



Real Time Loading for Sybase IQ

Sybase IQ: Target Markets in 2009

Real-Time Loading Valuable to All

Report Servers Play

- Horizontal market focused on enterprise or departmental reporting application users:
 - Ad-hoc and canned queries
 - Dashboards
 - Corporate Performance Management (CPM)
 - Key Performance Indicators (KPIs)
- **Real-Time data delivers more current picture of business conditions to executives and decision makers**

Advanced Analytics Play

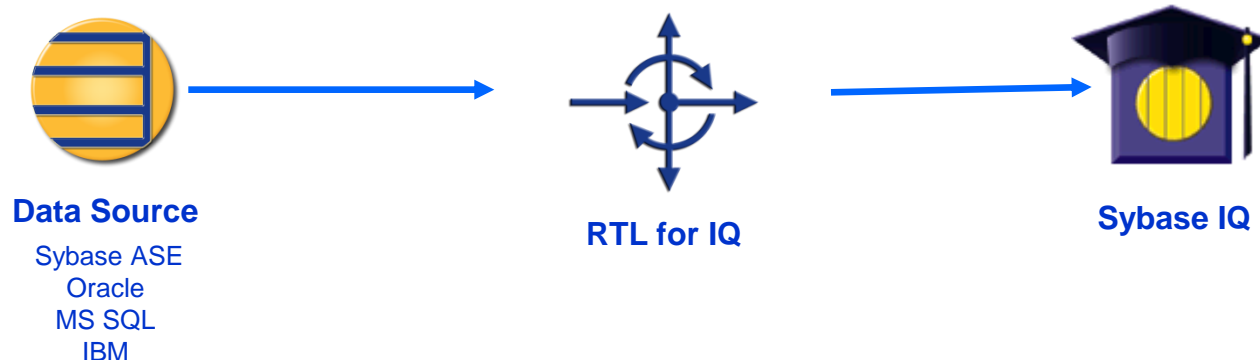
- Highly complex query environments (in-house) that support strategic/ operational decisions; used to gain a competitive edge by better understanding of customers, competition, risk positions, revenue leaks, fraud or scientific knowledge discovery
- **Real-Time data fuels the most accurate predictive analytics & risk mitigation**

Data Aggregator Play

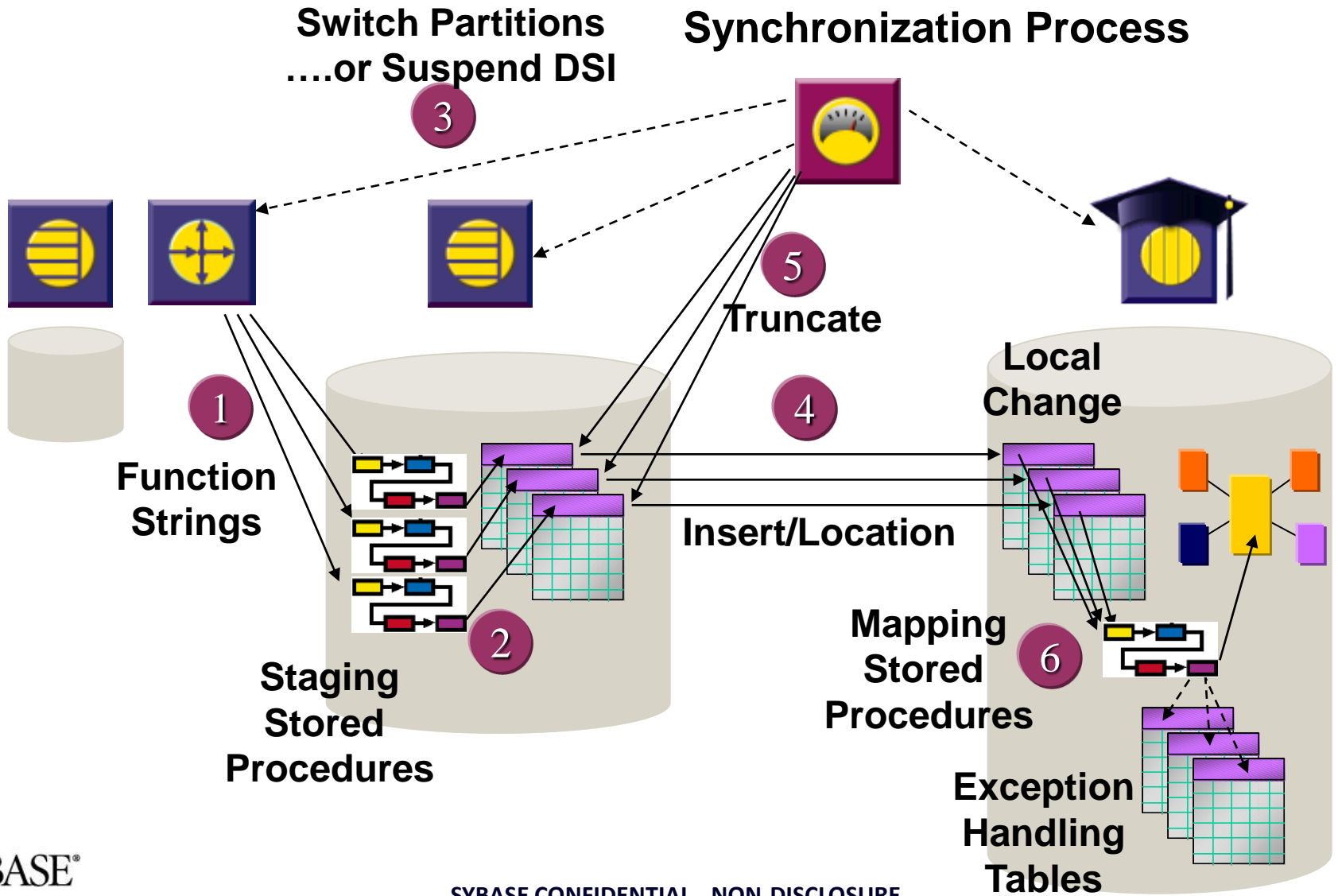
- Information services companies or information services divisions who offer analytics services to their customers – typically for revenue
- **Real-Time data powers analytics services that deliver true competitive differentiation**

Solution Overview

1. **The Real-Time Loading for IQ is a highly optimized solution to directly address customer needs for real-time analytics on current operational data**
2. **It leverages Replication technology for non-intrusive change data capture and optimizes loading for Sybase IQ**
3. **Releasing schedule:**
 - ASE to IQ: Q1, 2010
 - Other data sources to follow



Current Architecture



Issues With Current Architecture

Staging DBMS seen as overly complex

- Despite fact it is a clear part of architectures based on Kimball or Inmon
- Something else to administer

Customer has to supply synchronization process

- i.e. cron, CA Autosys, etc.

Latency in High-Volume Replication Environments

- Historically, cause is replicate ASE - in this case the staging database
- One problem is that in if only a single unpartitioned table was used per operation, you had to suspend the DSI to prevent DRI issues with orphaned rows in the reporting system
 - Suspending DSI = Latency
- Work-around is to use two sets of tables and flip with each synchronization - or to use partitioned tables (ASE 15) partitioned on the synchronization period (e.g. 10 min intervals) using a wrap-around partition scheme
- Adds to the complexity of the above

RS 15.5 RTL/HVAR Replication

Create a consolidated in-memory database for staging in RS

- RS 15.5 will be 64-bit largely for this reason
- Not a true DBMS with SQL, tables, etc. - essentially still structures
- Inserts, Updates, Deletes consolidated as with current staging database
 - Customizable common conversions such as u2di or d2none

Synchronization will be based on:

- When buffer is full or reaches configuration limit
- Will attempt to use target DBMS bulk load API's
- Possibly (caveat) may be usable in conjunction with periodic/scheduled (every n minutes) replication
 - RS 15.5 adding time-based scheduler for DSI universally as an independent functionality

The Two Modes

- Real-Time Loading (RTL): Sybase IQ (licensable option)
- High Volume Adaptive Replication (HVAR): ASE (may be licensable)
 - May be especially applicable for ASE 15.5 w/ In-Memory Database
 - Future possible support for Oracle, MS-SQL, DB2 UDB

What is RTL/HVAR Replication

RS normally sends DML as rows are replicated

When a threshold is crossed, RTL/HVAR replication will kick in

- HVAR - When DSI backlog crosses a configured threshold. For example, if DSI is more than #MB behind, HVAR will kick in.
- RTL - If the number of commands in a group of transactions is greater than a configured value, RTL will be used
 - You control this by setting the DSI grouping limits - the larger you set DSI grouping, the more likely RTL will be used

RTL/HVAR will consolidate net changes of DML for the current DSI transaction group

Consolidated changes are applied using bulk load + merge synchronization

RS starts next DSI transaction group

RTL/HVAR Synchronization

Data will be bulk-loaded into target DBMS work/temp tables

RS will perform synchronization between temp & real tables

- Delete & Updates using joins to work/temp tables
- Insert net new row images

Problem: Sybase IQ

- Does NOT support a bulk-load (push-style) API
- strictly file based bulk loaders
 - Both for local and client side files (IQ 15.)
 - In the case of client files, it uses a protocol API to transfer the file blocks across the network and still uses the file based loader
- There still is the [pull] bulk-loader via insert/location
 - Enhanced in IQ 15 to allow internal parallelism for insert/location
- This will require RS to be able to:
 - Connect to IQ and issue commands on session to connect back to RS itself
 - Mimic a database connection
 - Handle a basic select clause to extract data from the consolidated buffer
 - Cleanup without problem