



IOT およびエッジ開発者向けの 組み込みデータベース管理

By John K. Waters

実施の概要

IoT (internet of things : モノのインターネット) の出現とエッジ コンピューティングの急増は、多岐にわたる重要な課題をソフトウェア開発者に突きつけました。その中には、それらの急激な変化をもたらす傾向を助長する大量データの管理も少なからず含まれています。サードパーティ、組織のシステム、および、産業用センサーやアクチュエータからスマートフォン、自動運転車に至る数十億個のエンドポイント デバイスは、かつてない速度 (velocities) で空前のボリューム (volumes) と多様 (varieties) のデータを生成しています。

これらの「3つのV」は**ビッグ データ**の定義となっているため、開発者は構築する IoT およびエッジ アプリケーションでビッグ データを最大限に活用すること、つまり、ビッグ データを管理するだけでなく、ビジネス分析のために**アクション ポイント**でほぼリアルタイムの分析を生成すると共に、全体的なデータのガバナンスとセキュリティを提供することが期待されています。

このような分析レイヤーにより開発者にとっては複雑さが増すため、この環境で必要なものは古いクライアント / サーバー モデルでは得られないものになります。データをほぼリアルタイムで処理する必要がある場合には、すぐ隣にある地上のサーバーとクラウドのいずれにデータをプッシュするにせよ、アプリを配置するマシンと、必要なデータを配置するマシンとが異なるシステムに付き物のパフォーマンスラグが失敗の要因になります。製造、医療、金融サービスなどでは、ミリ秒が重要です。また、メモリに割り当てられた一部の領域のデータをファイル システムによって単に操作するだけでは、ゆくゆく解決策を得られなくなります。

開発者のソリューションは、データベースと管理ソフトウェアが実質的にアプリケーション自体の一部になっている、**組み込み可能なデータベース**にあることが増えています。このホワイト ペーパーでは、IoT およびエッジ コンピューティングの開発者が理解しておく必要のある、組み込み可能な DB の主要なコンセプトと実現技術について説明します。

定義

データベースの専門知識や、IoT およびエッジ コンピューティング分野の最新用語への精通度は開発者によって異なるため、いくつかの用語を定義しておくことは理にかなっていません。実のところ、データベースは、データベース管理者 (DBA) によって配置され、管理されるパッケージソフトウェアではもはやなくなっています。開発者としては、標準になりつつあるデータ集約型アプリを理解する必要があります。

組み込み

このホワイト ペーパーの焦点は**組み込み可能なデータベース**です。組み込み可能なデータ

ベースとは、操作するデータに直接または高速にアクセスする必要があるアプリケーションに統合されているか非常に緊密にバンドルされた、データベース (DB) およびさまざまなレベルのデータベース管理ソフトウェアです。DB とそれを使用するアプリは同じマシン上で実行されており、プロシージャ呼び出しを介して通信し合うため、待ち時間はほとんどありません。

このホワイト ペーパーにおいて、データ全体がメインメモリに格納される IMDB (in-memory database : **メモリ内データベース**) は、組み込みデータベースのカテゴリに含まれません。これらのシステムは軽量、シンプル、高速ですが、永続的なストレージを提供しません。

また、組み込みデータベースは**組み込み**システムではなく、ハードウェアに特定の機能がインストールされる専用ソフトウェアです。

この記事の執筆時点では、組み込みデータベースは主に ISV（独立系ソフトウェア ベンダー）、OEM（相手先商標製品製造会社）、および SI（システム インテグレーター）によって使用されています。しかし、IoT とエッジのエコシステムは急速に拡大しており、それに伴い、それらの市場も拡大することが予想されます。

データ分析は、実用的な洞察を得るためにデータを評価するプロセスです。一方、**組み込み分析**は、トランザクション型ビジネス アプリケーショ

生成および処理するものに、より近づけたもの」です。この用語は、エッジ コンピューティングの概念をマーケティング的に流行させるための用語として知られていますが、実際に定着しつつあります（「フォギング」は動詞にさえなっています）。

IoT

IoT (Internet of Things:モノのインターネット) は現在、人気のある用語の 1 つですが、意味を明確にするために、IoT により、インターネットおよび企業イントラネットを介して物理デバイス同士を接続した、拡大し続ける広大なネットワー

開発者のソリューションは、データベースと管理ソフトウェアが実質的にアプリケーション自体の一部になっている、組み込み可能なデータベースにあることが増えています。

ン内でレポート機能と分析機能を使用することです。組み込みデータベースがなければ、組み込み分析を行うことはできません。

エッジ コンピューティング

エッジ コンピューティングとは、データが生成された場所でデータの処理が発生する、つまり、集中管理されたデータ センターから遠く離れたネットワークの「エッジ」において処理が行われる、分散情報アーキテクチャのことを指します。IoT デバイスやリモート ゲートウェイで、このような処理が行われます。また、エッジ コンピューティングは、データの生成元のデバイス、またはその近くのデバイスで分散データの永続化が行われるものと説明されています。

フォグ コンピューティングは、2014 年にシスコ システムズによって造語された用語であり、事実上、エッジ コンピューティングと同じことを意味していますが、クラウド コンピューティングとより直接的に関係しています。フォグは、シスコが定義しているように、「クラウドを拡張して、IoT データを

クを指すことにします。このようなデバイスは、スマートフォンから温度センサー、台所用器具、さらには自動車まで、いまや何十億個もあります。それらのデバイスにインストールされたソフトウェアが、ローカルにおいて、または別のデータベースとの接続を介して、データを収集、保存、分析します。

IoT ゲートウェイ(別称:インテリジェント ゲートウェイまたは制御層)は、物理ハードウェア アプリアンスまたは専用ソフトウェアの形態を取ることができます。IoT ゲートウェイは、クラウドと、コントローラー、センサー、インテリジェント デバイスの世界とを接続する機能を果たします。

データベース管理システム

データベース管理システム (DBMS) は、データベースにおけるデータの保存、取得、削除、セキュリティ、および整合性を制御するソフトウェアです。組み込みデータベース ソリューションは、組み込み DBMS と呼ばれることもあります。

リレーショナル データベース

リレーショナル データベースは、情報を列と行から成る一連のテーブルに編成した DB です。各テーブルには、他のテーブルのデータに関係するデータが含まれています。このような関係は事前に定義されているため、データは、テーブルを再編成することなく、さまざまな方法でアクセスしたり、再構築したりすることができます。このようなタイプのデータベースは、トランザクションで強力な一貫性のある、オンライントランザクション処理 (OLTP) アプリケーション向けに設計されています。リレーショナル データベース管理システム (RDBMS) は OLTP には適していますが、オンライン分析処理 (OLAP) には適していません。

SQL (構造化問い合わせ言語)

構造化問い合わせ言語 (SQL) は、リレーショナル データベースに最も一般的に関連付けられている言語です。実際、「リレーショナル データベース」は「SQL データベース」とほぼ同義語です。SQL ステートメントは、リレーショナル データベースの情報に対する対話型クエリと、レポート用のデータ収集の両方に使用されます。SQL は、データベースを扱う開発者にとってマスターしなければならないツールです。

NoSQL

NoSQL データベースは非リレーショナル データベースであり、データが行と列に編成されたテーブルはありません。また、従来のリレーショナル データベースより緩い一貫性モデルを備えています。NoSQL データベースでは、ドキュメント、グラフ、Key-Value、インメモリ、検索など、さまざまなデータ モデルが使用されます。NoSQL データベースは、大量のデータ、短い待ち時間、および柔軟なデータ モデルを必要とするアプリケーション向けに最適化されています。

フラット ファイル

フラット ファイルは、文書処理およびその他の構造化のための文字やマークアップがすべて削除された、プレーン テキスト データベースです。これには、1 行に 1 レコードのデータから成る単一テーブルが格納されます。最も一般的なフラット ファイルの 1 例は、CSV (カンマ区切り値) ファイルです。このファイルのテーブル データは、ASCII (American Standard Code for Information Interchange) テキストの行に収集され、各テーブル セルの値はカンマで区切られます。各行は改行によって表されます。

ACID 対応

ACID (Atomicity : 原子性、Consistency : 一貫性、Isolation : 独立性、Durability : 永続性) は、データベース トランザクションの有効性を保証するための 4 つの属性を表す頭字語です。すべてのデータベース トランザクションは、データの整合性を保証するために **ACID 対応**である必要があります。各属性は次のとおりです。

Atomicity (原子性): 2 つ以上の異なる要素を含むトランザクションにおいて、1 つの要素が失敗するとトランザクション全体が失敗するという属性です。

Consistency (一貫性): トランザクションの一部としてデータベースに書き込まれるデータは、定義されているすべての規則と制限に従う必要があるという属性です。

Isolation (独立性): まだコミットされていない処理中のトランザクションは、他のトランザクションから分離されていなければならないという属性です。

Durability (永続性): トランザクションが正常に完了した時点で、データベースに加えられたすべての変更は永続的になるという属性です。

「データの量と速度が増加するにつれて、処理のためにこのすべての情報をクラウドまたはデータセンターにストリーミングすることの非効率性も増加します。」 –GARTNER 社アナリスト、SANTHOSH RAO 氏

展望

開発者が IoT、エッジ コンピューティング、およびこのような難解なデータベースの知識に関心を持つ必要があるのはなぜですか？開発者は結局、DBA ではなく、アプリケーションのリリース間隔の短縮化に関連した重要な課題に既に対処しています。

これらの傾向を無視することは絶対にできません。Gartner 社の業界アナリストによると、今後 4 年以内に、企業で生成されるデータの 75% が従来のデータセンターやクラウドではない場所で作成および処理されるようになります。Gartner 社アナリスト Santhosh Rao 氏は、2018 年のレポート（「What Edge Computing Means for Infrastructure and Operations Leaders」(インフラストラクチャおよびその運用を担当するリーダーにとってのエッジ コンピューティングの意味)）の中で、この傾向によって IT 組織に生じている問題を、「データの量と速度が増加するにつれて、処理のためにこのすべての情報をクラウドまたはデータセンターにストリーミングすることの非効率性も増加します」と総括しています。

また、地球上のほぼすべての産業は、これらの傾向に対応するために重要なリソースを投資する態勢にあります。最近公開された別のレポート（「Worldwide Semiannual Internet of Things Spending Guide」(世界のモノのインターネット支出ガイド (半期ベース))）では、IDC 社のアナリストが、2022 年に IoT 支出が 1.2 兆ドルに達すると予測しています。

これらの予測は、IoT およびエッジ コンピューティングに関する独自の戦略を検討している開発者に自信を与えるでしょう。開発者は自身が指定した結果を提供する機能を持つ API の

活用に最も関心を持っているかもしれませんが、IoT およびエッジ アプリケーションのデータを管理する API となると、組織は、ネットワークのエッジにより近い場所でのデータ処理を実現するテクノロジーへの投資により、それらのデータをほぼリアルタイムで分析できるようになります。この機能は、製造、医療、遠隔通信、金融など多くの業界ですぐに不可欠な機能となりました。そして、これらの業界では、それを実現するためのスキルを備えた開発者と、適切な API および基礎となるデータ管理機能が必要です。

使用事例

組み込み可能なデータベースの使用事例の分野は増え続けていますが、注目すべき最新例としては以下があります。

- 断続的に接続する環境や帯域幅が制限された環境。このような環境でデータを取り込み、後で分析するために少なくともデータの準備 / 処理 / フォーマットを行っている場合には、永続的に切断すると、壊滅的な状況になります。
- リアルタイムの意思決定に時間、頻度、またはその他の系列データを必要とする分析がデバイスまたはグリッドに組み込まれている場合（変化するベースライン データ セットに対する分析の場合はさらに適しています）。クラウド内のデータが氾濫することを避けるために、シンプルな分析はローカルで使用されることが多くなっています（AWS や Azure にデータを保持するのは低コストでは済みません。また、プライベート クラウド ストレージの減価償却でさえ、少額も積もれば相当な額となります）。

- 今後数年間で、シンプルな教師なし機械学習ルーチン（特に、ディープ ラーニング ネットワークにおけるノード処理用ルーチン）は、デバイス上か、デバイスのすぐ上のゲートウェイレベルに配置されるようになるでしょう。これを実現するには、ローカルの永続的な分散データ管理が必要となります。
- ビデオ ストリーミングなど、データの取り込み速度が速い場合。
- ピアツーピアの通信と制御を備えたメッシュセンサー ネットワーク。
- データ ガバナンス、セキュリティ、またはパフォーマンスの監視を必要とする、相互依存性（データまたはデバイス自体ではなく、それらに対応するメタデータ）のある異種セン

分析の複雑さが増しているためと、組み込み分析を介してそのリソースを活用できるかどうか競争上の主な優位性になろうとしているためです。開発者は、この格言を守りながら、選択を再考する必要があります。

今日の開発者の間でありがちな戦略は、プラットフォーム上で実行されているオペレーティングシステムに関連付けられているフラット ファイルシステム、または OS がない場合はローカルメモリ内のフラット ファイル システムを使用することです。フラット ファイルは、IoT およびエッジ環境においては厄介なものです。組み込みデータベース プロバイダーの Actian が最近の[レポート](#)（「フラット ファイルの使用を勧めない上位 10 個の理由」）の中で指摘している

「データはこれらのエンジンを駆動させる燃料であり、開発者はその基礎の燃焼エンジンの周囲に適切な部品を配置する必要があります。」

—ACTIAN 社マーケティング部長、LEWIS CARR

サーあるいはローカルのインテリジェント センサー グリッド。

- インストルメンテーション、複雑なプロセス、および規制監督が統合されたインテリジェントな資本設備。高価な機械や車両は、ますますサービスとして販売されるようになります。これを効果的に行う唯一の方法は、それらを広範囲にわたって装備することです。

どれを選択するかという問題

これらの傾向によって提起されるもう 1 つの疑問は、なぜフラット ファイルや、人気のあるオープン ソース データベースである SQLite を使用するだけはいけないのかということです。開発者は「シンプルにする」ことを推奨されており、組み込み可能なデータベースは複雑なもののように思われます。なぜ「車輪の再発明」をするのでしょうか？それは、データの管理と

ように、フラット ファイルのデータ一貫性を維持し、破損を回避するために、開発者はデータを注意して格納または取得するための作成、読み取り、更新、削除（CRUD）ロジックをコーディングする必要があります。このコーディングは、API やファイル システムによって変える必要があるかもしれません。

SQLite は、組み込みのデータ管理ソリューションとして企業に広く展開されているマルチプラットフォーム データベースです。SQLite は、手取り早くかつ簡単に使えて、SQL 開発者が持っている既存の専門知識を活用できるため、人気があります。SQLite がフラット ファイル システムよりも進歩したものであることは確かです。しかし、Actian が別の[レポート](#)（「SQLite から Actian Zen にアップグレードしなければならない 10 個の大きな理由」）で指摘しているように、SQLite には重大な欠点がある

あります。SQLite は、複数のアプリが同時に使用できるデータベースとして設計されていないのです。これは、アプリが進化し、マイクロサービスアーキテクチャを採用して、仮想サーバーインスタンスや、十分なリソースを備えたクライアント デバイスにまで展開されるようになった今日では、深刻な限界となっています。また、SQLite を使用するには通常、プラットフォーム間でのデータの変換とマッピングが必要になるため、複数の API を使用した設計とコーディングにかかる時間が長くなり、ETL のオーバーヘッドが増えて、保守サポートが面倒になります。SQLite は、サーバー機能がないため同時読み書きを処理することはできません。SQLite では ANSI SQL の完全サポートを提供していないため、アプリケーション コードに埋め込まれた一部の SQL 呼び出しでは、SQLite の内外に移動するという回避策が必要になります。

Actian のマーケティング部長である Lewis Carr は、最近のインタビューでこの状況を次のように簡潔に要約しています。「データはこれらのエンジンを駆動させる燃料であり、開発者はその基礎の燃焼エンジンの周囲に適切な部品を配置する必要があります。言わば、古いキャブレターをオーバーヘッド カムシャフト 24 バルブ燃料噴射装置に置き換える必要があるようなものです。」

組み込み可能なデータベースの要点

では、開発者は最新の組み込み可能なデータベース ソリューションから何を求めることができるのでしょうか？ 一般に、組み込み可能なデータベースは、アプリケーションの密結合を実現するためにコードへのパスの短い、プログラミング可能なインターフェイスを備えていて小型、高速、多用途、かつ信頼性が高い必要があります。ただし、IoT およびエッジ アプリケーションの開発者が組み込み DB を選択する作業に取り掛かる際に検討する必要がある特定の特徴と機能がいくつかあります。

SQL および NoSQL のサポート

NoSQL では、プログラミング API ベースの

データベース アクセスのサポート、SQL では、単一データ セットへのリレーショナル データベース アクセスのサポートが得られます。SQL アクセスがレポーティングおよびローカル トランザクションを扱うのに対し、NoSQL アクセスでは、高パフォーマンスとローカル分析サポートを提供しているほか、広範なプログラミング API を利用できます。これは必須機能です。

モジュール性

シングルユーザー クライアントのデータ管理が可能なライブラリのコア セットから、直接 Key-Value ストア エンジン、さらには本格的なエンタープライズ クラスのサーバー、つまり、マルチコアの VM クラウド環境で数千人のユーザーをサポートできるサーバーまで拡張可能な、モジュール式アーキテクチャを備えたソリューションが得られます。IoT およびエッジソリューションが異なれば、リソースのレベルが異なるため、モジュール性を重要要件と考える必要があります。ソリューションの例としては、16MB DRAM モジュールと 4GB フラッシュドライブを搭載して Android Things を実行する 32 ビット ARM ベースのセンサーもあれば、256 GB DRAM と 1 TB SDD ドライブを搭載して Windows Server 2019 を実行する、MRI マシンに組み込まれた 64 ビット Intel ベースのシングル ボード コンピューターの場合もあります。これらは文字どおり正反対ですが、共に IoT アプリというカテゴリに含まれています。

複数プラットフォームのサポート

重要な特徴：依然として Wintel が支配しているデスクトップの世界とは異なり、IoT とエッジ コンピューティングの世界はより細分化されています。開発者は、macOS、Android、Linux のバリエーションなどに対応する必要があります。これは、産業環境における運用技術 (OT) 環境にも当てはまります。IoT およびエッジ コンピューティングのプラットフォームをめぐる状況は、近い将来も、非常に細分化されたままになる可能性があります。

IOT およびエッジ ソリューションが異なれば、リソースのレベルが異なるため、モジュール性を重要要件と考える必要があります。

時系列データ

2018 年初頭に Eclipse Foundation の [IoT グループ](#)が実施した IoT 開発者への調査では、これらのアプリケーションによってあらゆる業界のさまざまなデータが収集されていますが、その約 62%が時系列データであることが判明しています。「時系列」とは、サーバー メトリック、アプリケーション パフォーマンス モニター、ネットワーク データ、センサー データ、イベントなど、時間経過と共に追跡、監視、ダウン サンプリング、集計される測定値またはイベントです。時系列データは、同じ測定値が定期的に記録される場所で発生します。なお、この調査における 2 位と 3 位は、それぞれデバイス情報とログ データでした。

以下のものも獲得できます。

- Intel および ARM での 32 ビットおよび 64 ビットのサポート
- すべての一般的なプログラミング言語から呼び出し可能なデータ管理 API
- 隠れた費用が突然かかることもない、無料の開発（多くの場合、オープン ソースと解釈されます）
- すぐに使える基本的な API のセット – シンプルかつ簡単で使いやすいサブセット（例：オープン、クローズ、読み取り、書き込み）
- 自分でわざわざ一から作り直したくないものを処理してくれる、より包括的な API セット：インデックス作成、ソート、検索、行列の転置など
- 開発がプラットフォーム間やプログラミング言語間で移行しても、API の変更はなし
- 異なるプラットフォームでの暗号化 / 復号化、

バックアップ、復旧など、基盤となるファイル システムの処理に必要なすべての変更は、データ管理システムによって抽象化され、処理される

- 組み込み DBMS を含む、展開されるソフトウェアは、開発者が構成可能である必要がある。そのため、展開は簡単であり、DBMS によって追加の構成と管理の複雑さが増すことはない
- レポーティングは、開発者がリモートで実行できるもの、または事前に構成および自動化するようにセットアップし、適切なデータが適切な形式で開発者、データサイエンティスト、または業務アナリストに配信できるものである必要がある
- ACID 対応
- DBA が不要：組み込み可能なデータベースを実装することで、開発者がデータベース管理者としての管理作業を常時行わなくても済むようになる

結論

世界は急速に「モノ」で満たされています。「地球上には人よりも多くのインターネット接続デバイスが既に存在する」と言う業界ウォッチャーもいます。また、食料品のバーコードをスキャンして毎週の食料品リストを作成できる冷蔵庫、ユーザーのスマートフォンに心拍数情報を送信するフィットネス トラッカー、または自身のメンテナンス チェックを指示する産業用機械のいずれのデバイスも、管理および分析する必要のあるデータを作用対象に近いところで生成しています。

ネットワークのエッジに発生するデータ管理および分析の必須要件を考えれば、組み込み可能なデータベースを IoT およびエッジ アプリケーション開発の実現技術と見なすことは妥当です。問題は、開発者がこのテクノロジーを採用すべきかどうかではなく、いつ採用するかです。最終的に、Volume (ボリューム)、Variety (多様性)、Velocity (速度) の3つの V が 4 番目の V である「価値」を生み出している必要があります。

詳しい情報：

<https://www.agtech.co.jp/action/product/>



John K. Waters は、シリコンバレーを拠点に活動している、フリーのジャーナリスト兼作家です。同氏は、『Application Development Trends』誌の報道記者であると同時に、『Campus Technology, T.H.E. Journal, and Law Technology News』誌の編集協力者です。同氏は IT における幅広いトピックを扱っており、十数冊の本の著者でもあります。同氏が執筆するニュースと特集記事の対象としては、ソフトウェア開発の動向、アプリケーションとネットワークのセキュリティ、インフラストラクチャとデータベースのテクノロジー、クラウド コンピューティング、オープン ソース ソフトウェア、規範とガバナンス、仮想化、IT ビジネスに従事する人々、および IT ビジネスに影響する問題があります。



株式会社エージーテック

本社：〒101-0054 東京都千代田区神田錦町1-17-5 Daiwa神田橋ビル6F

TEL：03-3293-5300（代表）FAX：03-3293-5270

カスタマーセンター TEL：03-3293-5283 Mail：info@agtech.co.jp

© 2019 Actian Corporation. Actian は、Actian Corporation およびその子会社の商標です。本資料で記載される、その他すべての商標、名称、サービス マークおよびロゴは、所有各社に属します。