

Optimizing Your Data Load Improves More Than You Think

The format of the data that is loaded to Essbase is often an after-thought. But, should it be? When requesting the data file from a source system, it is more important than you may think to have it sorted to mirror your outline.

Assume an outline has the following dimensions.

- Period [DENSE]
- Account [DENSE]
- Region [SPARSE]
- Category [SPARSE]
- Product [SPARSE]
- Organization [SPARSE]

The most efficient way to receive a data file would be to have it sorted by Organization, Product, Category, Region, and then Account. Data files load faster when the columns that hold the sparse members are sorted in reverse order of the sparse dimensions that exist in the outline.

The reason the data loads faster is because it opens a block of data only one time. If the data was sorted by the dense members first, then every block would have to be opened multiple times. If the same sparse member combinations have 3,000 dense members with data, the block would be opened up to 3,000 times.

There are some more important benefits of doing this, however. When the block is opened multiple times, the database becomes far more fragmented than it needs to be. Fragmentation causes calculations to be slower and retrieving data can also be impacted, which can lead to frustrated customers.

By not sorting the data when loaded, every time a data load

occurs, any performance issues that may exist are exacerbated. So, anytime possible, sort the data load files by the last sparse dimension in the outline, the second to last sparse dimension in the outline, and so on. You may be presently surprised at the benefits.