

# アシスト、超高速 ETL ツール「Syncsort DMExpress」の新バージョン 9.6 を提供開始

～Hadoop 関連を中心に強化を図り、ビッグデータ ETL 基盤としての機能をさらに拡充～

株式会社アシスト(本社:東京都千代田区、URL:<https://www.ashisuto.co.jp/>、以下、アシスト)は、取扱製品である超高速 ETL ツール「Syncsort DMExpress」(開発元:Syncsort Incorporated)の新バージョン 9.6 を 2月28日付けで提供開始します。

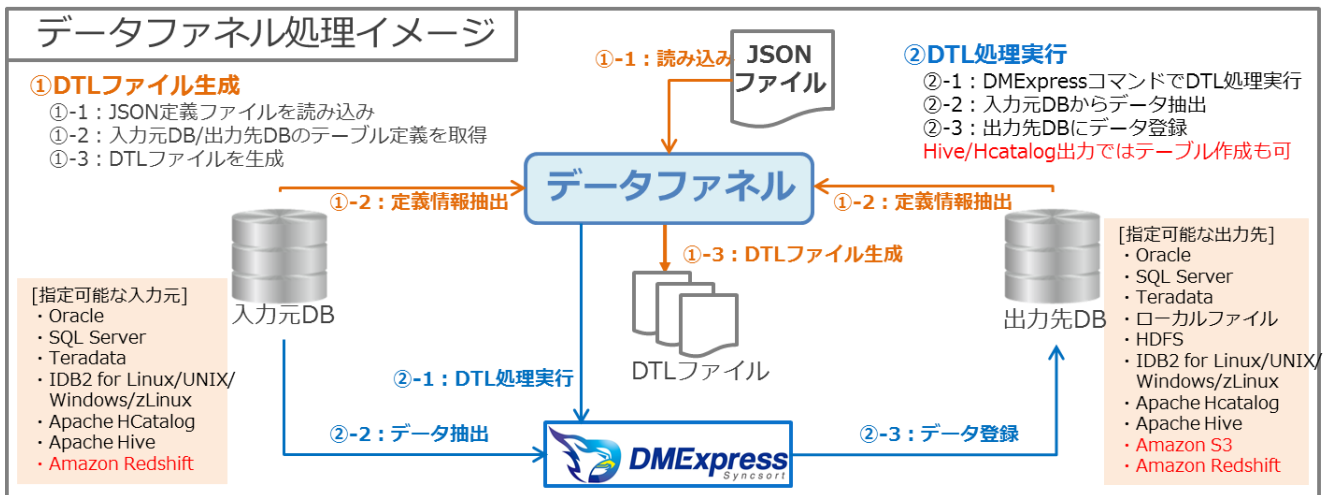
「Syncsort DMExpress」は、独自のスマート ETL オプティマイザを搭載した自動チューニング機構が、処理対象のデータ分析とシステムリソース分析を行い、最適なアルゴリズムを組み合わせることで ETL 処理を実行します。そのため、テスト運用での性能改善やチューニングに工数をかけることなく、ノンプログラミング & ノンチューニングで高品質/高性能な ETL 処理を誰でも簡単に開発することが可能です。

新バージョン 9.6 では、データファネル機能、ジョブ実行状況の監視等に関する機能が拡張され、また、Hadoop 環境上で並列分散処理を高速に実行する「Syncsort DMX-h」では、実行環境や接続先の拡張も実装されました。ニーズの高まるデータレイクの構築や、Hadoop 上でのデータ加工等、ビッグデータ ETL 基盤として、さらに充実した機能を提供します。

バージョン 9.6 の主な新機能は下記のとおりです。

## <Syncsort DMExpress>

### ● データファネル機能(\*)の拡張



※データファネルの実行には、Java 7以降が必要となります。

主要なデータベースのデータ移行時に必要な複数の処理定義を、DTL(テキストベースのコマンド言語)形式で自動生成する「データファネル」に、下記の機能が追加されました。

- 1)Amazon S3 を出力先に、Amazon Redshift を入力元/出力先に指定可能
- 2)データベースとの接続に Hive JDBC を含む接続方法を指定可能
- 3)Hive/HCatalog を出力先とする場合に、テーブル作成を行うパラメータを指定可能

※ データファネル機能の詳細は、下記ページ「データ移行処理の開発における 4 つの課題とその解決策とは～最も賢い超高速 ETL ツールが課題を解決!～」もご参照ください。

<https://www.ashisuto.co.jp/product/category/etl/syncsort-dmexpress/download/#tab>

- **ジョブ実行状況の監視機能の追加**  
DMExpress および DMX-h の MapReduce/Spark として実行されるジョブ／タスクのリアルタイムな進行監視と進捗状況のレポート作成が行えるようになりました。
- **キーブレイク処理関数の追加**  
キーが同じ場合、異なる場合の処理(キーブレイク処理)用の関数が追加され、累計計算のようなロジックを簡単に実装できるようになりました。
- **デバッグ機能の強化**  
データ変換ができない場合などのワーニング発生時に、レコード内容の他にレコード番号を実行ログに出力する機能が追加され、ワーニング対象のレコードの特定が容易になりました。
- **メタデータの DTL エクスポート**  
値、条件、レイアウト、データベース接続定義、照合順序といった GUI 上で設定したメタデータ定義を、テキスト実行形式の DTL ファイルとして出力する機能が追加され、定義を再利用することができるようになりました。

### <Syncsort DMX-h>

- **Spark 実行機能の拡張**
  - 1) Spark2.0 に対応
  - 2) Spark のデプロイモードとして、Client モードの他に Cluster モードでのジョブ実行に対応
- **MapReduce/Spark 実行形式に自動変換する「インテリジェント実行機能」の強化**  
DMX-h のジョブの中に外部プログラムを実行するカスタムタスクを含めた場合でも、インテリジェント実行が適用できるようになりました。
- **ETL 処理と Hadoop 処理の統合ワークフロー**  
ローカル実行/MapReduce 実行/Spark 実行を定義するタグをサブジョブやタスクに対して設定できるようになり、ETL 処理と Hadoop 処理を統合されたジョブとして定義できるようになりました。
- **Hive 接続の拡張**
  - 1) 検索入力データソースとしても使用可能
  - 2) タスクエディタから Hive テーブルを新規作成することが可能
  - 3) Hive 出力後にテーブル統計分析/カラム統計分析を実施するパラメータが追加
  - 4) Hive の Boolean 型カラムに対応
  - 5) Apache Ranger 認証に対応
- **入出力先の拡張**  
入出力データソースとして、HiveJDBC 接続が追加されました。

### ■「Syncsort DMExpress」および「Syncsort DMX-h」について

パフォーマンスに特化したソフトウェアを 40 年以上開発している Syncsort 社の“最も賢い超高速 ETL”DMExpress は、データ結合やソート／マージ処理を伴うバッチ処理、またデータウェアハウス、データ・マート構築時のデータ結合処理などの ETL 処理開発に威力を発揮する ETL ツールです。独自の高速処理アルゴリズム、並列処理などにより、システム資源を最大限に有効活用してバッチ処理時間を劇的に短縮します。メインフレーム環境からのダウンサイジング時に課題となる高速性と安定性を確保し、コモディティなオープン環境でもメインフレームと同等以上のパフォーマンスを実現します。アシストは、通信、金融、公共業界をはじめ 3,500 ライセンス、390 社以上の導入実績を誇ります。また、「DMX-h」は、Hadoop フレームワーク上での並列分散処理を実現します。Map/Reduce 処理を DMExpress の GUI で簡単に開発し、各 Hadoop データノードで稼動する C++ベースの「DMX-h」によって、Java、Pig、Hive などに比べて 2~5 倍※の処理性能向上が得られます。 ※Syncsort 社調べ

## ■ニュースリリースに関するお問い合わせ

株式会社アシスト 広報担当:田口

TEL:03-5276-5850 URL:<https://www.ashisuto.co.jp/contact/press/>

## ■「Syncsort DMExpress」および「Syncsort DMX-h」に関するお問い合わせ

株式会社アシスト 東日本技術本部 ビジネス推進部 担当:宇野

TEL:03-5276-3653

Syncsort DMExpress URL:<https://www.ashisuto.co.jp/product/category/etl/syncsort-dmexpress/>

Syncsort DMX-h URL:<https://www.ashisuto.co.jp/product/category/etl/syncsort-dmx-h/>

※ 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。

※ ニュースリリースに記載された製品／サービスの内容、価格、仕様、お問い合わせなどは、発表日現在のものです。その後予告なしに変更されることがあります。あらかじめご了承ください。