

An Oracle White Paper
February 2009

Oracle Data Integrator Enterprise Edition Executive Overview

Table of Contents

Summary	3
Introduction	3
The Three Rights.....	4
An Innovative Approach to High Performance Integration.....	5
One Platform That Evolves ON DEMAND	6
Keys to Flexibility.....	6
More Productive, Fewer Integration Resources.....	7
Enablers for Design and Maintenance Productivity	7
Fast Time to Launch, More Options in the Future.....	8
Layers of Reconfigurable Knowledge Modules	8
Oracle’s Vision for Data Integration	9

Summary

Data is one of your company's most important assets, and data integration constitutes the backbone of your enterprise's IT systems. Choosing the wrong technology for data integration can cause harmful long-lasting effects that impact not only the IT budget, but also the productivity and responsiveness of critical business divisions within your enterprise.

Oracle Data Integrator Enterprise Edition (ODI-EE), a key component of Oracle Fusion Middleware, provides a strong and reliable integration platform for your IT infrastructure. There are five key reasons why choosing ODI-EE may end up being one of the smartest decisions you make for your company:

- **Performance** – with its market leading unique architectural approach
- **Flexibility** – one platform, several data integration styles
- **Productivity** – declarative design tool drives reusability
- **Modularity** – hot-pluggable knowledge modules support diversity
- **Vision** – performance-driven, SOA-pluggable bulk data integration

"Oracle Data Integrator Enterprise Edition is helping us turn our data into gold"

—Jack Garzella, VP Data Warehousing, Overstock.com

Introduction

CIOs and IT leaders are responsible for a number of projects that rely on fast, accurate and flexible movement of data throughout the IT infrastructure. Data integration is the technology – or, more accurately, the set of technologies – that perform this movement in a variety of architectural approaches, depending on the individual requirements.

A recent survey of top IT priorities for 2008 cites a number of mission-critical projects, all of which demand a substantial data integration solution component:

- Business Process Improvement Projects
- Business Intelligence Projects
- Information Governance Initiatives
- Security and Risk Management Programs
- Service-Oriented-Architecture (SOA) Architecture Rollouts

Choosing a typical ETL (Extract Transform Load) to solve these data integration challenges might lead you down a dead-end path.

With so many years invested in the older data integration technology base, it is easy to overlook some of the nagging and persistent problems that it created. These problematic issues can be described as the following:

The “hub and spoke” architecture, still used by many data integration environments, is costly to implement and often becomes a bottleneck that slows down the overall IT systems

Many data integration projects are still treated as independent projects, with no overall consistency, resulting in silos of integration. This is exacerbated by most integration platform’s lack of support for multiple typologies of data latency and volume – for example, few systems support high volume data batches and trickle-feed real-time data flows within the same platform.

With top-down procedural design being the norm for many integration platforms, development and maintenance becomes very costly, this lack of flexibility can considerably hinder the time-to-market for a complete solution.

Tight coupling of data integration processes and operational systems results in a lack of flexibility: updates are much more costly and cannot be performed with the required agility of today’s fast paced business environment.

The best solutions on the market for unifying data across data-centric applications are ones that provide rapid productivity gains for the integration of data. Look for solutions that provide solution add-ons, or knowledge modules, for rapid-deployment and eliminate code creation. Also look for solutions that provide flexibility to work with multiple applications, databases, data warehouse frameworks, and BI applications.

"Oracle Data Integrator Enterprise Edition allows us to progressively transfer all of our group reporting to the data warehouse. We can thus anticipate difficulties with response times in light of the very strong growth of our activity and our resulting operational data exponential volume."

—Frédéric Rouiller, Head of Business Intelligence, NESPRESSO, NESTLE

The Three Rights

Getting the **right data** in the **right place** at the **right time** is the key to a successful data integration strategy.

- The right data – data that is not only appropriate for the use that is intended, but also that is accurate, and reliable.

- The right place – the overall information ecosystem consists of multiple operational and analytical systems, and they all need to benefit from data of the other systems regardless of their locale.
- The right time – data can become stale quickly. A decision support system that does not get the data in time is useless. A shipping application that does not get order information before cutoff time isn't efficient. Getting data in right time – with a latency that is appropriate for the intended use of this data – is one of the most important challenges faced by businesses today.

An Innovative Approach to High Performance Integration

ODI-EE provides all the elements of the solution needed to provide these three “rights”.

Its ability to connect to any source or target system – databases, files, packaged applications, etc. – ensures that no island of data will remain isolated from the data integration processes. And its built-in data integrity features guarantee that not only the data that is processed is accurate, but also that it is consistent and compliant.

Right time is about the performance of the integration processes. ODI-EE is the only data integration platform on the market to natively implement the high performance E-LT (Extract, then Load & Transform) architecture. Instead of moving all the data through an intermediate ETL transformation server, the E-LT approach leverages the power of the target RDBMS engines to perform the transformations, dramatically improving the performance at a much lower total cost of ownership. Data goes straight from sources to targets, and is transformed many times faster than other tools.

Unlike the “pushdown optimization” feature recently introduced by other ETL vendors, Oracle's native E-LT architecture is not a mere after-thought. Most of these so-called “pushdown optimized” transformations still occur inside the ETL engines and requires the physical data to transit over the network and through their engines anyway. While pushdown optimization may indeed improve the performance processes in some cases, the Oracle architecture is in a performance class of its own.

Other key features that address the right-time business need include ODI-EE's advanced capabilities such as Changed Data Capture (CDC) that transfers only updated records, and the ability to seamlessly change the latency of all integration processes – from batch to near-real-time to real-time.

¹⁶“...an incredibly powerful development environment, our development teams increased their productivity by at least 50%.”

—Leon Schurgers, Director of IT, **ABN AMRO Interfinance**

One Platform That Evolves On Demand

Data integration is not just batch ETL. It isn't only SOA data services either. Nor is it only event-driven real-time data flow. Data integration encompasses all of the above – and many more combinations of use cases. Oftentimes, data integration projects start with one specific need – i.e. populate the data warehouse, or exchange data in real-time between two operational systems. But then integration needs grow in new directions. Regardless of the scope of an initial project, choosing an integration platform that will grow with your integration needs is crucial to sustaining your operations and getting the most out of your IT investments.

"The first thing that struck us was the speed with which we ramped up our ETL developments with Oracle Data Integrator Enterprise Edition. The implementation and learning curve were extremely quick and in a few days we were fully operational with the solution."

—Leon Shaigorodsky, Project Manager, **iBasis**

Keys to Flexibility

ODI-EE supports several integration styles: data, events and services. It can move data directly from databases to different databases, to or from packaged applications such as ERP or CRM, to or from flat files; it can also provide data or events to an Enterprise Service Bus (ESB) or a Message-Oriented Middleware (MOM); it also becomes part of a Service-Oriented Architecture (SOA) by providing data services and transformation services.

ODI-EE also addresses all latencies, from monthly, weekly or daily batches, to hourly or more frequent batches, to near-real-time and real-time. Integration processes can be triggered by a pre-defined schedule, or by events from new data.

Unlike many “non-integrated” integration platforms available on the market, that were in fact built through successive patchworks of disparate technologies, ODI-EE is a single platform that addresses all these styles and latencies of integration within a common framework. The development approach and methodology is integrated across the platform, and reusability of shared components within ODI-EE and across Fusion Middleware is high.

ODI-EE is a single platform that addresses all data integration needs without compromise. When used in conjunction with other parts of the Oracle Fusion Middleware it becomes a central component of an end-to-end IT architecture and provides top-level shared data services across the enterprise.

"We needed a data integration tool that would reduce our dependency on manual coding of ETL scripts and leverage the power of our Teradata Warehouse for data transformation. With ODI-EE we have been able to deploy a standardized process for integrating large volumes of disparate data into the data warehouse with less development effort."

—Dan Hazel, Vice President, **Sabre Holdings**

More Productive, Fewer Integration Resources

The design phase of the integration processes is probably the most critical one in the entire IT infrastructure. Poorly designed integration will hinder dramatically the performance of the overall systems. Yet, it is also often a very time-consuming and costly process.

Declarative design, as introduced by ODI-EE, alleviates number of the concerns traditionally associated with integration processes design. It enables business users to participate into the design, decreases the workload of the developers, and significantly reduces development and maintenance times.

Enablers for Design and Maintenance Productivity

Conventional approaches to ETL design require you to design the entire flow of data, positioning transformation and mapping code at the nodes of this flow. Not only is this a time-consuming process, it is also prone to serious performance problems – one misplaced transformation can dramatically impact the overall system performance. As a result, for this legacy approach to actually produce acceptable performance, it requires highly skilled integration developers who are highly proficient in operating system and RDBMS tuning. But even worse, these approaches make your developers do repetitive work that needn't be done at all.

Suppose you want to populate a target table from three different source tables. With the conventional ETL approach, you need to design every step of the data flow: (1) connect to the 3 source tables, (2) sort the first source, (3) filter and look up data from the second source, (4) transform and stage the result of the lookup, (5) read and join the third source, (6) transform and stage the result of the join and then finally (7) write to the target. These time-consuming steps need to be repeated every time you want to populate a different target from different sources.

Declarative design, the ODI-EE approach, allows you to accomplish the same result in only three steps: (1) link the 3 sources together, (2) declare the mappings between source fields and target fields and (3) choose the built-in template that will generate the data flow. You don't need to worry about all the intermediate steps required in the conventional approach. They are generated for you by the ODI-EE Knowledge Modules.

Therefore, you concentrate on "What" your transformation is - declarative rules, rather than on "How" to do it - technical steps of the process. To populate different targets from different sources, you will always have only these three steps to complete – regardless of what the number of sources and targets are.

Further, suppose we want to update this previous example to add an extra source table, the conventional approach requires the following additional steps: (1) connect the new source table, (2) figure out the best place for it to get in the process, (3) read and sort it, (4) join it to the appropriate staging result, (5) transform and stage the new result, (6) reconcile this new result within the existing process.

But with declarative design, you simply need to: (1) link this additional source and (2) declare the new mappings between source fields and target ones. The Knowledge Module will figure out automatically the best place for the new table to get in the process and generate the appropriate internal flow for you.

ODI-EE will instantly provide your development team with:

- a shorter learning curve: concentrate on mappings only
- faster design: fewer steps to implement transformations
- drastically reduced maintenance costs

Since behavioral changes are localized within the ODI-EE Knowledge Modules, the incremental effects of change can be automatically inherited and propagated to all transformations within the ODI-EE system.

Fast Time to Launch, More Options in the Future

The unique productivity benefits achieved with ODI-EE derive from an industry-first approach to Data Integration modularity called Knowledge Modules (KM). KMs provide ODI-EE with unmatched flexibility, productivity, and modularity that directly translates to an improved Total Cost of Ownership (TCO).

Layers of Reconfigurable Knowledge Modules

ODI-EE's Knowledge Modules (KMs) are at the core of the product's integration architecture. They make integration processes modular, flexible, and extensible. KMs are master templates that define both the data flow implementation and the code generation.

Knowledge Modules are generic, because they allow data flows to be generated regardless of the transformation rules. At the same time, they are highly specific, because the code they generate and the integration strategy they implement are finely tuned for a given technology or application platform.

ODI-EE provides a comprehensive library of Knowledge Modules, which can be tailored to implement existing best practices. Examples of custom tailored Knowledge Modules might include: performance tuning to a specific DB instance, enforcing adhering to corporate standards and policies on the data, or to include specific vertical know-how within a given execution flow.

Because each KM is optimized for a given technology, porting a set of processes from one environment to another is as simple as picking the appropriate KM. Knowledge Modules are the foundation of ODI-EE's hot-pluggability – enabling the ODI-EE to support the widest possible number of platforms and applications without performance overhead or the requirement of a separate ETL server.

Knowledge Modules are the underlying components that enable declarative design, and accelerate the overall design, development, implementation and maintenance of the end-to-end data integration processes.

Oracle's Vision for Data Integration

In summary, it should be clear that ODI-EE provides a unique data integration platform, architected for performance and productivity, which provides a high degree of flexibility and modularity. Compared to conventional ETL approaches from our competition, the ODI-EE solution excels in the following areas:

- Performance – thanks to a market leading unique approach – the high-performance E-LT architecture
- Productivity – with declarative design that drives reusability, accelerates development and maintenance
- Flexibility – one single platform that natively supports several data integration styles and latencies
- Modularity – with hot-pluggable knowledge modules that support application and database platform diversity

Oracle's commitment to our customers has always been to provide the highest quality enterprise software with maximum value. ODI-EE is a cornerstone technology for providing comprehensive data integration solutions and highlights the Fusion Middleware commitment for hot-pluggability and the broadest possible support for diverse IT environments.

Regardless of the database or applications within your IT ecosystem, the Oracle Data Integrator Enterprise Edition can be optimized to drive the highest-performance bulk or real-time transformations. Oracle's vision is to combine and enable these capabilities from within a next-generation, unbreakable Service-Oriented Architecture that will continue to drive business value within your enterprise for many years to come.



Oracle Data Integrator EE: Executive Overview
February 2009
Author: Dain Hansen

Oracle Corporation
World Headquarters
500 Oracle Parkway
Redwood Shores, CA 94065
U.S.A.

Worldwide Inquiries:
Phone: +1.650.506.7000
Fax: +1.650.506.7200
oracle.com



Oracle is committed to developing practices and products that help protect the environment

Copyright © 2009, Oracle and/or its affiliates. All rights reserved. This document is provided for information purposes only and the contents hereof are subject to change without notice. This document is not warranted to be error-free, nor subject to any other warranties or conditions, whether expressed orally or implied in law, including implied warranties and conditions of merchantability or fitness for a particular purpose. We specifically disclaim any liability with respect to this document and no contractual obligations are formed either directly or indirectly by this document. This document may not be reproduced or transmitted in any form or by any means, electronic or mechanical, for any purpose, without our prior written permission.

Oracle is a registered trademark of Oracle Corporation and/or its affiliates. Other names may be trademarks of their respective owners.