

新機能ガイド Sybase® ETL 4.9

ドキュメント ID : DC00939-01-0490-01

改訂 : 2009 年 9 月

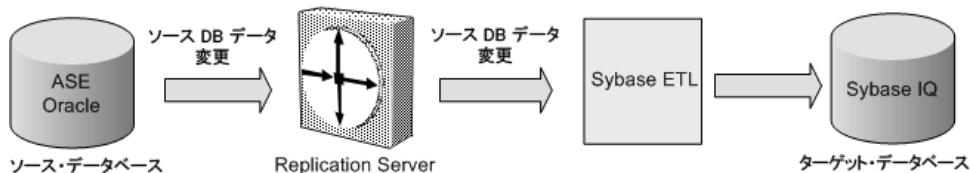
このマニュアルでは、Sybase® ETL 4.9 の新しい機能について説明します。

トピック	ページ
ETL と Sybase Replication Server を使用したインクリメンタル・ロードの機能強化	1
ETL スケジューラのサポート	3
実行時イベントに対するアラートの設定	4
Transformation コンポーネントの機能強化	6
SQL Executor コンポーネントの追加	6
Copy Splitter コンポーネントの追加	7
Data Splitter JavaScript コンポーネントの機能強化	7
トランザクション機能の強化	8
接続共有のサポート	10
DB Staging コンポーネントの機能強化	11
Data Sink コンポーネントの機能強化	12
複数のプロジェクトの同時実行	12
サポートされているリポジトリ	13

ETL と Sybase Replication Server を使用したインクリメンタル・ロードの機能強化

ETL と Sybase Replication Server® を使用して、Replication Server が Sybase Adaptive Server® Enterprise (ASE) データベースおよび Oracle データベースから取得したデータ変更を Sybase IQ にロードできるようになりました。このソリューションでは、Sybase ETL によるデータ変換と同様にソース・テーブルとターゲット・テーブル間の異なるスキーマもサポートしています。

図 1: Sybase Replication Server を使用したインクリメンタル・ロード



Sybase ETL 4.9 には、CDC Provider Sybase Replication Server という新しいコンポーネントがあります。このコンポーネントでは、Sybase Replication Server ベースのインクリメンタル・ロード・ソリューションがサポートされているため、パフォーマンスと利便性が向上しています。また、これ以外の場合での使用もサポートされています。

4.9 より前のバージョンでは、DB Data Provider Index Load コンポーネントでインクリメンタル・ロードがサポートされていましたが、昇順インデックス属性値に基づいて変更データを取得することしかできませんでした。

加えて、すべてのデータ変更（ソース・データベースの挿入、更新、削除アクティビティなど）をソース・テーブル・スキーマへの依存や変更することなく取得できるようになりました。そのため、ソース・データベースのパフォーマンス低下もありません。ETL のプロジェクト・レベルのトランザクション機能を有効にして、一貫性を確保できます。

ソース・データベースのすべてのデータ変更が Sybase Replication Server によって取得されます。ETL は、Replication Server からデータ変更を受け取り、変換して Sybase IQ にロードします。

CDC Provider Sybase Replication Server コンポーネントを使用すると、Sybase Replication Server の設定プロセスを自動化し、複写定義、複写の接続、ファンクション文字列、複写のサブスクリプションを作成および削除したり、ソース・データベースのテーブルを複写するようマーク付けしたりできます。これらは、すべての ETL タスクに対して実行できます。

ETL のインクリメンタル・ロード・ソリューションを使用すると、次の処理を実行できます。

- Oracle データベースおよび Sybase Adaptive Server Enterprise データベースからデータ変更を抽出する。
- ターゲットの Sybase IQ データベースまたはファイルにデータをロードする。
- 異なるテーブル・スキーマのターゲット・テーブルにソース・テーブルのデータ変更を転送する。

- Replication Server から受け取った変更データの各セットに対してさまざまな変換規則を適用する。
- 古い値を取得して変換規則を適用し、古いデータと変更データを比較する。

『Sybase ETL 4.9 ユーザーズ・ガイド』の「第5章 コンポーネント」を参照してください。

ETL スケジューラのサポート

Sybase ETL 4.9 には独自の ETL スケジューラが用意されており、すべての使用可能なプラットフォームに対応するようにスケジューラの機能が拡張されています。ETL スケジューラでは、ETL Development GUI で選択したエンジンを使用して、スケジュール・タスクを作成、一覧表示、編集、削除、実行、終了できます。

以前のバージョンの Sybase ETL Development では、Windows のタスク・スケジューラを直接使用していました。Sybase ETL Development 4.8 以前のバージョンからアップグレードするときにスケジュール・タスクをマイグレートする場合、Runtime Manager を使用して既存のタスクをインポートする必要があります。

また、ETL スケジューラでは、タスクの複数のインスタンスを同時に実行できます。すでにスケジュール・タスクが実行されている場合、次のオプションのいずれかを指定できます。

- [Execute new task concurrently] – 現在の処理と並行して新しいスケジュールの実行を開始します。
- [Execute new task sequentially] – 現在の処理が完了するまで次のスケジュールの実行を開始せずに待機します。
- [Do not execute new task] – 現在のタスクの処理を続行し、新しいスケジュールの実行の開始要求を無視します。
- [Cancel the running task before executing new task] – 現在の処理を終了して新しいスケジュールの実行をすぐに開始します。

ETL スケジューラは、ファイルベースの設定に従います。すべての設定は、各グリッド・エンジンと共にスタンドアロンのフラット・ファイルに保存されます。ファイル・ロックを使用すると、ホスト・マシンごとに1つの ETL スケジュール・サービスのみが開始されます。設定ファイルを手動で変更しないことをおすすめします。スケジューラの設定には、ETL Development GUI の ETL Runtime Manager を使用するようにしてください。

『Sybase ETL 4.9 ユーザーズ・ガイド』の「第4章 高度な概念とツール」を参照してください。

実行時イベントに対するアラートの設定

Sybase ETL 4.9 では、実行時イベント（プロジェクトまたはジョブの開始、完了、エラーなど）が発生したときのアラートを設定できます。次のいずれかのイベントが発生したときに電子メール・アラートを送信するように ETL を設定できます。

注意 ETL イベント通知のサブスクリプションを作成する電子メール・アドレスは複数指定できます。

- [Jobs] – ジョブの開始、正常終了、またはエラー終了時に通知します。
- [Job Start] – ジョブの開始時に発生します。ジョブ名、ジョブが実行されるグリッド・エンジン、ジョブの開始時刻が含まれます。
- [Job Finish] – ジョブの正常終了時に発生します。ジョブ名、ジョブが実行されていたグリッド・エンジン、ジョブの開始時刻と終了時刻、すべてのプロジェクトのリストが含まれます。
- [Job Error] – ジョブのエラー終了時に発生します。ジョブ名、グリッド・エンジン名、開始時刻、エラー発生時刻、エラー・メッセージなどの詳細情報が含まれます。
- [Projects] – プロジェクトの開始、正常終了、またはエラー終了時に発生します。
- [Project Start] – プロジェクトの開始時に発生します。プロジェクト名、プロジェクトが開始されたグリッド・エンジン、プロジェクトの開始時刻などの詳細情報が含まれます。

- **[Project Finish]** – プロジェクトの正常終了時に発生します。プロジェクト名、プロジェクトが実行されていたグリッド・エンジン名、プロジェクトの開始時刻と終了時刻、正常にロードされたロー数などの情報が含まれます。
- **[Project Error]** – プロジェクトのエラー終了時に発生します。メールには、プロジェクト名、プロジェクトが実行されていたグリッド・エンジン名、プロジェクトの開始時刻とエラー終了時刻、正常にロードされたロー数、エラー・メッセージなどの詳細情報が含まれます。

注意 同じグリッド・エンジンで実行されている複数のグリッド・ノード・プロセスが、同じアラート・サービス設定ファイルとアラート履歴ログ・ファイルを共有します。

ETL Development GUI の **Alert Manager** を使用して、アラート定義の作成、編集、削除を行います。プロジェクトのシミュレーションやジョブ、またはプロジェクトの実行に対応する任意のイベント・タイプのアラートは複数作成できます。ただし、アラート名は一意にする必要があります。

File Log Inspector や **ETL Web** インタフェースを使用してイベント・アラート履歴を表示できます。また、アラートのマッピング時にフィルタ条件を設定することもできます。たとえば、フィルタを定義して、特定のプロジェクト名でのみアラートがトリガされているかどうかを確認できます。

また、**Sybase ETL 4.9** では、すべてのアラート履歴を個別のログ・ファイルに保存できるため、トラブルシューティングに役立てることができます。

電子メール・アラート通知は 1 台のマシンで設定することをおすすめします。アラート設定ファイルはネットワーク上で稼働している別のすべてのグリッド・エンジンに手動でコピーできます。

『**Sybase ETL 4.9 ユーザーズ・ガイド**』の「第 4 章 高度な概念とツール」を参照してください。

Transformation コンポーネントの機能強化

この項では、Transformation コンポーネントの機能強化について説明します。

SQL Executor コンポーネントの追加

ETL 4.9 には、SQL Executor という新しい Transformation コンポーネントがあります。このコンポーネントを使用して、データベース・サーバに対してカスタム SQL 文や複数の文を実行できます。SQL Executor は、入力ポートや出力ポートのないスタンドアロン・コンポーネントで、他のコンポーネントとは別のプロジェクトに配置したり、1 つまたは複数の SQL Executor コンポーネントのあるプロジェクトに配置したりできます。

また、SQL Executor コンポーネントでは、1 つのトランザクションで複数の SQL 文を使用して複数の入力データ・ストリームから Sybase IQ データベースにデータをロードできます。こうすることで、すべての ETL ユーザが柔軟性の高い操作を実行できるようになります。

SQL Executor コンポーネントを使用すると、次の処理を実行できます。

- 他の各プロジェクトで同じデータベースのデータを読み込んだり、書込んだりする。
- SQL 文を使用してソース・テーブルのデータをロードし、IQ でサポートされているテキスト・ファイル形式にする。
- Load Table コマンドを使用して、テキスト・ファイルからターゲット IQ データベースに 1 つのトランザクションでデータをロードする。

SQL Executor コンポーネントには、pre-SQL 文および post-SQL 文のパラメータがあります。このコンポーネントによってターゲット・データベースで SQL スクリプトがステップ実行され、SQL スクリプトの構文が正しいかどうかを確認されます。また、スクリプトの実行が正常に実行された場合、または失敗した場合に実行する、異なるカスタム SQL スクリプトを定義することもできます。

『Sybase ETL 4.9 ユーザーズ・ガイド』の「第 5 章 コンポーネント」を参照してください。

Copy Splitter コンポーネントの追加

Sybase ETL 4.9 には、Copy Splitter という新しいコンポーネントがあります。このコンポーネントは、入力データを各出力ポートに無条件にコピーします。Copy Splitter コンポーネントを使用すると、出力ポート条件を評価することなく各出力ポートにデータを書き込めます。デフォルトでは、このコンポーネントには2つの出力ポートがあります。

Copy Splitter コンポーネントを使用すると、JavaScript を起動する必要がなくなり、条件評価のコストが削減され、パフォーマンスが向上します。

『Sybase ETL 4.9 ユーザーズ・ガイド』の「第5章 コンポーネント」を参照してください。

Data Splitter JavaScript コンポーネントの機能強化

Data Splitter JavaScript コンポーネントは、相互排他的な出力ポートの条件の評価や定義ができるように機能強化されました。この機能強化によって、メモリ割り付けのシステム・コールが不要になるだけでなく効率的な方法でメモリを管理できるようになり、パフォーマンスが大幅に向上しています。

Sybase ETL 4.9 より前のバージョンでは、Data Splitter JavaScript コンポーネントは、すべての出力ポートの条件に対して入力レコードを評価することしかできなかったため、潜在的なボトルネックとメモリ妨害を引き起こしていました。また、他の効率的な論理構成を使用できる柔軟性がありませんでした。

現在のバージョンでは、Data Splitter JavaScript コンポーネントを使用して相互排他的な出力ポートの条件を定義できるため、条件の評価が大幅に削減されています。各入力レコードは、0 または 1 の出力ポート式に一致します。条件に一致する最初の出力ポートで入力レコードが受信されます。後続のポート条件は評価されません。式に一致しない場合は次のポートの式が評価されます。これは、一致してそのポートのデータがコミットされるまで継続されます。

ある条件に偏っているデータの傾向がある場合、カスタム評価シーケンスを指定して、高確率の条件をリストの先頭に配置できます。これは、不要な条件の評価を回避できるため、出力ポート数が多い場合に特に有効です。

『Sybase ETL 4.9 ユーザーズ・ガイド』の「第5章 コンポーネント」を参照してください。

トランザクション機能の強化

Sybase ETL 4.9 では、書き込み操作 (ジョブおよびプロジェクトの実行またはシミュレーション内の複数のコンポーネントが対象) の終了時におけるデータのコミットおよびロールバックのサポートが拡張されています。正常に実行されると、書き込み処理の終了時にデータがコミットされます。正常に実行されなかった場合は、ロールバックされます。また、Multi-Project および Synchronizer Job コンポーネントでは、処理が終了したらすぐに中間作業をコミットできるようになっており、包含プロジェクトの管理性が向上しています。Multi-Project コンポーネントのダウンストリーム・プロジェクトが失敗しても、ダウンストリーム・プロジェクトのみがロールバックされ、すでにコミットされたプロジェクトには影響しません。

4.9 より前のバージョンでは、複数のコンポーネントまたはプロジェクトにまたがるトランザクションは使用できませんでした。また、ターゲット・コンポーネントのすべての変更セットをロールバックすることもできませんでした。これは、コンポーネント・レベルとバッチ・レベルでのみトランザクション機能がサポートされていたためです。

現在のバージョンでは、複数のプロジェクトにまたがってトランザクション機能を使用できるようになりました。そのため、データベースの更新またはロードをすべて実行するかまったく実行しないという方式を実行または適用できます。また、トランザクション機能をサポートするすべてのコンポーネントに対してコミット機能およびロールバック機能が有効になるように、ジョブまたはプロジェクトのプロパティを設定することもできます。[Propagate Rollback] オプションを選択していないときは、1 つ以上のコンポーネントにエラーが発生した場合、プロジェクトまたはジョブは、正常に実行されたコンポーネントに対してはトランザクション・ロールバックを適用しません。エラーが発生したコンポーネントごとに、独自のトランザクションがロールバックされます。

また、プロジェクトのシミュレーション時にトランザクション・コンポーネントで実行された操作をコミットするかロールバックするかを選択することもできます。実行の終了時に成功条件または失敗条件のいずれかをシミュレートして、プロジェクトを初期状態にリセットできます。すべてのポート・バッファがクリアされ、テンポラリ・テーブルが解放され、データベース接続とテンポラリ・ファイルが閉じられます。ジョブまたはプロジェクト内のすべてのコンポーネントに対して新しいトランザクション機能を無効にすることもできます。

改善されたトランザクション機能では、次の処理を実行できます。

- ソース・データベースから Sybase IQ にデータをインクリメンタル・ロードする。成功すると、Sybase IQ に新しいローが挿入され、変更対象のローが更新されます。実行に失敗した場合は、IQ のターゲット・テーブルおよびステージング・テーブルに対するすべての変更が、1 つのトランザクションの一部としてロールバックされます。
- 複数のテーブルを Sybase IQ にロードする。成功すると、ソース・データベースのデータの新しい情報で IQ のすべてのテーブルがロードされます。実行に失敗した場合は、すべてのターゲット・テーブルに対する変更がロールバックされます。
- 1 つのトランザクションでターゲット IQ データベースから複数のローを削除する。
- Load Table を使用して 1 つのトランザクションの一部として IQ をロードする。
- プロジェクトの実行の終了時にソース・コンポーネントの前処理 SQL および後処理 SQL のデータをターゲット・テーブルからロールバックする。
- Text Data Sink コンポーネントを使用してローカルのテキスト・ファイルにデータを書き込むときに DBMS 形式のトランザクションをシミュレートする。
- マルチプロジェクト・ジョブのすべてのプロジェクトを 1 つ分散トランザクション内でコミットする。
- トランザクション・プロジェクトをシミュレートし、シミュレーション・プロセスの任意の段階でコミットまたはロールバックする。
- トランザクション・ジョブを停止して、ジョブ内のトランザクション・リソースに書き込まれたデータを確実にロールバックする。
- マイグレーション・ウィザードを使用してジョブを作成し、すべてのプロジェクトおよびターゲット DB コンポーネントに対してトランザクション機能を定義する。

プロジェクトの実行に失敗した場合、pre-SQL 文や post-SQL 文 (データベース・トランザクションの一部) などのすべてのデータベースの変更は、プロジェクトの他のすべてのトランザクション・コンポーネントと共にロールバックされます。また、プロジェクトのシミュレーション中にコミットされていないタスクを手動でロールバックすることもできます。

ただし、データ定義言語 (DDL) SQL コマンド (`create`、`drop`、`alter`) を実行してから、ロールバックに必要な `pre-SQL` コマンドを実行する必要があります。これは、DDL SQL コマンドを実行すると、データベース・サーバで暗黙的にトランザクションがコミットされるためです。たとえば、データをテーブルに挿入して `CREATE TABLE` コマンドを実行すると、挿入されたデータはデータベースによって暗黙的にコミットされるためロールバックできません。

『Sybase ETL 4.9 ユーザーズ・ガイド』の「第 5 章 コンポーネント」を参照してください。

接続共有のサポート

ETL 4.9 では、次の Destination コンポーネントの接続共有がすべてサポートされています。

- DB Bulk Load Sybase IQ
- DB Data Sink Insert
- DB Data Sink Update
- DB Data Sink Delete

この新しい機能では、複数のデータベース・コンポーネントを使用して同じソース・テーブルまたはターゲット・テーブルにアクセスできます。テーブル・ロックのエラーは発生しません。この機能を使用すると、コンポーネントは、データベースに接続する場合に、接続パラメータとデータベース・パラメータの定義が同一である別のターゲット・コンポーネントとの間で単一の接続を共有できます。4.9 より前のバージョンでは、複数のコンポーネントで同じテーブルに同時にアクセスしようとする、プロジェクトの実行に失敗していました。現在のバージョンでは、新しい接続共有機能が備わっているため、複数のソース・テーブルからデータを抽出したり、データを 1 つのターゲット・テーブルに同時にアップロードしたりできます。また、Sybase IQ データベースへのバルクロードを実行できるため、パフォーマンスも向上しています。

接続共有機能は、同じプロジェクト内のコンポーネント間でのみ使用できます。同じジョブ内でもプロジェクトが異なるコンポーネント間では、接続は共有できません。また、既存の ETL 複数ライタ機能と接続共有機能は併用できません。

注意 同じデータベース・インタフェースとログイン情報が定義されていてもデータベース・オプションが異なるコンポーネント間では、接続は共有できません。プロジェクトが実行されるかシミュレートされた場合にエラーが発生します。

『Sybase ETL 4.9 ユーザーズ・ガイド』の「第 5 章 コンポーネント」を参照してください。

DB Staging コンポーネントの機能強化

ステージング・プラットフォームとして Sybase IQ を使用すると、Sybase ETL 4.9 では Load Table 文を使用してステージング・テーブルにデータをバルクロードできるため、パフォーマンスが向上します。以前のバージョンでは、ターゲット・データベースに受信データをロードするには、DB Staging コンポーネントで INSERT 文を使用してローベースの操作を実行することしかできませんでした。

DB Staging コンポーネントの機能強化によって、次の処理を実行できるようになりました。

- 実行時に、入力ポート構造に基づいて自動的にテーブルを作成する。プロジェクトで処理が終了したら、実行時に作成されたテーブルを削除できます。
- 出力が作成される前に、追加の SQL スクリプトを実行してステージング・テーブルを変更または更新する。入力ポートに関連するテーブルに変換フローのすべてのデータがロードされてからクエリの結果セットが取得されるまでの間に、SQL 変換スクリプトを実行できるようになりました。
- 正常に実行されると、書き込み操作の終了時にデータがコミットされ、正常に実行されなかった場合はロールバックされるようにする。

Sybase IQ 15.0 ステージング・データベースで、インタフェースとして ODBC を使用するクライアント側ロードが有効になっている場合は、[Load Stage Path] を指定する必要はありません。IQ サーバはデフォルトの LOAD TABLE 文を使用して、リモート・ホスト・マシンにあるファイルから Sybase IQ テーブルにレコードを自動的に追加します。

『Sybase ETL 4.9 ユーザーズ・ガイド』の「第 5 章 コンポーネント」を参照してください。

Data Sink コンポーネントの機能強化

DB Data Sink Insert、DB Data Sink Update、DB Data Sink Delete コンポーネントが機能強化され、Load Table 文を使用して Sybase IQ データベースのターゲット・テーブルにデータをバルクロードできるようになりました。

『Sybase ETL 4.9 ユーザーズ・ガイド』の「第 5 章 コンポーネント」を参照してください。

複数のプロジェクトの同時実行

Sybase ETL 4.9 では、プロジェクトの並列実行がサポートされており、1 つの ETL サーバを使用して複数のプロジェクトをスケジュールする場合は特にパフォーマンスが向上します。以前のバージョンでは、グリッド・エンジンおよびジョブごとに 1 つのプロジェクトしか実行できなかったため、リソースの使用率が低くなっていました。現在のバージョンでは、1 つまたは複数のグリッド・エンジンで複数の ETL プロジェクトを同時に実行できます。また、複数の ETL ユーザが同じリモート・グリッド・エンジンでプロジェクトを同時に実行することもできます。

同時に実行する最大プロジェクト数を指定するには、インストール・フォルダの *etc* サブディレクトリにある *Default.ini* ファイルを編集する必要があります。デフォルトでは、同時に実行可能な最大プロジェクト数は 10 です。ただし、プロジェクトの最大値 (MAXPROJECTS) を 0 または負数に設定すると、特定のグリッド・エンジン上で同時に実行可能なプロジェクト数は無制限となります。

『Sybase ETL 4.9 ユーザーズ・ガイド』の「第 6 章 Sybase ETL サーバ」を参照してください。

サポートされているリポジトリ

Sybase ETL 4.9 でサポートされているリポジトリ・データベースは SQL Anywhere® 11 のみです。ETL の以前のバージョンで他のリポジトリを使用している場合、既存の ETL リポジトリを SQL Anywhere 11 にマイグレートする必要があります。

『Sybase ETL インストール・ガイド』の「第 4 章 アップグレード」の「既存のリポジトリから SQL Anywhere へのマイグレート」を参照してください。

