

## Tuning ETLs for Database Transformations

Coming up on Thursday, May 28, our senior applications tuner Chuck Ezell will present a webinar on Tuning Your ELTs for Better BI. His presentation offers high level ideas on how to tune ETLs to reach the two main goals for tuning your ETLs for better performance; enhance speed and consistency.

Here is a sneak peek at what some of the presentation will cover.

### What is ETL?

- Extract data
- Transform data
- Load transformed data

An ETL extracts data from a variety of sources, often performing merges and transformations on the data, then loading it into a data warehouse.

Some of the sources you might see in the Extraction portion of an ETL could be:

- Cloud Data (Big Data)
- Data Warehouse Data
- Flat File Data
- EBS Data

### The 5S Approach to Database Tuning

Whenever you are performing large-scale ETLs, you want both the source and the target to be tuned as much as possible. We often use a "5S" approach to health checks which is designed to mitigate database issues before they occur. The 5S framework includes:

- SQL Code
- Statistics
- Segmentation
- Sessions
- Scheduled Processes

When tuning the source, the presentation explains some of the things we look for

- SQL Code (better filtering, filter as much as you can as soon as you can)
- Statistics
- Indexing & Fragmentation
- Conflicting Sessions and Processes
- Offload/Replication for better Data Isolation

During the transformation of the database, we monitor the activity, using the 5S approach. Much of the transformation activity happens in memory, so temp space is critical. Lack of proper temp space will cause failures, so keep an eye on your logs & ASM.

During transformation, filesystem lookups can be slow. Filesystem lookups perform better if they're converted to database table lookups. Indexing can also slow processes down. Filesystems can become fragmented

Tuning the destination database after loading is recommended. Focus on write speeds and I/O. Confirm Async I/O settings in the operating system and the database. Indexing and Stats can be problematic. Stats and Indexing can be added after loads and performed in parallel to other loads.

Chuck will conclude with samples scripts, screencaptures of monitoring tools, and reports generated to measure the performance of ETLs. You'll find Chuck's complete presentation on our website and the webinar will be recorded and available for playback soon after the conclusion of the webinar.

We encourage you to register for the webinar to take advantage of the live version so that you are able to ask questions during the webinar. Register today to save your spot. Be sure to also browse our resources section for more excellent webinars and our frequently updated blog for DBA topics.

**About the Author:** John Kaufling

*Vice President and Practice Leader of Application Services, Datavail*

John Kaufling has more than 20 years of experience in the IT industry, including more than 12 years as an Oracle EBS database administrator at Level 3 Communications and at Oracle Corporation. His specialties include implementations, upgrades, performance tuning and extensive capability to support the product. John's work with Oracle apps database administration has included experience with SOA suite, Veritas Cluster, Oracle DataGuard, Load Balancing from Resonate, Cisco and BigIP and extensive experience with Oracle self-service applications and self-service framework technology.