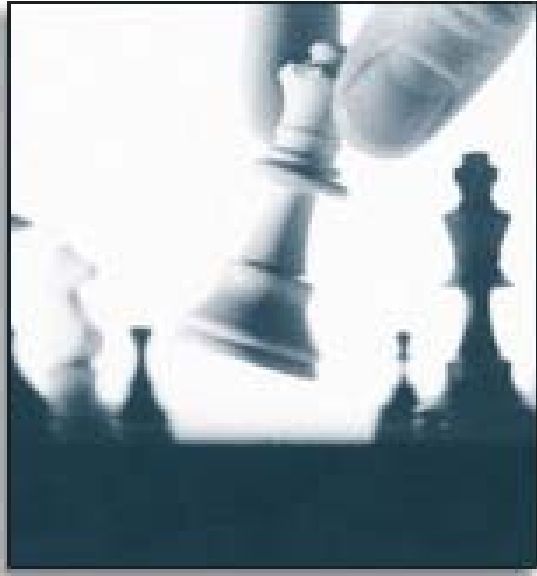


**ORACLE®**



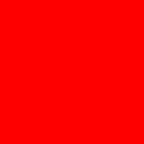
# ORACLE®

Live on Oracle Database 11g XML DB

**Carlo Tiu (NCPA)**

**Geeta Arora (Oracle)**

**Bhushan Khaladkar (Oracle)**



**The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions.**

**The development, release, and timing of any features or functionality described for Oracle's products remains at the sole discretion of Oracle.**

# Agenda

- **XMLDB 11.2 enhancements**
  - **Schema and XQuery**
  - **Structured usecases**
  - **Semistructured and Unstructured usecases**
  - **Repository**
- **NCPA demo**

# XML Use Cases

## Structured

**“Data-Centric”**

**Static Schema  
with  
Occasional  
Variability**

**No  
ANY or MIXED  
content**

## Semi Structured

**“Variable Data”**

**Dynamic  
&  
Complex  
Schema**

**Islands of  
ANY or MIXED  
content**

## Unstructured

**“Doc-centric”**

**No Schema  
Variable  
&  
Flexible  
Schema**

**Repeating  
ANY & MIXED  
content**

# Agenda

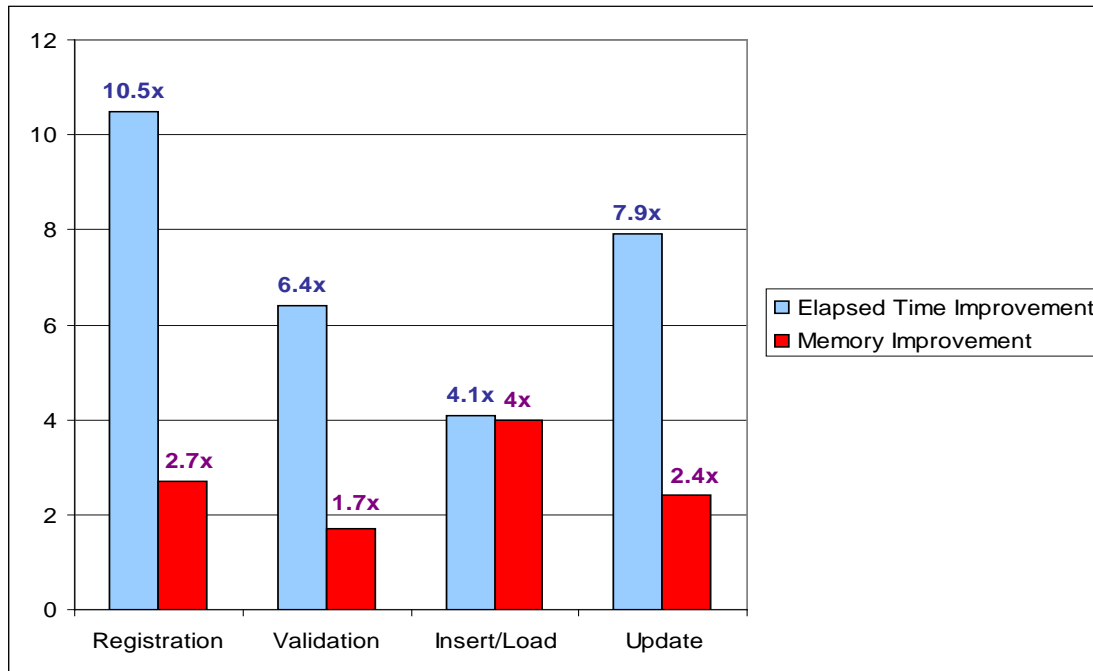
- **XMLDB 11.2 enhancements**
  - **Schema and XQuery**
  - **Structured usecases**
  - **Semistructured and Unstructured usecases**
  - **Repository**
- **NCPA demo**

# XML Schema Enhancements

- **Schema registration performance**
  - Eliminate internal and external memory fragmentation
  - Optimized schema loading
  - Time and memory improved by **50x** for US-GAAP, HL7, NIEM
- **Schema Validator Cache**
  - Improves XML schema validation by around 5x, more for small docs
- **Can handle complex industry schemas**
  - **GJ-XML, GML, US GAAP, NIEM, HL7, FixML, MPEG-7, KML**
  - **ACORD, SDMX, FPML, Reed, OAGIS, MPEG7: Binary & O-R**

# XML Schema

## Performance Improvements over 11gR1



**Avg. improvement for  
NIEM, HL7, US GAAP, FPML**



# XQuery and SQL/XML

## Storage Independent Enhancements

- XQuery 1.0 / SQL/XML standard compliant
- Handling Large XQueries
  - Re-write enabled for large size Xquery (upto 32K)
  - **8x** increase in size and complexity of supported XQuery operations
- **20x** improvement in functional evaluation of Xquery
- Up to **60x** improvement for XML Generation from relational data

# Agenda

- **XMLDB 11.2 enhancements**
  - Schema and XQuery
  - **Structured usecases**
  - Semistructured and Unstructured usecases
  - Repository
- **NCPA demo**

# Object-Relational Storage

## Enhancements

- **Inheritance constructs like `xsi:type` and “instance of” now rewrite**
- **Default storage of collections is “table” across the board**
  - leads to relational query performance
- **Document ingestion and fragment retrieval faster and more scalable**

# Ingestion of Large Documents in Object-Relational Storage

- **Files upto 4G can be ingested in an O-R table**
  - up from 500M pre-11gR2
- **“Direct Insert” : Greedily flushes document fragments directly to destination tables**
- **Scalable and Performant**
  - Minimizes fragments in memory
- **Applied when documents are uploaded to XDB using**
  - FTP protocol
  - PL/SQL function `DBMS_XDB.createResource`

# Agenda

- **XMLDB 11.2 enhancements**
  - Schema and XQuery
  - Structured usecases
  - **Semistructured and Unstructured usecases**
  - Repository
- **NCPA demo**

# XMLIndex Table-based (New in 11gR2)

## Example

### XML Data

```
<Address>
  <city>Fremont</city>
  <state>CA</state>
</Address>....

<Address>
  <bold> <city>Melbourne</city></bold>
  <font size="21">
    <state><!-- state is not in US --
  ></state>
  </font>
  <country>Australia</country>
```

### XMLTable Index

Key	City	State	Country
10	Fremont	CA	
20	Melbourne		Australia

# XMLIndex Table-based

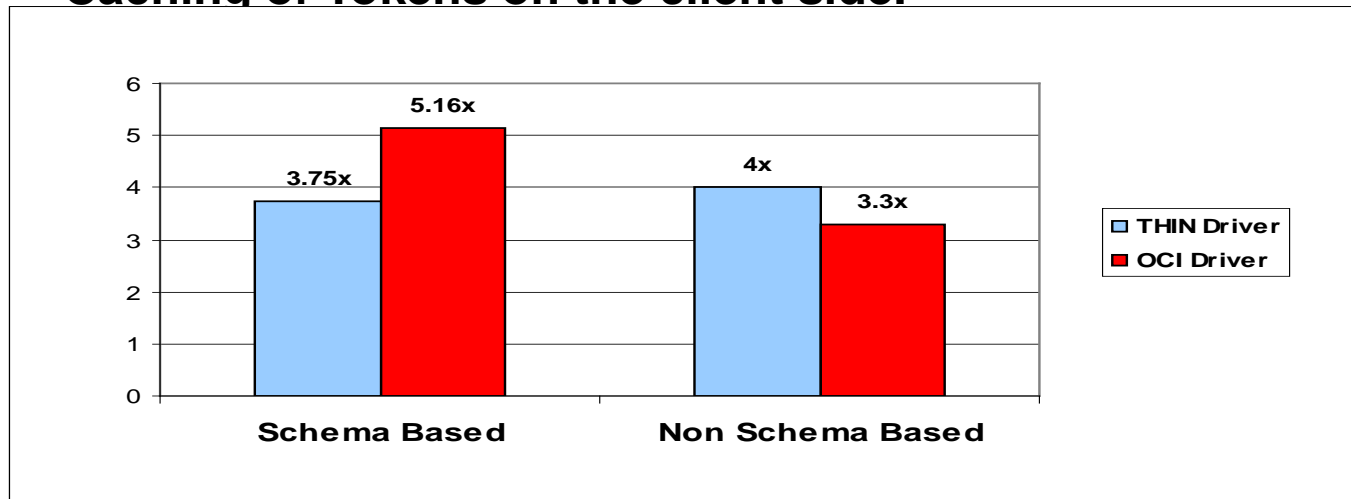
- **Complements Path based XML Index**
- **Project commonly searched structured data into a table**
- **Secondary relational indexes can be created on top of index content tables**
- **Text Index can be created on projected text column**
  - **Speed up text search in XML content**
- **Rewrites queries for relational performance**

# XMLIndex Path-based Enhancements

- Partitioning, parallel index creation and parallel query supported
- Physical rewrite for path subsets
- Many queries improve **5-20x**
- Asynchronous DML performance improves **2.5x**

# Binary XML Enhancements

- Improved query performance without indexes
  - New Search based decoder
  - Document level Summary
  - Improved XPath Caching
- Improved End to End Binary XML performance for Thin and OCI drivers
  - Caching of Tokens on the client side.



# Agenda

- **XMLDB 11.2 enhancements**
  - Schema and XQuery
  - Structured usecases
  - Semistructured and Unstructured usecases
  - **Repository**
- **NCPA demo**

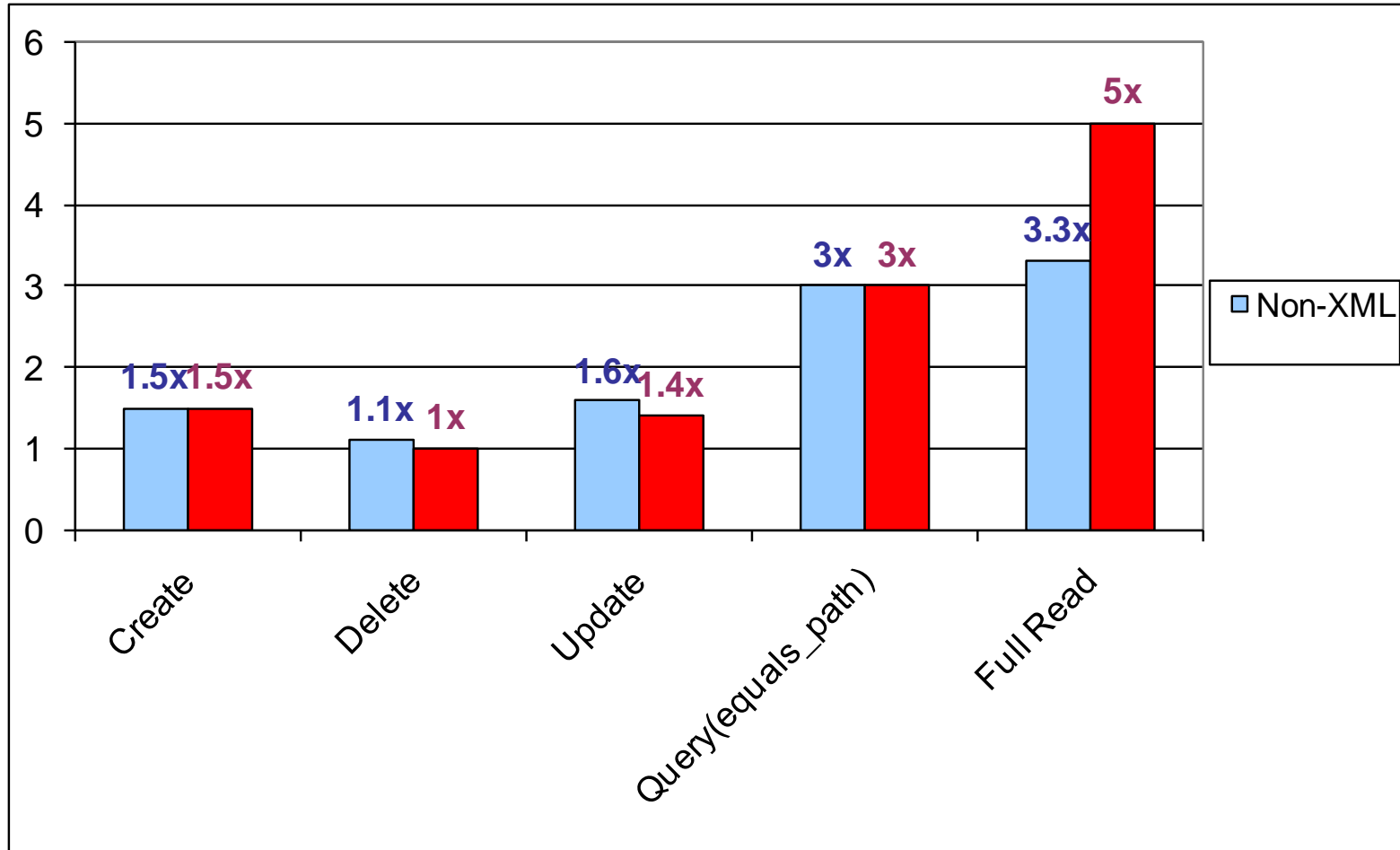
# XML DB Repository

## Enhancements

- Secure files for all repository content
- Create Operations
  - Improved creation performance of documents conforming to system schemas (ACLs etc).
  - Avoid creating intermediate DOMs while creating resources
  - Optimized in-memory structures
- Retrieval
  - Improved resource\_view queries where predicates involve equals\_path and under\_path – **3x** improvement.
  - Avoid copies during full retrieval of documents - **2-3x** improvement.
- Delete
  - ACL Deletion performance increase by **5x** using ACL OID Index on resources.

# XML DB Repository

## Performance Improvements over 11gR1



# Agenda

- **XMLDB 11.2 enhancements**
  - **Schema and XQuery**
  - **Structured usecases**
  - **Semistructured and Unstructured usecases**
  - **Repository**
- **NCPA demo**

# Northern California Power Agency (Structured)



- NCPA supplies and purchases power from the California grid on behalf of it's members
- Settlement system leverages SOA and XML
- “XML DB Short Circuits the Data Transformation”
  - NCPA is able to process large volumes of XML using commodity hardware
  - System has proven flexible and adjusts quickly to ongoing changes
  - Source code available as ‘open-source’ – being adopted by other market participants
- In-house solution developed at a fraction of the cost of an off-the-shelf solution
- Currently live on 11gR1 with RAC



# Live on Oracle Database 11g XML DB

*Presented by:*

**Carlo Tiu**

Senior Programmer, NCPA

Carlo.Tiu@ncpa.com

651 Commerce Drive, Roseville, CA 95678

ORACLE OpenWorld US 2009

San Francisco, CA

October 11-15, 2009

## Brief Background on NCPA

- NCPA supplies and purchases in the California grid on behalf of it's 16 cities in Northern California
- Our load is about 1000 MW
- We celebrate 41 years of service this year
- Our members' value is derived in large part from:
  - Our ability to be a cost leader
  - Our ability to be agile under different market conditions and regulations

## Start of XMLDB Project

- CAISO was formed in 1997 as a result of market restructuring in California
- The meltdown of 2000 formed the basis for the current Market Redesign.
- The aging equipment at the CAISO was the basis for a system refresh of the equipment and technology under MRTU.
- Web Services, B2B Communication and XML Payload Delivery

## Project Challenges for NCPA and MRTU

### MRTU Challenges

- Evolving Requirements
- Complex Market
- Large Payloads

### Business Challenges

- High Reliability
- Agility Built In
- Cost Control

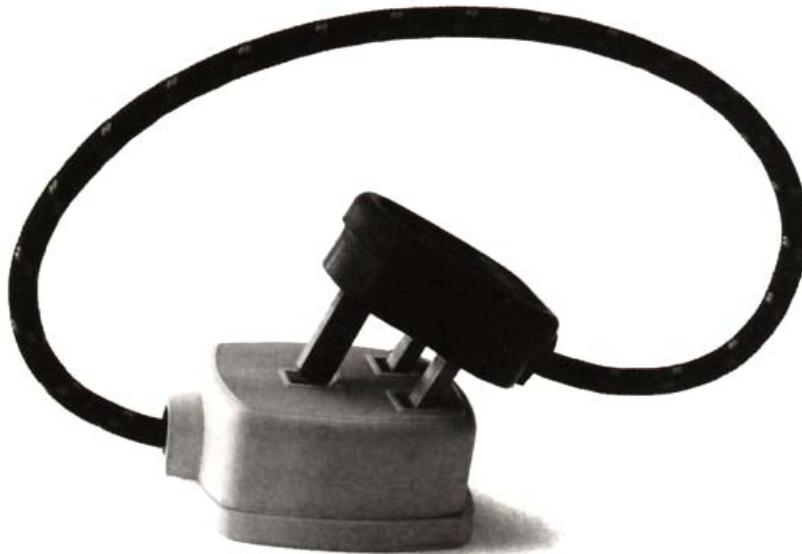
# XML DB Short Circuits the Data Transformation

## XML DB

- Register XSD
- Load XML
- Write Queries

## Traditional Approach

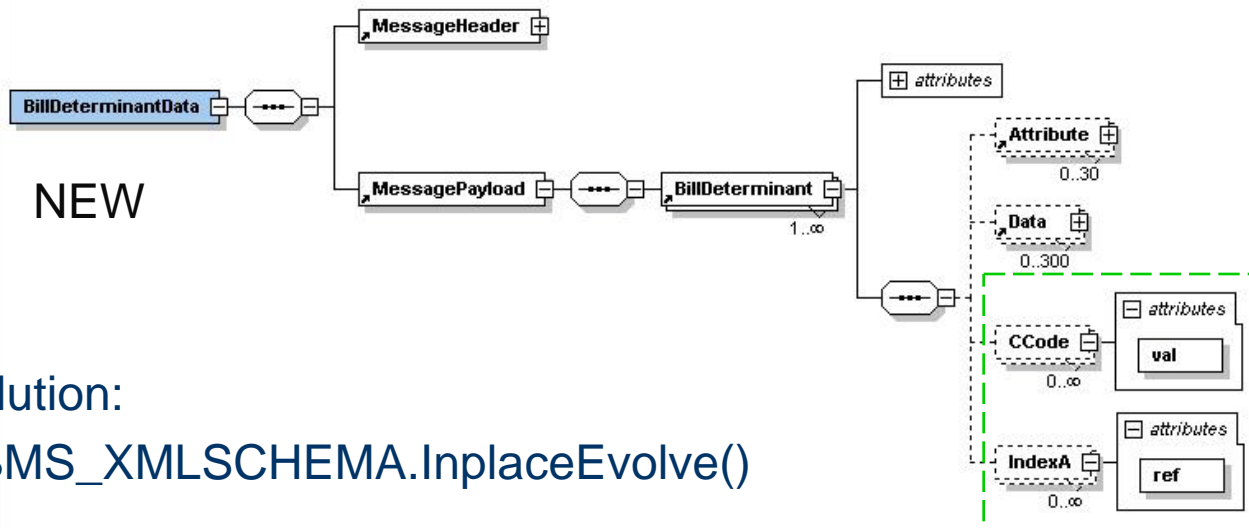
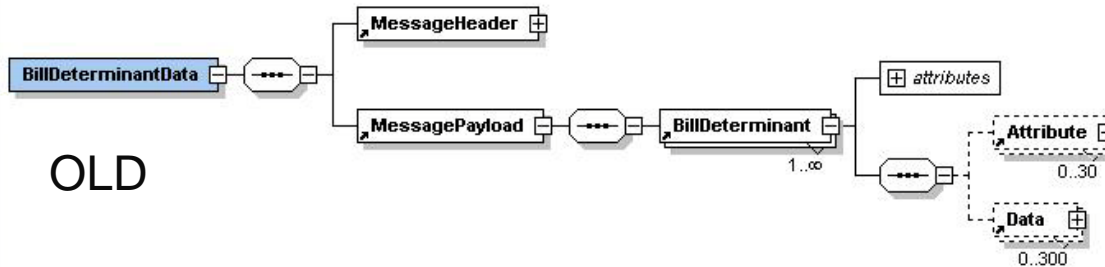
- Design Tables
- Build Tables
- Parse Data
- Write Queries



# DEMO

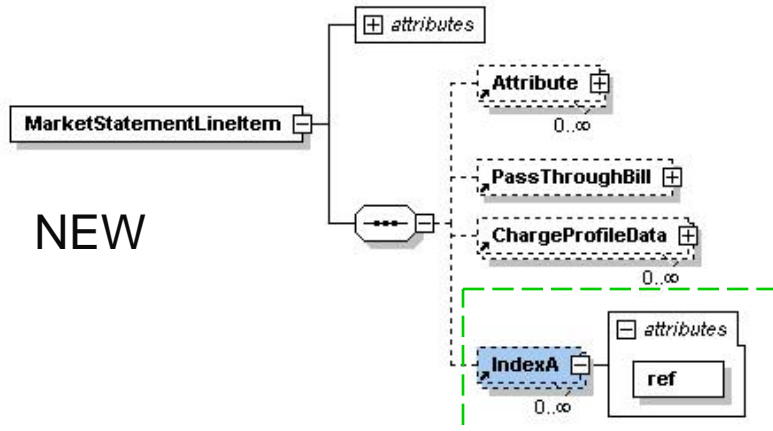
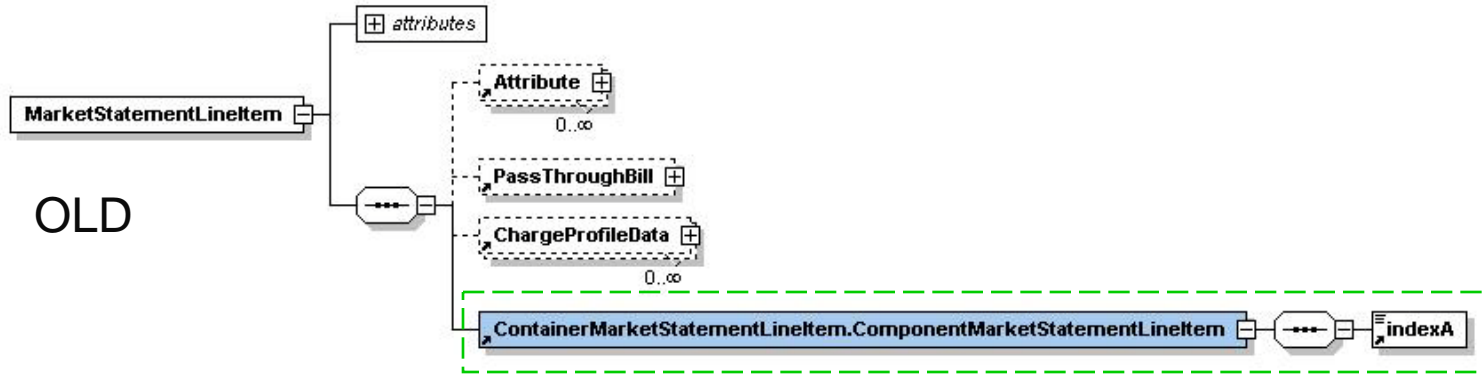
- XML Schema Registration
- XMLType tables generated from the registration
- Loading XML Instances
- Xquery Views
- In-Place Schema Evolution
- Copy-Evolve Schema Evolution
- Schema-Based Binary XML

# BillDeterminantData.xsd Schema Evolution



Solution:  
`DBMS_XMLSCHEMA.InplaceEvolve()`

# StatementData.xsd Schema Evolution



Solution:  
`DBMS_XMLSCHEMA.CopyEvolve()`

## Major Systems Built

- 26 Different Systems Were Impacted
- Major Systems Built For MRTU:
  - Bidding System
  - Settlement System
  - Master File System
  - Dispatch Retrieval System

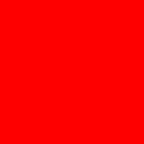
## Costs

- Hardware, Software, and Third Party Components
  - \$300 thousand (one time cost)
- Labor (two year effort)
  - \$900 thousand (internal)
- \$1.2 Million Invested
- \$250K Per System.....A Real Bargain

## Hardware and Software Configuration

- 64-bit Windows Server 2003 SP 2 on 2 Intel Xeon 5160 dual-core CPUs, with 8 GB of RAM.
- The database is 220 GB (today).
- The Oracle software is Oracle Database 11g Release 11.1.0.7.0 – 64bit Production

# End




**The preceding is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions.**

**The development, release, and timing of any features or functionality described for Oracle's products remains at the sole discretion of Oracle.**

For More Information

**search.oracle.com**



or

**oracle.com**

**ORACLE®**