

大量の実世界データから「今」を分析する ストリームデータ処理基盤 「uCosminexus Stream Data Platform」

電子マネーやICカードによる電子決済、センサやRFID^{※1}を利用した物流管理などが普及する中、多様な実世界の情報が大量かつリアルタイムに情報システムへと流れ込んでいます。これらのデータを瞬時に分析したいという市場ニーズに応えるため、日立は継続的に発生するデータをDBに格納せず、メモリ上で高速処理するストリームデータ処理基盤を開発。ビジネスに新たな価値をもたらす“瞬発力”を提供しています。

※1 Radio Frequency-Identification

瞬発力を実現するストリームデータ処理

実世界から絶え間なく流れ込んでくる膨大な情報をリアルタイムに分析し、より迅速な意思決定やリスク管理、現場状況の把握などに生かしたいという動きが本格化しています。金融業界のトレンドとなったアルゴリズムトレードや、通信サービス分野におけるSLA^{※2}管理などの分野で適用が進みつつある「複合イベント処理(CEP^{※3})」「イベントストリーム処理(ESP^{※4})」などの技術が注目されているのはこのためです。

現在、多くの企業で採用されているDBMS^{※5}などのデータ処理方式では、分析する情報を一度DBに格納した後、バッチ処理などで一括して集計・分析を行うため、情報の発生から集計・分析までには、どうしてもタイムラグが生じてしまいます。これまでの一般的なビジネス用途では問題ありませんが、“いま

起きている事象”への即対応という新しい要求に対応するためには、このタイムラグを極小化し、いかに“瞬発力を高めるか”が最大のプライオリティとなります。

そこで日立が開発したのが、CEPやESPの適用範囲をカバーしながら、大量データを瞬時に分析することでビジネスに新たな価値を創造する「ストリームデータ処理」という新技術です(図1)。

※2 Service Level Agreement
 ※3 Complex Event Processing
 ※4 Event Stream Processing
 ※5 DataBase Management System

集計・分析シナリオを記述が容易なCQLで定義

日立のストリームデータ処理は、データ発生時に、あらかじめ登録した「集計・分析シナリオ」に従って必要なデータを抽出し、即座に結果を出力します。その

基盤製品となる「uCosminexus Stream Data Platform」(以下、uCSDP)では、対象となるデータを高速なメモリ上に展開した上で、時系列データの中で必要な差分計算のみを行う手法を採用しています。これにより、従来のDBMSに対して二桁以上もの高速性を実現するほか、分析済みのデータをHDDなどに蓄積しないため、ストレージ容量の削減にも貢献します。

このストリームデータ処理において最も特徴的なのが、多様な分析ニーズに対応できる「集計・分析シナリオ」と、その定義を簡単に記述できるスクリプト言語「CQL^{※6}」の存

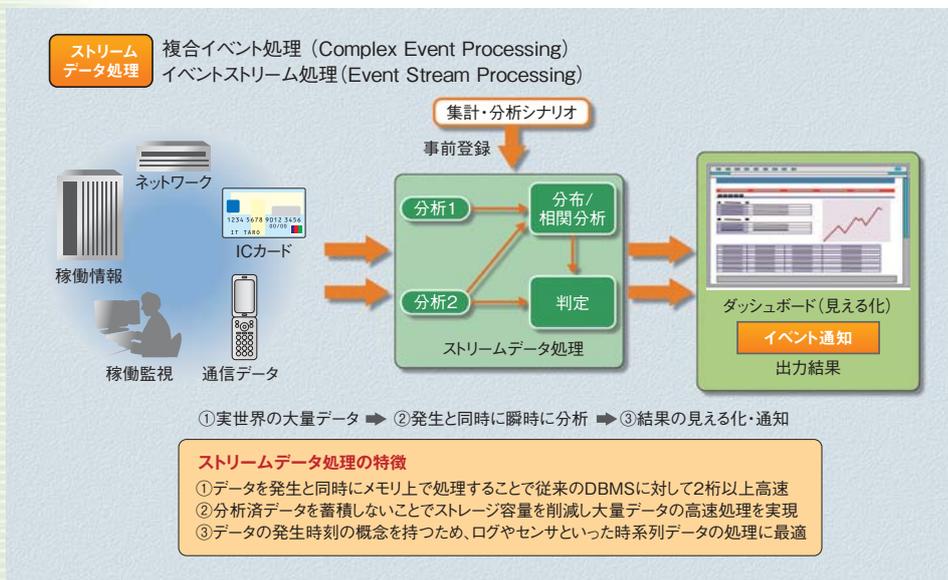


図1 瞬発力を実現するストリームデータ処理

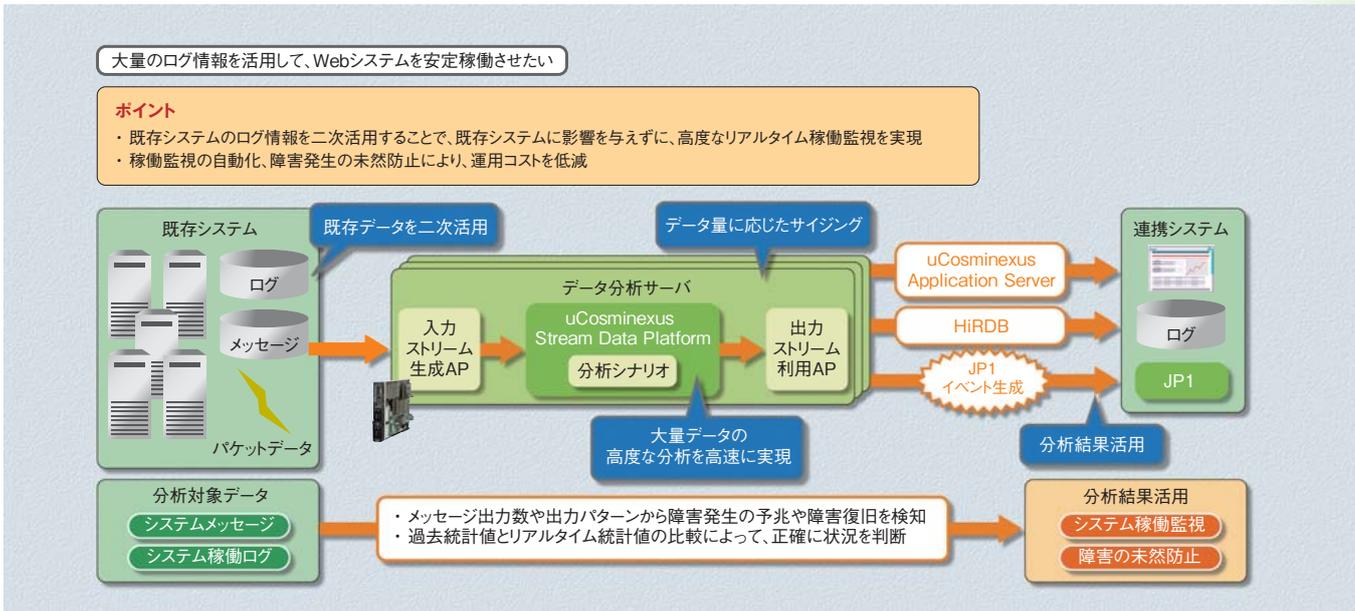


図2 ストリームデータ処理によるWebシステム稼働監視の例

在です。一般的なデータベース言語SQL^{※7}を拡張したCQLは、汎用的な言語であるため誰にでも扱いやすく、幅広い業種に対応できます。CQLを連結することで、複雑な時系列処理も簡単に記述することが可能です。

また、シナリオベースであるため、集計・分析用のアプリケーション開発を行う必要もなく、シナリオをテキストファイルに定義し、uCSDPに登録するだけで使用できます。このため、開発効率が大幅に向上し、シナリオの追加・変更だけで分析内容の変化に柔軟に対応できます。

uCSDPに加えて、実世界データを活用したモニタリングを実現する入出力アダプターや、集計・分析結果表示用の環境設定機能などのアプリケーションフレームワークも提供されるため、短期間で容易にシステム導入が行える点も大きな魅力といえます(図2)。

※6 Continuous Query Language
※7 Structured Query Language

データの活用範囲を飛躍的に拡大させ 新たな価値を創出

ストリームデータ処理は、大量のフローデータを瞬時に分析することで、「いつもと違う」現象の変化を即座に検知できるソリューションです。このため、不正アクセスや機密ファイルのダウンロードなどをリアルタイムに検知・ブロックする「コンプライア

●この記事は、はいたっく別冊「Open Middleware Report vol.54」に掲載されたものです。

お問い合わせ先

HMCC(日立オープンミドルウェア問い合わせセンター)

☎0120-55-0504

利用時間 9:00~12:00, 13:00~17:00(土・日・祝日・弊社休日を除く)

携帯電話、PHS、一部のIP電話など上記フリーダイヤルがご利用いただけない場合

TEL (03) 5439-2733

情報提供サイト

<http://www.cosminexus.com/>

<http://www.hitachi.co.jp/cosminexus/>

スチェック」のほか、株価や出来高の解析により売買注文を自動化する「アルゴリズムトレード」、GPS端末のセンサ情報を利用したリアルタイム位置情報監視による「渋滞検出」、人の滞留状態に合わせて広告内容を瞬時に切り替える「デジタルサイネージ」など、幅広い分野への応用が期待されています。

もちろん、RFIDデータを活用したリアルタイムな物流管理や流通トレーサビリティ、大量のログ情報解析による業務システムのモニタリングなども得意分野の1つです。例えば、仮想化やクラウドの進展でシステム構成が一段と大規模化・複雑化しているITシステムの稼働監視に適用すれば、膨大なログの中から正常時とは異なる障害発生の予兆を検知できるため、障害が起こってからへの対応ではなく、障害を未然に防ぐことでダウンタイムを最小化する、文字どおりの「プロアクティブな予防保守」が恒常的に可能となります。

このように、データの流れを瞬時に分析し、必要なアクションを即座に実行できる「瞬発力」の高さは、まさにストリームデータ処理ならではの大きなビジネスバリューといえます。

現在、あらゆる業務分野において、より鮮度の高いデータに基づいた迅速な意思決定や状況分析、リアルタイム監視によるリスクマネジメント、スピードを付加価値とした新ビジネスの創出が強く求められています。その意味でも、ストリームデータ処理の可能性は今後さらに拡大し、従来のDBMSと並ぶ新たなメインストリームの技術へと進化していくことは間違いありません。