TechTarget
テックターゲットジャパン
Powered by ITmedia Inc.

HITACHI
Inspire the Next

日立が提供するETLツールのスタンダード「DataStage®」

データ統合をスピーディかつ低コストに実現する秘訣は

開発ツールの生産性

DataStage

日々データは蓄えられているのに、各システムに散在しており、戦略的に活用できていない。
社内データの統合を試みたものの、膨大なコストが……。
こうした課題を解決するのがETLツール「DataStage」だ。

データ統合の切り札「ETLツール」

今日、企業では多種多様な情報が社内システムに日々蓄積され、そのデータ量はとどまることなく増え続けている。しかも、これらのデータは社内の複数のシステムに散在し、各システムが個別の形式でデータを保管しているため、多くの企業において業務効率の低下を招く一因となっている。

例えば、顧客データ。営業活動やマーケティング業務に顧客データを活用しようとしても、各部署でそれぞれ独自に顧客データベースを構築しているため、どれが最新のデータでどこに自分の欲しいデータが格納されているのか見当が付かずに手間取ってしまう、といった具合だ。

一方で、このように社内に散在しているデータを統合・連携し、業務効率の改善を図るとともに、戦略的な意志決定に有効活用しようという動きも活発化し始めている。ただし、ここで課題になってくるのが、従来の汎用プログラミングツールを使った手組み開発では多大な時間とコストが掛かってしまうという点だ。また、いったんデータを統合した後も、経営環境の変化とともにデータ構造の変更・追加作業が定常的に発生する。そのため、手組み開発ではその都度プログラミング作業が発生し、保守費用がかさんでしまう点も無視できない。このようなコストをかんがみた場合、中堅企業においてはデータ統合の必要性に迫られながらも、なかなか実行に踏み切れていないのが実情のようだ。

そこで、こうした課題を解決するソリューションとして大きな注目を集めているのが、散在するデータを「抽出」（Extract）、「加工・編集」（Transform）、「格納」（Load）する処理を効率的に開発できる「ETLツール」である。これを活用することにより、データ統合におけるデータ移行、データ連携、データ加工の処理をノンプログラミングで開発でき、情報活用の基盤となるデータの統合・連携をスピーディーかつ低コストに構築することが可能となる。

そして、数あるETLツールの中でも、国内市場で売り上げシェアNo.1*の実績を持つ製品が、日立製作所の「DataStage」だ。同製品は、近年拡大を続ける国内ETLツール市場において、2003年度以降売り上げ金額ベースでトップシェアを維持し続けている。

開発生産性と保守性が鍵を握る

DataStageがシェアNo.1をキープし続けているのには、もちろん理由がある。まず同製品の最も大きな特徴として挙げられるのが、データ統合における開発生産性と保守性の高さだ。

直感的に理解しやすい開発環境

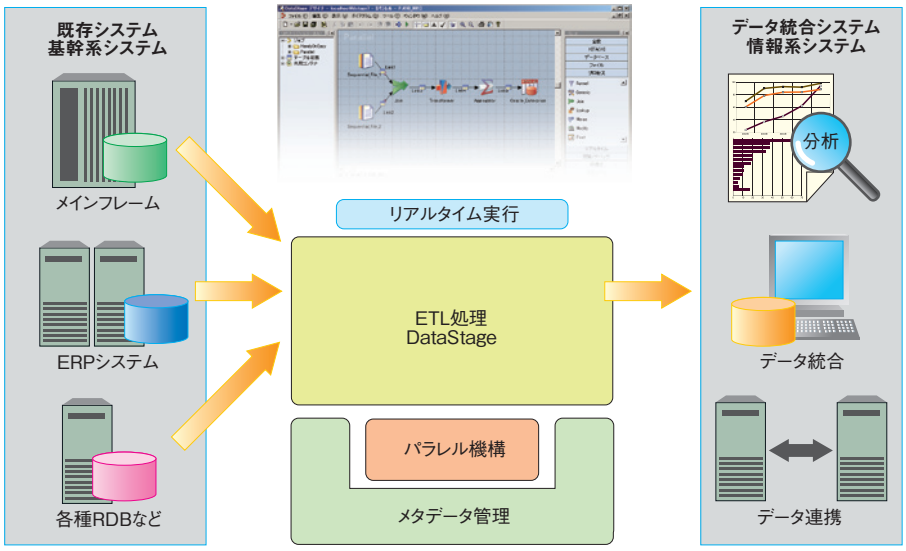
DataStageでは高い開発生産性を実現するために、誰にでも直感的に分かりやすいGUI操作によるビジュアルなノンプログラミング開発環境を用意している。

同製品はデータの抽出・変換・統合の一連の処理を「ジョブ」という単位で管理している。このジョブを開発するには、データ抽出・変換・集計などの処理単位を表すアイコンをドラッグ＆ドロップで画面上に配置し、それらの間を線で結んでデータの流れを指定するだけだ。詳細な設定も、各アイコンのプロパティとして簡単に指定できる。視覚的に理解しやすく、極めて直感的なインタフェースだといえよう。

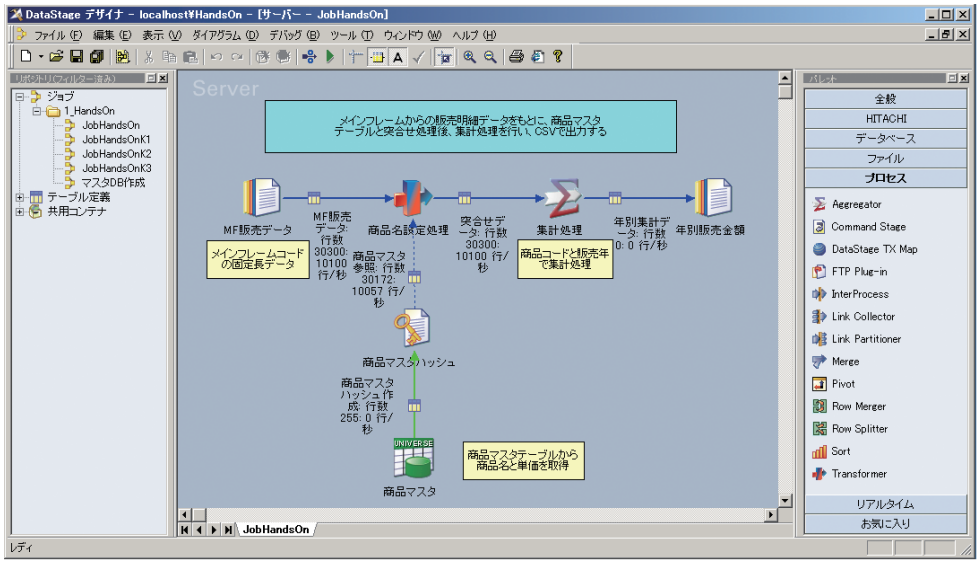
複数のシステムからデータを抽出し、豊富な加工機能を用いて整合性のあるデータを作成するまでの一連の処理の流れを直感的に表現できるため、そのジョブを開発した本人以外の開発者でも内容を

* 富士キメラ総研「パッケージソリューション・マーケティング便覧」（2004年～2009年発行）から引用。

DataStageの位置付け



DataStageのデザイン画面



すぐ理解できるようになっている。

また、HiRDB、Oracle Database、Microsoft SQL Server、IBM DB2、Teradata、Sybase、Informixなどの各種DBMSを幅広くサポートしており、それぞれに対応したSQL文が自動的に生成される。そのため、開発者は各DBMS特有の高度な知識が不要になる。

メンテナンスやエンハンスにも威力を発揮

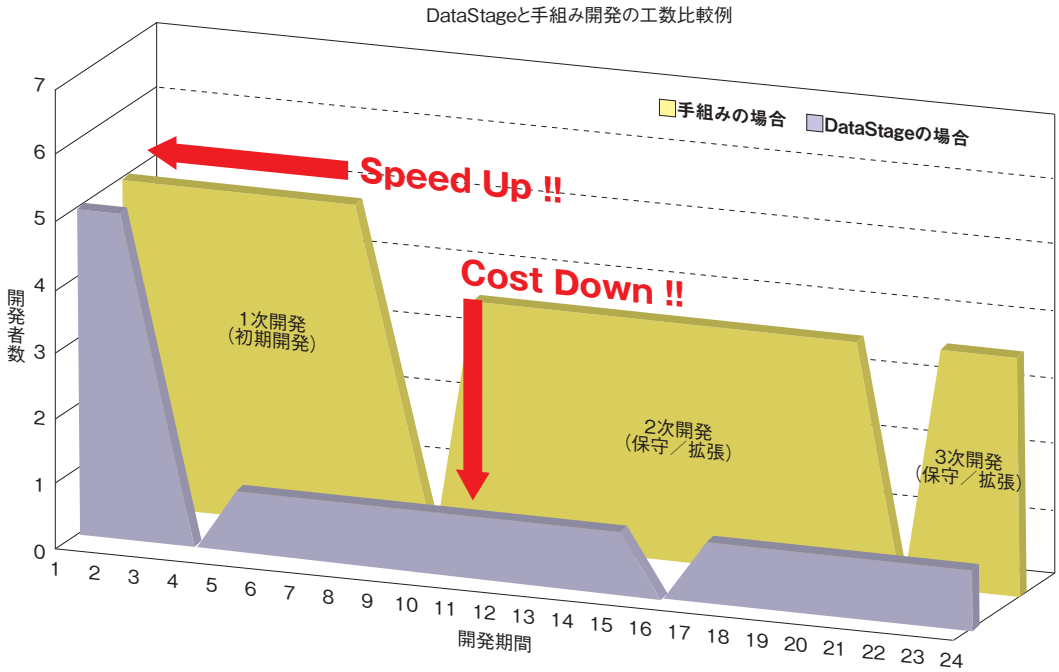
DataStageの強みが発揮されるのは、データ統合処理の実装時だけにとどまらない。実装後のメンテナンスやエンハンス時にも大きな威力を発揮する。

DataStageは、データ統合に関する情報をメタデータ化し、一元管理する。ジョブやテーブル定義、サブルーチンなどの個別要素、さらにはそれぞれの要素間の相関関係までメタデータ化し、単一のリポジトリで一元管理している。そのため、どの要素がどのジョブで使われているかといった依存関係を、即座に確認することができる。

これは、開発者にとっては実に便利な機能だといえる。例えば、データ抽出元のデータベースのある項目に変更が加えられたとしよう。手組み開発の場合であれば、プログラム中でその項目を参照している個所をすべて洗い出し、1つ1つ手作業で修正しなければならない。これは手間が掛かるだけでなく、修正のミスや漏れによって新たなバグが入り込む危険性も高い作業だ。最初にプログラムを作成した開発者と後にプログラムを修正する開発者が異なる場合は、なおさらリスクが高くなる。

その点DataStageであれば、そのデータ項目の変更が影響を及ぼす個所を、ツールがメタデータを基に自動的に検出し、一覧表示してくれる。そのため、最初に開発を担当した開発者でなくともスムーズに処理の変更作業を行うことができ、システムの保守性や拡

DataStageと手組み開発の工数比較例



張性を大幅に向上させることが可能になるのだ。

日立製作所が、実際にDataStageを導入したとある国内製造企業の事例を基に、DataStageによるノンプログラミング開発と汎用プログラミングツールによる手組み開発の生産性を比較してみたところ、データ統合プロジェクトの開発期間が6カ月から2カ月にまで短縮できたという。また、1次開発完了後の2次開発では、手組み開発の場合であれば3人の開発要員が必要だったところを、DataStageを導入したことにより1人の開発要員で対応できるようになったという。

このようにDataStageは、手組み開発に比べてデータ統合の開発期間とコストを大幅に削減できるとともに、開発完了後に定常的に発生するシステム変更や拡張にも効率よく対応できるため、使い続けるほどその高い費用対効果を実感できるツールとなっている。

294万円の小規模用ライセンスも用意

もう1つDataStageが広く選ばれている理由として、日立製作所という国内有数の大手メーカーが販売とサポートを手掛けている点が挙げられる。

DataStageの開発元ベンダーでは、ユーザーの幅広いニーズに対応するため、導入規模に合わせた複数のエディションをラインアップしている。日立製作所ではこれに加え、同社のみが提供する「DataStage Common Edition」という独自エディションを、294万円（税込み）からという導入しやすい価格帯で販売している。

このエディションを活用することにより、限られた範囲からデータ統合を始め、段階的にデータ統合の範囲を広げていくスモールスタートが可能となり、初期投資を最小限に抑えながらデータ統合に着手することができる。事実、これまでコスト面からETLツールの導入に踏み切れなかった中堅企業、あるいは学校、病院、自治体などといったユーザーが次々とDataStage Common Editionを導入し始めているという。

製品を知り尽くしているからこそのサポート品質

さらに日立製作所では、1999年から現在に至るまでDataStageの開発元ベンダーと販売および開発のパートナーシップを結んでおり、長年にわたる日本語化作業や品質テストなどで培ってきた豊富な製品知識とノウハウを保有している。今や同社では、DataStageに関するユーザーからの問い合わせの大部分を、開発元ベンダーにエスカレーションすることなく自社で対応しているという。そして、こうした強みを生かして、他社には追随できない充実したサポート体制を築き上げている。

サポートサービス契約者向けのサイトでは、DataStageに関する豊富なFAQ情報が公開されているほか、サンプルのジョブやシェルプログラムも入手できる。さらには製品のバージョンアップやエンハンスなどに関する情報も詳細に提供されている。

また、日立ソリューションサポートセンター（HSSC）の問い合わせ窓口では、先述の通り豊富に蓄積されているDataStageの製品知識をフルに生かしたユーザー対応を行っている。同社内に常時用意されているDataStageの検証環境で迅速に問題の調査に当たっているほか、開発元ベンダーとも密接に連携して技術的に高度な問題にも迅速に対応できる体制を整えているという。

パッケージ製品を導入する際には、その機能や価格のみならず、導入後のサポートサービスの品質がものをいう。これは、パッケージ製品を使ったことがある方なら、身に染みて実感されていることだろう。国内有数の大手メーカーが全面的にサポートする安心感は、そのまま自社システムに対する安心感に直結するのだ。

データ連携やデータ移行ニーズにも

以上で見てきたように、データ統合への取り組みが急務になっているさまざまな企業にとって、DataStageは現時点でのベストプラクティスの1つだといえるだろう。また、データ統合のみならず、システム間のデータ連携やデータ移行といったニーズにもDataStageを活用することができる。

こうしたニーズを抱えながらも、いまだ解決策を見いだせていない情報システム担当者を対象に、日立製作所ではDataStageの定期開催セミナー**を実施しているという。企業システムのデータ管理にまつわるさまざまな課題に、最新のETLツールがどのように役立つのか、一度体感してみてはいかがだろうか。

**DataStageセミナー開催情報 http://www.open-middle.com/hitachi/DataStage/

■価格：DataStage Common Edition　294万円～（税込み）
- サーバ側OS：Windowsのみ　• CPU数：2コアまで
- 同時接続クライアント数：1クライアント（追加費用で2クライアントも可）
- 開発機、待機機ライセンスはなし

※DataStage Standard Editionへのアップグレード（有償）により、同時接続クライアント数の3クライアント以上への増加、オプション製品（PACK for SAP R/3他）の追加、CPU数の追加、Windows以外のOSへの変更が可能です。

●製品販売元

株式会社 日立製作所

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）
TEL:0120-55-0504
http://www.hitachi.co.jp/soft/hmcc
利用時間 9:00～12:00 13:00～17:00［土・日・祝日・弊社休日を除く］

製品情報サイト
http://www.hitachi.co.jp/soft/datastage/

●製品開発元
日本アイ・ビー・エム株式会社

※本記事は、TechTarge（http://techtarget.itmedia.co.jp/）へ2009年8月に掲載されたコンテンツを再構成したものです。
http://techtarget.itmedia.co.jp/tt/news/0908/28/news01.html