b> .....	<b>275</b>
10-3-1 Mongoクエリ言語の概要 .....	275
10-3-2 CRUDのサンプル .....	276
10-3-3 CRUDの特徴 .....	278
10-3-4 集計 .....	279
10-3-5 アプリケーションからの使い方 .....	281
10-3-6 インデックス .....	282
<b>10-4 性能拡張</b> .....	<b>284</b>
10-4-1 シャーディングによる性能拡張 .....	284
10-4-2 セカンダリ読み込みによる読み込み負荷分散 .....	287
<b>10-5 高可用</b> .....	<b>287</b>
10-5-1 レプリケーションの概要 .....	288
10-5-2 フェイルオーバ .....	290
10-5-3 セカンダリの種類 .....	290
10-5-4 書き込み台数指定クエリ .....	291

<b>10-6 運用</b> .....	<b>292</b>
10-6-1 バックアップ .....	292
10-6-2 ヒューマンエラー対策(遅延レプリケーション) .....	293
10-6-3 監視・稼働統計 .....	293
10-6-4 バージョンアップ .....	294
10-6-5 MongoDB Ops Manager .....	295
<b>10-7 セキュリティ</b> .....	<b>296</b>
10-7-1 通信暗号化 .....	297
10-7-2 データ暗号化 .....	297
10-7-3 アクセスコントロール .....	297
10-7-4 監査 .....	297
<b>10-8 出来ないこと</b> .....	<b>298</b>
<b>10-9 主なバージョンと特徴</b> .....	<b>299</b>
10-9-1 バージョンのつけ方 .....	299
10-9-2 主なバージョンとその機能 .....	300
<b>10-10 国内のサポート体制</b> .....	<b>300</b>
<b>10-11 ライセンス体系</b> .....	<b>301</b>
<b>10-12 効果的な学習方法</b> .....	<b>302</b>
<b>10-13 その他</b> .....	<b>302</b>
10-13-1 便利な機能一覧 .....	302

## 第11章 Couchbase

<b>11-1 概要</b> .....	<b>306</b>
11-1-1 Couchbase という言葉 .....	306
11-1-2 Couchbase Server の主な特徴 .....	307
<b>11-2 データモデル</b> .....	<b>310</b>
<b>11-3 API</b> .....	<b>311</b>
11-3-1 データへのアクセス方法 .....	311
11-3-2 クライアントライブラリの各API実行サンプル .....	320
<b>11-4 性能拡張</b> .....	<b>324</b>
11-4-1 データ分散 .....	325
11-4-2 データアクセスの分散 .....	326

11-4-3 リバランスによる無停止でのクラスタ伸縮	328
<b>11-5 高可用</b>	<b>329</b>
11-5-1 クラスタ内レプリケーション	329
11-5-2 物理構成を意識したレプリケーション	330
11-5-3 複数クラスタ間のレプリケーション(XDCR)	331
<b>11-6 運用</b>	<b>334</b>
11-6-1 バックアップ	334
11-6-2 監視・稼働統計	335
11-6-3 バージョンアップ	336
<b>11-7 セキュリティ</b>	<b>337</b>
11-7-1 通信暗号化	337
11-7-2 管理者ユーザー、LDAP連携	338
11-7-3 監査ログ	338
<b>11-8 出来ないこと</b>	<b>339</b>
<b>11-9 主なバージョンと特徴</b>	<b>340</b>
11-9-1 バージョンの振り方	340
11-9-2 主なバージョンとその機能	341
<b>11-10 国内のサポート体制</b>	<b>341</b>
<b>11-11 ライセンス体系</b>	<b>342</b>
<b>11-12 効果的な学習方法</b>	<b>343</b>
<b>11-13 その他</b>	<b>344</b>
11-13-1 モバイルソリューション	344
11-13-2 便利な機能	346
11-13-3 ロードマップ	348
11-13-4 Couchbase Serverアーキテクチャ詳細	349

## 第12章 Microsoft Azure DocumentDB

<b>12-1 概要</b>	<b>358</b>
12-1-1 Microsoft Azure	358
12-1-2 Microsoft Azureのデータベース関連のサービス	360
12-1-3 Microsoft Azure DocumentDB	362

12-1-4	DocumentDBを使ってみよう	365
12-1-5	Azureの他の機能との連携	368
<b>12-2 データモデル</b>		<b>369</b>
12-2-1	リソースモデル	369
12-2-2	データモデル	371
<b>12-3 API</b>		<b>374</b>
12-3-1	REST API	374
12-3-2	クライアントSDK	378
12-3-3	インデックス	379
12-3-4	SQLクエリ(DocumentDB SQL)	383
12-3-5	ストアドプロシージャ、トリガ、UDF(ユーザー定義関数)、 トランザクション	389
12-3-6	文字列の検索	393
<b>12-4 高可用</b>		<b>394</b>
12-4-1	高可用性のためのアーキテクチャ	394
12-4-2	整合性レベルとレプリケーション	396
12-4-3	クライアントからの接続	401
<b>12-5 性能拡張</b>		<b>403</b>
12-5-1	コレクション	403
12-5-2	パーティション分割	405
12-5-3	パーティション分割に対応したアプリケーションの開発	407
12-5-4	.NET SDKを使用したパーティション分割	409
<b>12-6 運用</b>		<b>412</b>
12-6-1	管理と監視	412
12-6-2	バックアップ/リストア	413
<b>12-7 セキュリティ</b>		<b>414</b>
<b>12-8 出来ない事</b>		<b>416</b>
<b>12-9 国内のサポート体制</b>		<b>417</b>
<b>12-10 効率的な学習方法</b>		<b>417</b>

**第13章 Neo4j**

<b>13-1 概要</b> .....	<b>420</b>
13-1-1 グラフDBに向いている処理 .....	421
13-1-2 グラフDBに向いていない処理 .....	422
<b>13-2 データモデル</b> .....	<b>422</b>
13-2-1 グラフを構成する要素 .....	423
13-2-2 グラフデータの格納形式 .....	425
13-2-3 グラフデータモデル .....	427
<b>13-3 API</b> .....	<b>430</b>
13-3-1 Cypherクエリ .....	430
13-3-2 アプリケーションからのアクセス方法 .....	441
<b>13-4 性能拡張</b> .....	<b>443</b>
13-4-1 HAクラスターによる処理性能向上 .....	444
13-4-2 キャッシュキャッシングによる処理性能向上 .....	444
<b>13-5 高可用</b> .....	<b>445</b>
13-5-1 HAクラスターのアーキテクチャ .....	445
13-5-2 システム構成 .....	446
<b>13-6 運用</b> .....	<b>449</b>
13-6-1 バックアップ .....	449
13-6-2 リストア .....	449
13-6-3 パルクロード .....	450
13-6-4 監視 .....	450
13-6-5 ログ出力 .....	452
13-6-6 疆衝統計 .....	452
<b>13-7 セキュリティ</b> .....	<b>452</b>
<b>13-8 出来ないこと</b> .....	<b>453</b>
<b>13-9 国内のサポート体制</b> .....	<b>454</b>
<b>13-10 主要バージョンと特徴</b> .....	<b>455</b>
13-10-1 Neo4jのエディション間の比較 .....	455
13-10-2 ライセンス体系 .....	456
<b>13-11 効果的な学習方法</b> .....	<b>457</b>
13-11-1 公式ドキュメント .....	458

13-11-2 ユーザ会	458
13-11-3 書籍	459
13-11-4 他の日本語の資料	459

## 第14章 想定されるNoSQLのユースケース

<b>14-1 キャッシュ(Redis)</b>	<b>462</b>
14-1-1 RDBのスケールアップ・スケールアウトによる対処	462
14-1-2 RedisによるWebアプリケーションキャッシュ	463
14-1-3 まとめ	465
<b>14-2 IoT(モノのインターネット)基盤(Cassandra)</b>	<b>465</b>
14-2-1 RDBを用いた場合の課題	465
14-2-2 NoSQLによる課題解決	467
<b>14-3 メッセージ基盤(Cassandra)</b>	<b>471</b>
14-3-1 なぜNoSQL向きなのか	472
14-3-2 NoSQLを用いたメッセージ基盤の具体例	474
14-3-3 まとめ	477
<b>14-4 Hadoop連携(HBase)</b>	<b>478</b>
14-4-1 RDBの課題	478
14-4-2 HBaseによる解決	479
14-4-3 HBaseとMapReduceアプリケーションの連携	480
<b>14-5 モバイルアプリケーションに代表されるアプリケーションでの利用(DynamoDB)</b>	<b>483</b>
<b>14-6 AWSサービスとの連動性を意識した利用(DynamoDB)</b>	<b>484</b>
<b>14-7 ログ格納システム(MongoDB)</b>	<b>485</b>
14-7-1 RDBだと大変	486
14-7-2 MongoDBだと楽	487
14-7-3 まとめ	489
<b>14-8 ECサイトのカタログ管理(MongoDB)</b>	<b>490</b>
14-8-1 RDBだと大変	490
14-8-2 MongoDBだと楽	491
14-8-3 まとめ	492
<b>14-9 高速開発(MongoDB)</b>	<b>493</b>

14-9-1	RDBだと大変	493
14-9-2	MongoDBだと楽	493
14-9-3	Webフレームワークに組み込まれるMongoDB	495
14-9-4	まとめ	495
<b>14-10</b>	<b>業界横断型アプリ(MongoDB)</b>	<b>496</b>
14-10-1	RDBだと大変	496
14-10-2	MongoDBだと楽	497
14-10-3	まとめ	498
<b>14-11</b>	<b>Webアプリ(ユーザプロファイル/セッションストレージ(Couchbase))</b>	<b>498</b>
14-11-1	RDBで実現しようとした時の課題	499
14-11-2	Couchbaseによる解決	499
<b>14-12</b>	<b>Webアプリ(オムニチャネル/パーソナライズ(Couchbase))</b>	<b>500</b>
<b>14-13</b>	<b>データベースのグローバル展開/ディザスタークリアリティ(Couchbase)</b>	<b>502</b>
<b>14-14</b>	<b>モバイルとサーバのデータ同期(Couchbase)</b>	<b>503</b>
<b>14-15</b>	<b>リアルタイム詐欺摘発システム(Neo4j)</b>	<b>505</b>
14-15-1	概要	505
14-15-2	Neo4jによる解決	506
14-15-3	RDBでは実現が難しい	507
<b>14-16</b>	<b>適材人材の検索システム(Neo4j)</b>	<b>508</b>
14-16-1	概要	508
14-16-2	Neo4jによる解決	512
14-16-3	RDBでは実現が困難	513
<b>14-17</b>	<b>経路計算システム(Neo4j)</b>	<b>511</b>
14-17-1	概要	511
14-17-2	Neo4jによる解決	512
14-17-3	RDBでは非効率	513

## 第15章 NoSQLの選び方

<b>15-1</b>	<b>データ処理の課題を見極める</b>	<b>517</b>
15-1-1	NoSQLで解決するのが最適な課題はどれか?	518

15-1-2	NoSQLでは解決できないRDB(OLTP)の課題	520
15-1-3	NoSQLで解決するかわからないRDB(OLTP)の課題	521
15-1-4	NoSQLで解決が期待できるRDB(OLTP)の課題	524
<b>15-2</b>	<b>高い処理性能を出すためのNoSQLの選び方</b>	<b>525</b>
15-2-1	小規模なキーバリューならRedis	525
15-2-2	マルチデータセンタでどこでも書き込めるようにしたいなら CassandraかCouchbase	526
15-2-3	MongoDBは柔軟なデータ分散やレンジ指定クエリを速くしたい場合 …… 528	
15-2-4	Hadoopと一緒にならHBase	528
15-2-5	クラウド上でのスケーラビリティ獲得ならばDynamoDBや Microsoft Azure DocumentDBを検討	529
<b>15-3</b>	<b>半構造データを処理しやすいNoSQLの選び方</b>	<b>530</b>
15-3-1	ドキュメントDBはどれを選ぶべきか	531
<b>15-4</b>	<b>その他の選定の観点</b>	<b>536</b>
15-4-1	可用性の高いNoSQLの選び方	536
15-4-2	セキュリティの高いNoSQLの選び方	537
<b>15-5</b>	<b>本書にないNoSQLを選ぶ時のポイント</b>	<b>537</b>
15-5-1	ありがちな誤い文句に捕られない	538
15-5-2	性能比較を当てにしない	539
15-5-3	最新ドキュメントを見る	541
<b>Index</b>	<b>索引</b>	<b>542</b>
<b>Profile</b>	<b>著者プロフィール</b>	<b>547</b>

-----

第1章

---

---

前提

---

NoSQL

## 1-1

### この本で伝えたいこと

本書で伝えたいことは3つあります。

- NoSQLというバズワードの実態を理解
- エンタープライズ視点のNoSQL活用方法
- NoSQLの最新情報

#### 1-1-1

##### NoSQLというバズワードの実態を理解

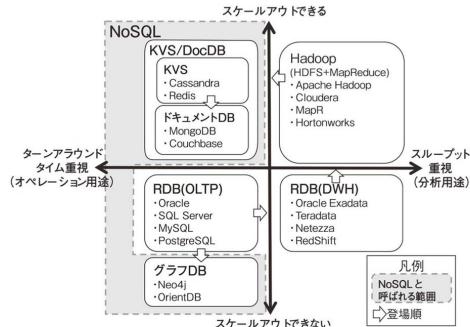
世の中のデータベースのほぼ全てはRDBです。読者の方もRDBだけを使ってきた人がほとんどでしょう。しかし、近年NoSQLと呼ばれるRDBではないデータベースが人気になってきています。昔は、NoSQLといえばオープンソースソフトウェアしかなく、最先端の技術というイメージでしたが、近年では大手ベンダからNoSQLの商用製品やクラウドサービスが続々と登場しており、いまやNoSQLは当たり前の技術の一つになってきました。

しかしNoSQLとはなんでしょうか?それを正しく説明するのはかなり困難です。言葉通り「RDB以外のデータベース」と思ってしまうと、大きな誤解になります。

NoSQLはバズワードです。NoSQLに明確な定義はなくNoSQLと呼ばれるデータベースの中でも各々性質が大きく異なります。

本書では、NoSQLという言葉はできるだけ使わずに、具体的なKVS、ドキュメントDB、グラフDBといったデータベース種別で分類します。そし

て関連するHadoopやDWHも絡めて、図1-1のように整理しています。この図を見てわかる通り、NoSQLはRDB以外の全てのデータベースを含んでいるわけではなく、かといってある特定の性質があるわけでもないのです。これがNoSQLがバズワードであることを物語っています。この図については本書で詳しく説明しますので、これを理解することでNoSQLの実態がはっきりと浮かび上がるでしょう。



●図1-1 NoSQLと呼ばれるデータベースが所属する範囲

## 1-1-2 エンタープライズ視点のNoSQL活用方法

NoSQLに分類されるデータベースは数多く存在します。オープンソースだけでなく商用製品やクラウドサービスでもNoSQLと名のつくプロジェクトが数多く登場しています。また、既存のデータベース製品もNoSQLのインターフェースを設けるようになってきました。NoSQLが氾濫していると言つてもいいでしょう。

そこで本書では、NoSQLの中からエンタープライズの本番運用に耐えるプロダクトに絞り込んで詳細に紹介します。一部の人しか使っていないミニマックなNoSQLや、開発体制やサポートが十分でないNoSQLは、エンタープライズでは利用できませんので紹介からは外します。広く普及して品質が安定しており、サポートが受けられ、エンタープライズでの事例があるNoSQLだけに絞り込んで紹介します。具体的には、キーバリューストアからはRedis、Cassandra、HBase、Amazon DynamoDB、ドキュメントデータベースからはMongoDB、Couchbase、MySQL、PostgreSQL、Oracle NoSQL Database、MarkLogic、Microsoft SQL Server、IBM DB2、Teradata、Google Cloud Datastore、Amazon ElastiCache、Microsoft Azure Redis Cache、Amazon DynamoDB、Microsoft Azure DocumentDB、IBM Cloudantを紹介します(図1-2)。

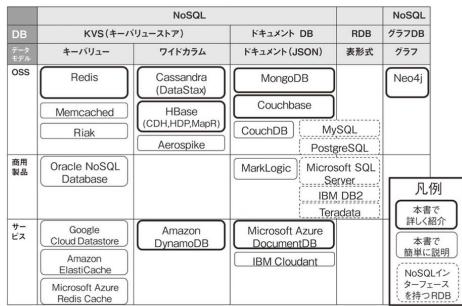


図1-2 NoSQL一覧と紹介するプロダクト

各プロダクトの説明は、プロダクトを選定する立場の読者を意識して、アーキテクチャや処理方式といった高い目線から説明します。よって、実作業で必要となるインストール手順やサンプルコードなどの詳細な情報は本書では省略します。プロダクトを選定するに最低限必要な知識を提供する本だと思ってください。

また、各プロダクトを横並びで比較しやすいように、各プロダクトの説明では観点をそろえています。具体的には、以下の観点でプロダクトを説

明します。

●概要

●機能

- データモデル
- API

●非機能

- 性能拡張
- 高可用
- 運用
- セキュリティ

●その他有用な情報

- 出来ないこと
- 主なバージョンと特徴
- 国内のサポート体制
- ライセンス体系

- 効果的な学習方法

加えて、NoSQLの想定されるユースケースを数多く紹介します。ユースケースは「商品カタログ管理」や「リアルタイム詐欺摘発」など、可能な限りアプリケーションに近い視点で紹介します。この説明を読んでいただくことによりどのプロダクトがどのユースケースに向いているか理解できるとともに、NoSQLを利用するイメージがアップするでしょう。

以上のように、本書を読んでいただければ、読者が直面している問題を解決するのに最適なNoSQLプロダクトを選択して活用できるようになるでしょう。

### 1-1-3 NoSQLの最新情報

現在、NoSQLは非常に活発に開発されており、多くのプロダクトは年に数回バージョンアップしています。また変化も激しく、1年前はサポートされていなかった機能が今ではサポートされていることは多いです。

本書では、NoSQLの代表プロダクトについて、それぞれのスペシャリストを集め最新の情報を記載しています。他のNoSQLの日本語書籍もありますが、それらとの違いは、本書が2015年12月時点の最新情報を集めた書籍であるという点です。具体的にはRedis 3.0、HBase 1.1、Cassandra 3.0、MongoDB 3.2、Couchbase 4.1、Neo4j 2.3、Amazon DynamoDB(2015年12月時点)、Microsoft Azure DocumentDB(2015年12月時点)であり、これらの日本語の正確な情報というのではありません。

また、各プロダクトの最新情報については、各プロダクトをエンタープライズでサポートしているスペシャリストの方々に執筆してもらっています。そのため、情報は正確で的確です。インターネットのブログの情報や英語のドキュメントを読むよりも確かに効率的に最新情報をキャッチアップできるでしょう。

## 1-2 想定する読者

本書が想定する読者は、RDBだけでシステムを開発しており、NoSQLについては初心者の方を対象としています。RDBの概念や基本用語については理解していることを前提としていますので、データベースを全く触ったことのない読者には向いていません。

また、NoSQLをエンタープライズの本番環境で運用しようと考えている方を前提とします。そのため、純粋な技術以外の学習コストや国内保守体制といった内容についても言及します。エンタープライズ以外の領域でNoSQLを学習する方にとっては少し的外れな内容になってしまいますので、その点はご了承ください。

最後に、想定している読者の職域は、技術を選定する立場にある方です。現場のアプリケーション開発者やデータベース管理者の方に対しては、ある程度有益な情報はあると思いますが、具体的な構築手順や操作手順など細かい内容は掲載されていません。

## 1-3 本書の構成

2章「イントロダクション」では、NoSQLの必要性や、NoSQL自体の簡単な説明、NoSQLを使うと嬉しいこと、NoSQLのよくある勘違い、など気軽に読める内容になっています。

3章以降は、体系立てた説明をしていきます。

3章「データベースの中のNoSQLの位置づけ」では、データベース全体を俯瞰した際にNoSQLがRDB(OLTP)、RDB(DWH)、Hadoopなどのように違うか説明します。また、NoSQLというあいまいな言葉をKVS、ドキュメントDB、グラフDBの三つに分けてそれぞれの特徴を説明します。

4章「データモデルごとのNoSQLプロダクト紹介」ではNoSQLをデータモデルで分類し、分類の中の具体的なプロダクトを紹介します。

5章「NoSQLの代表プロダクト紹介を読む前に」では代表プロダクトの詳細な説明をする上で、先に読んでおいてほしい点について記載しています。

す。そして、6章「Redis」以降から13章「Neo4j」までは具体的なプロダクトの詳細な説明になります。

14章「想定されるNoSQLのユースケース」ではNoSQLの想定されるユースケースを多数紹介します。これにより、NoSQLを利用するイメージをより具体化してもらえるでしょう。

最後に、15章「NoSQLの選び方」では、RDBの課題を整理した上で、それがNoSQLで解決できるのか、解決できるとしたらどのNoSQLを選ぶべきなのかを説明します。

---

無料サンプルはお楽しみいただけましたか？

この続きをお読みになるには...

[今すぐ購入](#)

または

[Kindleストアで本の詳細をチェック](#)

---