Mark Roush
Business Intelligence Specialist

ETL vs ELT:
Don't worry, Oracle's got them both!

A debate is taking center stage over two similar, yet very different, data integration methods: ETL (Extract, transform and load), the crafty veteran of the market, versus ELT (Extract, load and transform) the up-and-coming rookie. So what are the differences? Why should you choose one over the other? And, more importantly, does Oracle provide tools to perform each task?

The most obvious difference between the two is the acronym discrepancy (who knew one letter could mean so much). ETL is an approach that relies primarily on a separate, proprietary engine being deployed between the source and target systems. Thus, the "transformation" takes place between source and target while the data is still on the move. In other words, data is extracted from the source, moved to this intermediate platform where transformation logic (business rules, data cleansing, etc.) takes place, and finally, the data is loaded in the target system. ELT, on the other hand, erases the need for this separate platform, and performs all of the complex transformation logic on the target system itself.

With the obvious advantages of ELT, it is important to know that ETL still out-shines ELT when:

• Volumes are medium to low sized.

• The time-line for loading the data is not super critical.

• The source systems provide data slower than 60,000 rows per second - or the networks are restricting data flow.

• The target system won't perform at more than 50,000 or 60,000 rows per second, even if it's been tuned.

• Code and reference tables are needed to populate default values or provide simple data replacement through key matching.

• Small micro-bursts of data are arriving in real-time, and transformation needs to be in-line between systems, or between queue's[1].

So which method/tool should you choose?

Well the short answer is that it depends. Your choice of tool depends on:

• The amount of data you're pushing from source to target.

• Where you can afford to put your system load.

• Where you can place your resources (employee, money, etc.).

• About twenty to thirty other variables that only you have the answer for.

With either choice, Oracle has you covered. Oracle Warehouse Builder (OWB) is Oracle's comprehensive tool for ETL (extract, transform and load), fully integrated relational and dimensional modeling, data quality, data auditing, and full life cycle management of data and metadata[3]. Oracle Data Integrator is a comprehensive data integration platform that covers all data integration requirements — from high-volume, high-performance batches, to event-driven, trickle-feed integration processes, to SOA-enabled data services[4].

If you're interested in learning more about Warehouse Builder, Data Integrator, or which technology is right for you and your business, contact Zirous.

ETLPROS
• Can balance the workload with the RDBMS.
• Can perform more complex operations in single data flow diagram (maps).
• Can scale with separate hardware.
• Can handle partitioning and parallelism independent of the data model and architecture.
• Can process data in-stream on route between source and target.
• Captures huge amounts of metadata lineage today.
• Can run on SMP or MPP hardware.

ETLCONS
• Requires separate and equally powerful hardware in order to scale.
• Can "bounce" data to and from the target (could result in performance issues).
• Cannot perform slower ELT unless running twice the size of hardware.

ELTPROS
• Leverages RDBMS engine hardware for scalability.
• Keeps all data in the RDBMS all the time.
• Can achieve 3x to 4x faster throughput on appropriately tuned MPP hardware.

ELTCONS
• Relies on proper database tuning.
• Relies on MPP hardware.
• Can be resource intensive during processing.
• Cannot balance the workload.
• Could easily double, triple, or quadruple disk storage requirements.
• In some cases, is NOT 100% metadata lineage traceable.
• Can take longer to design and implement[2].

[1]ETL, ELT – Challenges and Metadata written by Dan E. Linstedt accessed on March 24, 2008 at http://www.b-eye-network.com/blogs/linstedt/archives/2006/12/etl_elt_challen.php.

[2]Information in the table is from: ETL, ELT – Challenges and Metadata written by Dan E. Linstedt accessed on March 24, 2008 at http://www.b-eye-network.com/blogs/linstedt/archives/2006/12/etl_elt_challen.php.

[3]Oracle Corporation, Oracle Warehouse Builder site accessed on March 24, 2008 at: http://www.oracle.com/technology/products/warehouse/index.html.

[4]Oracle Corporation, Oracle Data Integrator site accessed on March 24, 2008 at: http://www.oracle.com/technology/products/oracle-data-integrator/index.html.