



# クラウド向けデータストア 最適配置技術

## RDBや分散KVSの特徴を活かした自動選択・切替を実現

分散KVSは、AmazonやGoogleが大規模なサービス向けに採用しビジネスを成功に導いたことで、大きな注目を集めています。富士通研究所は、この分散KVSやRDBといった複数のデータストアを、扱うデータの特性に応じて自動的に選択し、Webアプリケーションからデータストアへのアクセスを適切に切り替えることが可能な、クラウド向けデータストア最適配置技術を開発しました。

### ①RDB

Relational DataBase。  
1件のデータを複数の項目(フィールド)の集合として表現し、データの集合をテーブルと呼ばれる表で表す方式。

### ②一貫性

参照時に、最後に書き込んだ情報を確実に読み取れることを保証すること。

### ③分散KVS

Key-Value Store(キーバリュ型データストア)の略。RDBで標準的にサポートしているトランザクション管理やテーブル結合等の機能は持っていない、シンプルなデータストアであるが、高いスケーラビリティを有している。

### ④NoSQL

Not only SQL、SQL(RDB)だけでなく、異なる種類のデータストアと、その使いこなしが重要という考え方。

### RDBの限界と分散KVSの台頭

現在、ICTシステムの多くで採用されているRDB<sup>①</sup>は、複雑な検索や集計を得意とし、データの一貫性<sup>②</sup>

を保証するロック機構を備えています。一方で、複数のデータベースサーバで処理を行う分散環境では、このロック機構のために、サーバの台数に比例した性能を得ることが困難です。運用開始後に柔軟にサーバの台数を増減することもできません。

さらに、アプリケーションが扱うデータの中には、構造化されていなかったり、書き込みが膨大であったりと、RDBに適さないものも多く、こうしたデータも含めてRDBに配置してしまうと、RDBの性能が低下します。

クラウドの時代となり、このような問題を解決しうる技術として注目されているのが、「分散KVS<sup>③</sup>」です。

分散KVSは、データの一貫性保証を行わない代わりに、複数のサーバでデータを管理することによるスケーラビリティや、仮にサーバが1台ダウンしてもサービスを継続できる高可用性を備え、かつRDBよりも低コストで実現することが可能です。

### データストアの特徴を活かしたデータ最適配置技術

富士通研究所では、NoSQL<sup>④</sup>という考え方に基づき、RDBや分散KVS等のデータストアの特徴を活

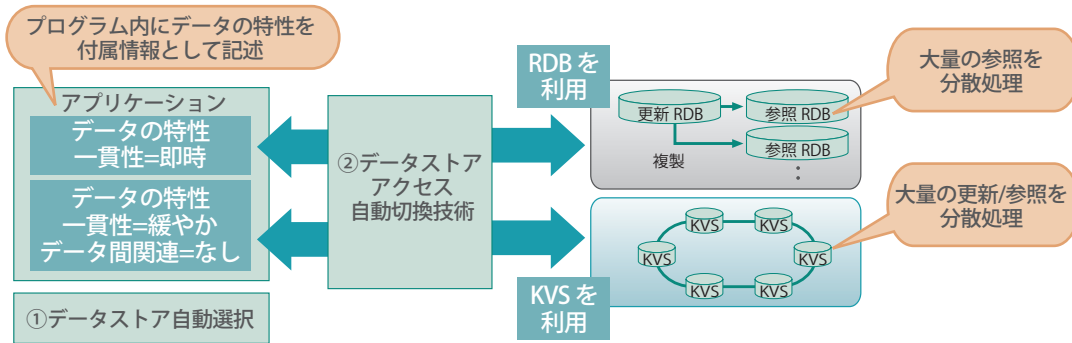
かし、データを最適なデータストアに配置する技術を開発しました。本技術では、業務データの特性に応じて、サーバ台数に応じた性能の向上を図れるシステムを自動的に構築します。プログラムロジックを変更することなく、データストアに対するアプリケーションのアクセスを切り替えることができるので、仕様変更等への迅速な対応が可能になります。

### 適切なデータストアの自動選択と実行環境の自動構築

データの一貫性やデータ間の関連性、検索条件の複雑性といったデータ特性をもとに、クラウドが適切なデータストアを選択。選択したデータストアにあわせて仮想マシン(VM)をプロビジョニングし、クラウド上の実行環境システムを自動的に構築します。運用中の高負荷発生時にはサーバ台数を動的に変更し、常に最適な構成を維持することが可能です。

データストアとしてはRDBと分散KVSを利

## ■ クラウド向けデータストア最適配置技術



用します。RDB固有の機能を必要としないデータは分散KVSで管理します。RDB利用時にも、一貫性を必要としない検索ではレプリケーション<sup>⑨</sup>で分散処理を行います。

データ特性の記述には、Java5で導入された付属情報を記述する「アノテーション(注釈)」を利用します。

### データストアの抽象化

複数のデータストアへのアクセスを、Java Persistence API (JPA) と呼ばれるアクセス方式で共通化しました。使用するデータストアにあわせて、データアクセス用のプログラムが動的に切り替えられるため、プログラムロジックの変更なくRDBや分散KVSを使い分けることが可能です。

#### スモールスタートと適正な投資が可能

本技術により、クラウド上でシステムを運用する企業では、エンドユーザーからのアクセスの増減に

対して柔軟な対応が可能なシステムを素早く構築できるようになります。また、サービス開始時は初期費用を抑えてスモールスタートし、ビジネス規模の拡大に応じてシステムを拡張していくといったことも可能です。

例えば、ニュースサイト等の有料情報提

供システムの立ち上げに際して、「可用性の維持と検索性能」「スモールスタートによる初期コスト抑制」といった要件があったとします。この場合のデータストア構成は、検索処理に必須となるインデックスのみをRDBに格納し、必要に応じてレプリケーションにより検索を高速化し、販売するコンテンツ情報は高可用の分散KVSで管理する、といった構成を素早く構築できます。

このように、アイデア次第で、適正な投資と短い開発期間で容易に市場に参入できるようになりますので、より付加価値の高いサービスや新ビジネスの提供が可能になります。

#### クラウドのさらなる有効活用に向けて

富士通研究所では現在、所内のクラウド環境にて本技術の検証を行っています。今後は、本技術をお客様に広く知っていただきながらフィードバックをもとに改良を加え、2011年度以降の実運用化を目指します。

富士通研究所は、長年蓄積してきたリソースを最大限活用するとともに、常に業界をリードする先端技術で、これからもお客様のビジネスに最適なクラウド技術をご提供していきます。

<sup>⑨</sup>レプリケーション  
データベースを複製し、参照系の負荷分散を可能とする技術。