

# ETI HIGH PERFORMANCE CONNECTIVITY FOR MICROSOFT SQL SERVER 2005 AND SSIS

---

PROVIDING READY-TO-RUN CONNECTORS  
FOR HETEROGENEOUS DATA SYSTEMS

By Dan Linstedt

Daniel E. Linstedt is an internationally known expert in data warehousing, business intelligence, analytics, very large data warehousing, OLTP, performance and tuning. He has been the lead technical architect on enterprise-wide data warehouse projects and refinements for many Fortune 500 companies. Linstedt is an instructor for The Data Warehousing Institute (TDWI) and a featured speaker at industry events such as Oracle Developer Tools User Group, Informatica World, Wilshire Conferences, Teradata Partners and Rocky Mountain Oracle User Group. Dan can be reached at [danl@DanLinstedt.com](mailto:danl@DanLinstedt.com).

---

## INTRODUCTION

---

Businesses are demanding higher performance, heterogeneous connectivity and single-point management solutions to their data integration problems. They are continuously faced with ever growing data sets into and beyond terabyte levels and these same businesses typically have legacy systems that are difficult to access and nearly impossible to integrate.

To meet today's challenges of ever growing data sets, rapid data integration and legacy access, it is important to have access to a single control interface (for data integration) with high speed inter-connects and ease of use (visual flow formatting). Microsoft is answering the challenge with SQL Server 2005 and SSIS (SQL Server Integration Services) providing a standards-based approach to connectivity. By supporting OLE DB, ODBC and ADO.NET, Microsoft enables customers to plug in the appropriate adapters for their environment.

ETI then boosts Microsoft performance by providing easy-to-use, plug-and-play high speed connectivity and visual management for the SSIS ETL (extract transform and load) design studio. These components are essential to consolidation, centralization and lowering the total cost of ownership for moving and transforming data. The harmony created by these solutions is hard to beat, given that SSIS is included in the SQL Server environment.

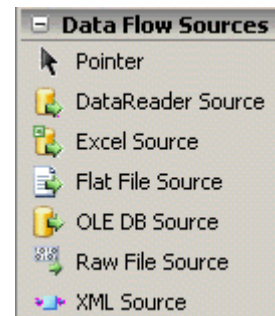
ETI offers unsurpassed high speed connectivity to legacy systems, 3<sup>rd</sup> party systems and non-standard file systems. Their products include both ready-made and customized and are capable of running in native format on the source and target systems as well as plug-and-play from the centralized SSIS management console.

---

## SSIS AND OUT-OF-THE-BOX CONNECTIVITY

---

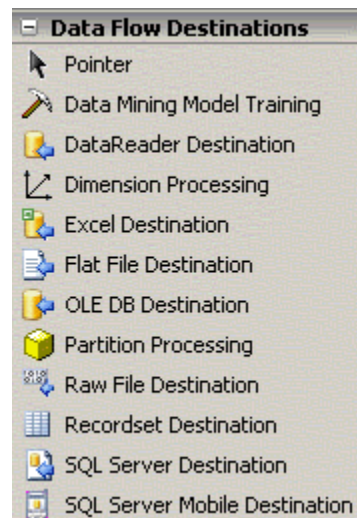
SSIS installs with standard connectivity out of the box. Microsoft provides the connectors for SQL Server sources and targets, plus standards-based connectors to other data sources (Oracle and DB2 OLE DB). The out-of-the-box Data Flow Sources include:



- ✚ **DataReader Source** extracts data from a relational database by using a .NET provider.
- ✚ **Excel Source** extracts data from any Excel spreadsheet or tab on the Excel spreadsheet.
- ✚ **Flat File Source** reads data from a delimited or fixed width data set.
- ✚ **OLE DB Source** reads data from any OLE (Object-linking-and-embedding) supported source.
- ✚ **Raw File Source** reads data from any binary, text or random unstructured data file.
- ✚ **XML Source** reads data from any file specified in XML format.

These connectors are easily accessible and configurable and reside on Microsoft platforms. (They do not reside on UNIX, Linux, mainframe or midrange systems.)

The out-of-the-box Data Flow Destinations include:



- ✚ **Data Mining Model Training** passes data to Analysis Services for a data mining model structure. It trains the data mining models during this process of loading.
- ✚ **DataReader Destination** writes data to a .NET process or component.
- ✚ **Dimension Processing** writes data to a dimension-type table and includes the options for add, full and update.
- ✚ **Excel Destination** writes data to Excel-based spreadsheets or spreadsheet tabs.
- ✚ **Flat File Destination** writes delimited or fixed width data files.
- ✚ **OLE DB Destination** writes data to any OLE-supported target.
- ✚ **Partition Processing** writes data to any Analysis Services cube and partitioned or aggregated target.
- ✚ **Raw File Destination** writes data to any binary, text or random unstructured data file.
- ✚ **Recordset Destination** creates and populates an in-memory ADO recordset. The shape of the recordset is defined by the input to the Recordset Destination at design time.
- ✚ **SQL Server Destination** connects to a SQL Server database and bulk loads data into SQL Server tables and views. This destination offers the same high-speed insertion of data into SQL Server that the Bulk Insert task provides; however, by using the SQL Server destination, a package can apply transformations to column data before the data is loaded into SQL Server.
- ✚ **SQL Server Mobile Destination** writes data to SQL Server Mobile databases.

---

## ACCESSING NON-MICROSOFT PLATFORM DATA SETS WITH ETI

---

What's missing from this picture? Most organizations have a need to work with data on many different platforms in native format and that's where ETI comes in. It's not just a matter of accessing (reading/writing) the data on alternate platforms, but also executing **native transformations** on **native platforms** before the data ever leaves the platform. These standard systems include: Teradata, iSeries (AS/400), VSAM, IMS, COBOL, current and legacy versions of Oracle, DB2, Informix and Sybase. By executing transformations, de-duplication, merging and aggregation **on** the native platform using ETI connectors, network traffic can be reduced and performance can be dramatically increased.

Many corporations have source data on legacy systems where it has existed for years. Access to these systems is critical for a heterogeneous environment. Consolidation and a single, unified master platform is a nice pipe-dream for many of these organizations; however, this dream rarely becomes a reality due to the nature and complexity of native data sets and the operational applications that have been written on top of them.

Much more important is the accessibility and interchangeability of these data sets. Being able to effectively move the data from one platform to another and transform, process and integrate it from a single consolidated console is the power play for the organization and an ETI strength. Furthermore, most organizations have cube-based analysis services and most business users understand and use Excel for data visualization and drill-in graphing. The pervasive nature of the front-end analysis tools drives the need to easily access and update Microsoft Analysis Services data sets with legacy data, sometimes in near-real-time.

Accessibility to non-Microsoft platform data sets is critical and ETI offers solutions that are plug-and-play high performance connectors (HPCs) or ETI Built-to-Order™ (BTO) connectors. BTO connectors are custom-made software that integrates mainframe, legacy and proprietary applications your current business intelligence, data warehousing and SOA-based infrastructures. BTO integration meets the exact needs of the customer, with no platform to buy and no hand coding required. While the option of *build-it-yourself* exists, it isn't an attractive alternative. ETI's connectors are constructed on a series of base-libraries written with thousands of hours of engineering and best practices. These native base-libraries are configured, compiled and tested for specific legacy environments. They also provide the Microsoft interface to work with SSIS in a plug-and-play manner. ETI's integration products can be tuned specifically to meet the needs of legacy platforms within an organization.

### Connecting To Teradata

Teradata is a challenging platform to connect to. The API is incredibly difficult for most programmers to write code against.

An entire set of knowledge regarding semaphores, thread-safe and memory buffering operations must be understood before a successful connection to Teradata can be encoded.

ETI solves this problem by having an engineering team with years of experience coding against the Teradata APIs. These engineers keep up with the latest releases and performance enhancements that Teradata publishes.

Making Teradata data accessible to Microsoft OLAP cubes is an extremely powerful solution: high-speed parallel SQL operations, pulling data through in-database data mining algorithms and then writing it directly into OLAP cubes for further what-if analysis – the business user becomes empowered with huge volumes of data accessible at their fingertips.

Furthermore, the bi-directional capability of ETI connectors allows the business users to generate what-if analysis and write the information back to Teradata!

Organizations today realize that they cannot simply “sunset” their legacy applications and platforms. Instead, they understand the value of unlocking the raw data sets housed within the legacy systems and integrating these systems seamlessly with the target environment. High speed, parallelism, native access and customized parameters are all benefits of ETI’s products. Customizations include: change data capture, filtering, aggregation, joining, transformation, encryption/decryption, compression and complex regular expression parsing. All customizations are executed in native code on the source system which keeps the network traffic down when reading or writing to legacy systems. Furthermore, all connectors provided by ETI are bi-directional – providing seamless, high speed read/write integration.

#### ETI BTO Connectivity

ETI’s BTO integration solutions enable rapid access to heterogeneous data, including:







- Databases
- Mainframes
- Applications
- File Systems
- Messages/Transactions
- Unstructured Data
- Web Services
- BI/Reporting
- Industry Protocols

#### ETI Platform Support

- Mainframe z/OS
- Mainframe VSE
- iSeries - OS/400
- Windows
- Linux Red Hat
- Linux SUSE
- IBM AIX
- Sun Solaris
- HP-UX
- Tandem
- DEC UNIX
- DEC VMS

Unlocking legacy data in a seamlessly integrated environment is vital to the success of regulatory compliance, web-services (SOA) and the dynamic build-out of BI solutions like Master Data Management, Data Warehousing and Business Intelligence. By using ETI connectors, you can automatically integrate data from Teradata and ISAM/VSAM formats directly into Microsoft OLAP cubes or Excel spreadsheets. These can be automated, controlled and plugged into the SSIS workflow interface for ease of use, scheduling and visual maintenance.

If an internally hand-coded solution is still being considered, then consider the following requirements for building from scratch (in house):

-  Coding knowledge and expertise on the legacy system
-  Knowledge of data sets, file system layouts and accessibility
-  Knowledge of compression, encryption, transformation and data movement
-  Understanding of compilation technology
-  Debugging procedures and best practices
-  Previous work with SQL Server 2005 and SSIS API

Even with all of these elements in place the resulting build becomes a non-standard, one-off custom solution *for a single platform*. Most likely it would not use cross-platform libraries, nor would it be sustainable in the long term. It becomes a high cost maintenance component and when the coder leaves the organization, the knowledge of the “custom coded connect” leaves too. It is a much better (and cheaper) value proposition to outsource to ETI for a custom-built connector which mitigates all of the above risks.

The complexity of internal coding of integrations like this (particularly with enormous data sets) increases dramatically with the functionality required by the front-end system such as SSIS. When coding complexity increases, cost increases. ETI HPCs and BTO integrations utilize best of breed, standardized libraries where much of this functionality has already been built, tested and applied with the latest patches to the compilers that live within the legacy systems.

---

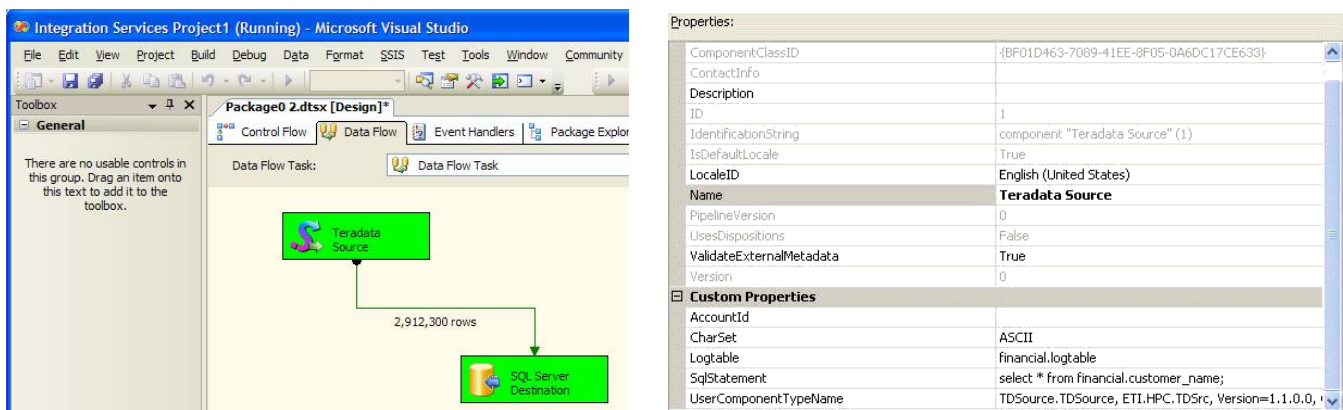
## MICROSOFT PARTNER CONNECTIVITY: ETI CONNECTORS FOR SQL SERVER

---

As a Microsoft managed partner, ETI receives architectural support to bring many of these connectors to the market. Denise Draper, Microsoft Unit Manager for SSIS said, “We know that customers have an almost endless array of different systems they need to integrate. We don’t believe that any one vendor could serve all those needs, so Microsoft’s strategy has been to make SSIS an extensible platform and to rely on our partners, such as ETI, to provide the individual components, whether they be pre-packaged or custom built, to complete the solution.”

ETI has built standard (out of the box) high performance connectors (HPCs) for Teradata, DB2, Oracle (both current and classic versions of Oracle) and iSeries (AS/400). These are all native, high-speed connectors that run on the SSIS system and access the data using source system libraries and/or utilities.

Below are two screenshots of the ETI Teradata connector and its corresponding configuration parameters.



How do these connectors work? They are provided as libraries and compiled on the native platform on which they will be installed. The SSIS (Microsoft components) are then plugged in to the Microsoft SQL Server environment for use with the SSIS GUI. Any client machine that is running SSIS can then easily configure, access and use the source connectors. The request for data, the transformations, SQL statements and CDC commands are all sent (at run time) over the network to the connector. The connector then executes native file access or executes the SQL (in the case of relational databases) and performs the data commands. Be it read or write, the instructions are executed natively on the source or target platform.

To install the connectors, simply insert the CD on the Microsoft SQL Server SSIS platform (the components must be installed at the server level in order to run the jobs) and follow the instructions provided by the installer. Once the installation has completed, click the “Tools,” then “Choose Toolbox Items” menu option in the SQL Server Business Intelligence Development Studio, click on the “SSIS Data Flow Items” tab and choose the “High Performance Connector” from the list to make it available from the “Data Flow” toolbox.

---

## CONCLUSION AND SUMMARY

---

Projects that combine ETI high performance connectivity and Microsoft SSIS benefit from the best of all worlds solution set. Access through a common platform and visualization (IDE), bi-directional high speed interconnectivity and the use of heterogeneous environments provide tremendous power to both the business users and the data integration knowledge workers. By purchasing the ETI connectivity solutions and employing the visual execution and management of Microsoft SSIS, you multiply the power and ability to complete projects on-time and within-budget.

The option to buy either ready-made HPCs or ETI Built-to-Order™ integration allows flexibility in the delivery model that is easily extended when the project team is ready to access specific, hard to reach data sets. There is a strong play in the market for heterogeneous, high speed connectivity and ETI has the right solution at the right time. ETI's high performance connectivity solutions allow the large and growing SQL Server install base to leverage the full power of SSIS (ETL/Data Integration) while empowering knowledge workers with heterogeneous information.

ETI, Evolutionary Technologies International, the ETI logo, MetaStore, SuccessFirst, AnswerLink, Dialogue Coach, ETI Solution, ETI Data Profiler, ETI Data Cleanser, ETI Change Data Capture, ETI Data Monitor, ETI Built-to-Order, ETI Integration Center and ETI iSpec are trademarks or registered trademarks of Evolutionary Technologies International, Inc. Software and/or information encodings included in ETI Solution are subject to the following U.S. Patent: 6,523,172. All other company and product names may be trademarks or registered trademarks of their respective owners. © 2007 Evolutionary Technologies International, Inc.