



KIMBALL  
GROUP



KIMBALL  
UNIVERSITY

# ETL Architecture in Depth

Altis Consulting, the Kimball Group's partner in Australia and New Zealand, is excited to announce that Kimball University is delivering the *ETL Architecture in Depth* (ETLD) course in Sydney, 13-16 September 2010. The course runs for 4 days and will be presented by **Ralph Kimball** and **Bob Becker**.

## Why Attend

This class makes sure that you understand all the factors necessary for effectively designing the back room of a data warehouse that can gracefully evolve over time as your needs mature and new technologies become available.

This class is taught by Ralph Kimball, co-author of *The Data Warehouse ETL Toolkit* book, and Bob Becker, a leading expert in the design of data warehouses for the health care industry and a member of the Kimball Group. See [www.kimballgroup.com](http://www.kimballgroup.com) for teacher profiles.

## Who Should Attend

This course is designed for data warehouse implementers, who are responsible for building the back room, or ETL portion, of a data warehouse environment. This would include ETL developers, ETL architects, data warehouse operational staff, compliance tracking data warehouse professionals and real time data warehouse designers. Every attendee in this class will receive Ralph Kimball and Joe Caserta's book, *The Data Warehouse ETL Toolkit*.

## What You Will Learn

This is not a microscopic code-oriented implementation class. Rather, it is an architecture class for the designer who must keep a broad perspective, and who needs to know what the latest technologies and techniques make possible. The course is organised around 34 necessary ETL subsystems which are developed in detail as the course progresses. See the course outline over the page for the names of the 34 subsystems.

## Prerequisites

Familiarity with the basic principles of dimensional modeling is helpful since dimensional models are designed as the ultimate ETL deliverables. The student can gain this familiarity by reading the first four articles in the Fundamentals series of articles found on the Kimball Group web site. The class will include selected brief reviews of the principles of dimensional modeling so that everyone has the same vocabulary.

## Course Outline (see next page for details)

## Course Details

### COURSE

ETL Architecture in Depth

### LOCATION

Bayview Boulevard Hotel  
90 William Street, Sydney

### WHEN

13-16 September 2010

### REGISTRATION

[www.altis.com.au](http://www.altis.com.au)

### PRICE\*

- Full rate: AU\$3700
- Early bird\*\*: AU\$3500
- Group rate<sup>^</sup>: AU\$3500
- 2 course rate<sup>^^</sup>: \$AU3400

\* All prices above are GST exclusive  
\*\* Early bird rate available until 5pm, 31 May 2010

<sup>^</sup> 3 or more students from the **same** organisation registering at the **same** time

<sup>^^</sup> Register for ETLD and DMD at the **same** time

### INFORMATION

Email [training@altis.com.au](mailto:training@altis.com.au) for more information or check [www.altis.com.au](http://www.altis.com.au) for the latest information

### OTHER COURSES

Dimensional Modeling in Depth (DMD), Sydney, 7-10 September 2010.

DMD brochure available at [www.altis.com.au](http://www.altis.com.au)

Sydney

Canberra

Melbourne

Auckland

(Numbered items refer to the 34 subsystems taught in this course)

#### DAY 1

##### Surrounding The Requirements

- Business needs
- Compliance
- Data profiling
- Security latency
- Archiving
- End user profiles
- Skills
- Licenses
- Coding vs. tool choice
- The restaurant analogy
- Data types used in ETL systems

- (1) Data Profiling
- Source to target map
- Access methods, source types
- Software, techniques
- (2) Change data capture
- (3) Extract window
- (3) Immediate transformations
- (3) Extract staging table designs, table types, retention, backup
- (22) Job scheduler
- (22) Exception handling architecture
- (23) Backup, (24) recovery, (24) restart
- Historical versus incremental load
- Team Responsibilities

#### DAY 2

##### Cleaning

- (4) Data quality architecture
- (4) Data quality screens
- (5) Error event fact table
- (6) Audit dimension, compliance tracking
- (28) Sorting
- Module designs: (7) customer deduplication, address validation, ...
- Final clean data table designs
- (8) Conforming
- Definition of conformed dimensions and facts
- Using the matrix
- Master data management
- Mapping incompatible structures into common structure
- (25) Version control
- (26) System and version migration, testing and regression
- (27) Workflow monitor
- (23) Job scheduler
- (29) Lineage and dependency analyzer
- (30) Problem escalation system

##### Modifying your ETL architecture for Real Time data warehousing

- The Hot Partition
- Streaming ETL vs. batch ETL
- Streaming extract
- Streaming cleaning and conforming
- Streaming delivery, query, reporting, dashboards, notifications
- EII architecture (Enterprise Information Integration)
- CTF architecture (Capture, Transform, and Flow)
- EAI architecture (Enterprise Application Integration)
- MBETL architecture (Micro Batch ETL)

#### DAY 3

##### Delivering Dimension Tables

- Referential integrity
- (9) Time variance designs (Slowly Changing Dimensions)
- (10) Surrogate key generator
- (15) Multi-valued dimensions, bridge tables I
- Special cases (extreme dimensionality, extreme dimension width, many incompatible members)
- (11) Hierarchical dimensions (fixed, variable, ragged), bridge tables II
- (12) Special dimensions (mini, junk, outrigger, shrunken, step, text fact)

##### Delivering Fact Tables

- (13) Fact table builder (transaction, periodic, and accumulating grains)
- (14) Surrogate key pipeline
- Graceful extensibility (add attributes, add facts, add dimensions to existing schemas)
- Handling structure changes
- (16) Late arriving data design

#### DAY 4

- (17) The dimension manager, responsibilities and procedures, real time complexities
- (18) the fact provider, responsibilities and procedures, real time complexities
- Distributed, federated data warehouses
- Delivering remote dimensions and attributes
- Delivering remote facts
- (19) Aggregations
- (20) Feeding OLAP cubes
- (21) DI (Data Integration) manager (feeding data mining, presentation layer extracts, 3rd party flat files)

##### Development and Operations

- (31) Parallel processing and pipelining
- (32) Security
- (33) Compliance
- (34) Metadata
- Process metadata
- Run results, exception handling, immediate schedule
- Technical metadata
- System inventory, data models, data definitions,
- Business rules, ETL jobs, transformations, batch parameters
- Business metadata
- Business definitions, source system info, DW data dictionary,
- Responsibilities
- Team roles
- Stepping back from the details
- The important decisions
- The important deliverables



#### Sydney

Level 6  
219-223 Castlereagh St  
Sydney NSW 2000  
Phone: +61 2 9211 1522  
Fax: +61 2 9211 3634

[www.altis.com.au](http://www.altis.com.au)  
[www.altis.co.nz](http://www.altis.co.nz)

#### Canberra

Level 6  
33-35 Ainslie Ave  
Canberra City ACT 2601  
Phone: +61 2 6262 5422  
Fax: +61 2 6262 5055

#### Melbourne

Level 3  
480 Collins Street  
Melbourne VIC 3000  
Phone: +61 3 8610 6973  
Fax: +61 3 8610 6976

#### Auckland

201 Victoria Street West  
Auckland 1001  
Phone: +64 9 369 1910  
Fax: +64 9 369 1940