



# Large Scale ASO: Making Essbase Silky Smooth with Terabytes of Data



Brian Marshall  
US-Analytics

# Agenda

- Introduction
- Major ASO challenges
- Getting the most from your hardware
- Q&A

# Introduction – About Brian Marshall

- 10+ Years IT and EPM/BI Experience
- Began career as a software and database developer at a small software firm.
- Developed specialization in Microsoft BI offerings.
- Focused on Oracle EPM, primarily Hyperion Essbase and Planning, with some HFM.
- Presented at Kaleidoscope 2010, 2011, and now 2012 and various regional events.

# Introduction – About US-Analytics

## ***“Focused and Committed”***

- Dallas-based Industry Leaders, Pioneers and Trustworthy for 13 years
- Focused on enterprise performance management applications
- Over 50 professionals dedicated to EPM and BI
- Strategic Oracle Partner and Oracle BI Pillar Partner
- Advanced degrees and certifications (CPAs, CMAs, MBAs)
- Seasoned Infrastructure practice: 400+ installations/migrations
- Unique blend of deep technical expertise and business acumen with hundreds of implementation cycles, driven towards a results-oriented, customer ROI
- Strong project leadership & proactive account management
- Corporate culture of integrity with 100% customer commitment
- Full Service Solution Provider

# Introduction – About US-Analytics

## Performance Applications

- Design and development of EPM and BI solutions

## Operational Infrastructure

- Infrastructure design and installation services
  - Change management, disaster recovery, load balancing, fail-over

## Continuity Services

- Specialized Placement Services
- Helpdesk/Hotline support
- Education/Mentoring/"Expert-on-site"
- Software re-sell
- Managed Services

## • Leadership in Hyperion Community

- Founding sponsor of the Hyperion Women's Forum
- Presenters at the Kaleidoscope Conference
- Sponsor of Dallas Hyperion User Group (HUG)

# About The Data

- Built on Hyperion Essbase 11.1.2.2
- SQL Server 2008 R2 Data Source
- Millions of members in the outline
  - 2.2 Million Customers
  - 132,000 Products
- Hundreds of millions of rows of data
- Benchmarks use a subset

# Large Scale ASO Challenges

- Large dimensions
- Long restructure times
- Large data sets
- Slow data loads
- Slow query performance
- Long running aggregation materialization

# Large Dimensions

- Hundreds of thousands or even millions of members
  - Sku-level data
  - Customer-level data
- Large dynamic dimensions will kill performance, so always make these stored where possible
  - Multiple Hierarchies are your friend
- Partitions can help to break out dimensions that grow with time
- The more levels the better...
  - Phone-book flat dimensions



# Long Restructure Times

- First the simple solution...deferred restructure
- Clearing data can be faster...sometimes
  - Creates a slightly more complex process
  - Great for instances where you have to reload data anyway
- Partitions will allow you to split the data out and limit the amount of data being restructured

| No Data      |
|--------------|
| • 20 Seconds |

| 2GB Data      |
|---------------|
| • 235 Seconds |

# Large Data Sets

- How do we load large data sets efficiently?
  - Parallel Loads
  - Data Slices (incremental loads)
    - No parallel loads supported
    - Create multiple scripts to run in parallel with new slices
  - Optimize your Temp tablespace (more on this later)
  - Partitions (yes, I keep mentioning this!)

# Large Data Sets (Cont.)

- Enabling SQL Parallel Loads
  - Create this file:  
“C:\Oracle\Middleware\user\_projects\epmsystem1\ESSbaseServer\essbaseserver1\bin\esssql.cfg”
  - Use the DataDirect driver for SQL

System Data Sources:

| Name         | Driver                                  |
|--------------|---|
| EPM11122SQL  | SQL Server                              |
| EPM11122SQL2 | DataDirect 6.1 SQL Server Wire Protocol |

```
[
Description "DataDirect 6.1 SQL
Server Wire Protocol"
DriverName ARMS25
UpperCaseConnection 0
UserId 1
Password 1
Database 1
SingleConnection 0
IsQEDriver 0
]
```

# Large Data Sets (Cont.)

- Parallel Load Performance
  - MaxL SQL Load Sample:

```
import database ASODemo.ASODemo data
connect as hyperion identified by 'password'
using multiple rules_file 'dQ12011','dQ22011' to
load_buffer_block starting with buffer_id 100
on error write to 'C:\Process\Logs\dQ1Q22011.err';
```

- Performance depends on your source too:

| One Thread    | Two Threads   | Four Threads  |
|---------------|---------------|---------------|
| • 110 Seconds | • 141 Seconds | • 240 Seconds |

# Partitions

- Why would I use partitions?
  - Reduces the amount of data and meta-data that have to be touched for regular processes
  - Still faster than slices
  - Lets aggregations stick around for the majority of your data
  - Helps make better use of optimized hardware
- What do I partition by?
  - Time or other predictable patterns of analysis
  - Data elements that will reduce the size of dimensions per partition

# Aggregations

- How do we decide what to aggregate?
  - Usage-based aggregations
  - Size-based aggregations
  - Hint-based aggregations
- What gets aggregated?
  - Only stored dimensions are considered
  - Dynamic dimensions will remain just that...dynamic
- How do we speed up aggregation builds
  - Optimize your hardware (more on this later)
  - Partitions...so we don't have to rebuild them

# Hardware Optimization

- Three Main Factors
  - CPU Threads
  - Memory
  - Hard Disk
- As hardware gets cheaper, we start to throw more of it at Essbase
  - Without the right settings, all that power means nothing
  - Default ASO settings use 2 CPU threads and 32mb of RAM...seriously...

# Hardware Optimization (cont.)

- So I have a server with 64 threads, 256GB of RAM, and more hard drives than you can shake a stick at...now what?
  - CALCPARALLEL
    - Defaults to 2 threads for ASO.
    - New to 11.1.2.2, this can be set all the way to 128!
    - Previously this maxed out at 8...sorry old versioners.
  - Aggregate Storage Cache
    - Pending Cache Size Limit (set per application)
    - Defaults to 32mb...why not just set it to 128gb?



# Hardware Optimization (cont.)

- Why not just set the CALCPARALLEL and cache settings all the way up?
  - Your hard disks probably can't keep up
  - If you have 128gb of data processed in cache, it then needs to be committed to the temp tablespace (more on this later)
  - When that data starts to get written, you start wasting CPU cycles
  - Find the balance!

# Tablespaces

- Every ASO application will require at least two tablespaces
  - Default – The tablespace that holds all of your data and aggregations
  - Temp – The tablespace where data is initially loaded before being committed to the default tablespace
- Hard drives are generally the weakest link
  - Default and Temp should be on separate hard drives, raid arrays or HBAs (SAN cards).
  - During the final commit, data is transferred from one to the other, so separate them

## Putting it all together...

- Deferred restructures to speed up dimension builds
- Parallel data loads to get large sets of data in faster
- Partitions to split data and meta-data into more than one cube
- Optimize your server and application settings to make the best use of your hardware

## Q&A



# **Large Scale ASO: Making Essbase Silky Smooth with Terabytes of Data**

**Please fill out your evaluations!**



Brian Marshall  
US-Analytics