

クラウド時代の新しいデータベース

Web上で提供されるサービスの規模が巨大化するに伴って、従来のデータベースシステムとはアーキテクチャが異なる新しいタイプのデータベースが開発され、実際にサービスに導入されるなど実用性を高めている。本稿では、それら新しいデータベースの種類や特徴を解説するとともに、サービスサイトに適用する際のポイントや注意点を紹介する。

既存データベースシステムの限界

インターネットでサービスを提供するサイトは、データをデータベースサーバーに格納し、これをRDBMS（リレーショナルデータベース管理システム）で管理しているのが普通である。ところが会員数や取引量の急激な拡大にRDBMSが追いつけないケースが出てきた。負荷の分散などによってある程度は対応もできるが、サーバーが数十台を超えるようになると、従来のRDBMSによる管理は限界に達する。

この問題に対してはさまざまな手が打たれている（図1参照）。既存RDBMS製品でも、インメモリー化（データをDRAMに格納）や専用チップによるデータウェアハウスアプライアンス化（ハードウェアと一体化したデータベース）などによって処理の高速化を図っている例がある。オープンソースのアプリケーション「memcached」を用いて、キャッシュサーバーによりデータベースへの直接のアクセスを減らす工夫も可能である。

新しい考え方のデータベース

従来のデータベースはSQLという標準言語

によって操作するが、サービスの拡大に対応するためにNoSQLと呼ばれる新しいタイプのデータベースも開発されている。

(1) キー・バリュー型データストア

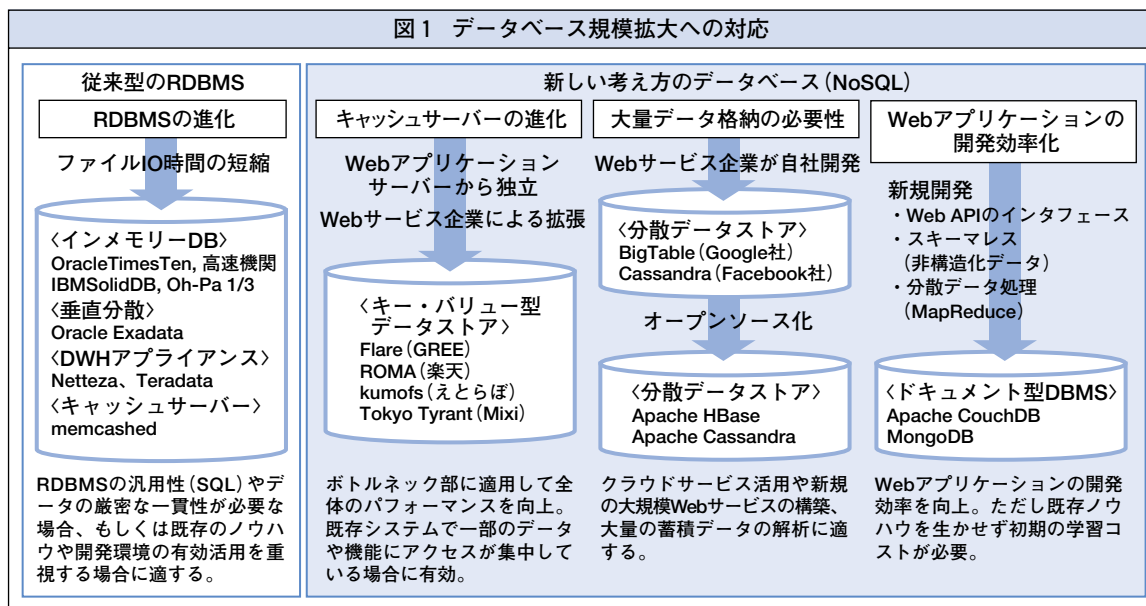
キャッシュサーバーは、システムの規模がさらに大きくなると不都合な点も出てきた。そこで楽天などの大規模なWebサービス事業者は、「memcached」を独自に進化させて自社サービスに適用するようになった。これはキー・バリュー型データストアと呼ばれるもので、①キャッシュではない永続的なデータ保存、②SQLという汎用性を捨て単純なデータ入出力による高速処理、③多ノード分散かつ動的拡張が容易、④ソースコードの公開という特徴を共通して持つ。開発言語や冗長化の実装方法は各サービス事業者ごとに異なる。

(2) 分散データストア

さらに最近になって、Web上のサービスに最適化された分散データストアと呼ばれる新しいデータベースが開発されている。これもサービス事業者が独自に開発したものである。多ノードの分散環境やクラウドコンピューティング（以下、クラウド）環境での大量データの格納や集計・加工に適し、高可用性、



図1 データベース規模拡大への対応



動的なノードの追加といった特徴がある。

(3) ドキュメント型データベース

上記のほか、Webアプリケーション開発の効率化を目指したドキュメント型データベースがある。スキーマレスでコンテンツデータを格納するのに適し、データ形式の変換やSQLの知識が不要である。

導入には慎重な検討が必要

新しいデータベースの多くは、サービス事業者が自らのサービスに適用することで効果が証明されている。しかし新しいデータベースは、どのサービスでも同様の効果が得られるとは限らない。それぞれに適した分野があり、システムの規模や特性によっては、従来型のRDBMSが最適な場合もある。

例えば「memcached」系のキー・バリュー

型の製品であればシステムのネック部分に適用することが有効であり、Google社のクラウドサービスを利用するのなら「BigTable」の利用が自然である。開発効率向上のために開発手法を見直すという目的であればドキュメント型も視野に入るだろう。

新しいデータベースのほとんどはオープンソース製品であり、大規模なクラウド環境で導入しても比較的成本を低く抑えられる。その反面、従来のRDBMSのような手厚い製品サポートや参考になる稼働例が少ないため導入にはそれ相応の技術力が必要である。新しいデータベースの導入には、長期的な規模の推移、必要な可用性のレベル、処理集中の見通し、データ項目の拡張の有無など、必要なシステムの特性を踏まえて適用する製品を見極めるなどの慎重な検討が必要である。■