

Agenda



- Introduction
- RBDR Approach - Rule Based, non-conforming dimensional Data Reconciliation modeling
- Business Case – An Asset Reconciliation Engine
- Benefits
- Questions

Ravi Ananthanarayana



- 16 Years of Data Modeling and Business Intelligence
- Enterprise Data Architect with Rolta's Business Intelligence Practice
- Our 50+ BI Practice is primarily based in Chicago
- Active member of various DW/BI groups

3

Highlighting the Use Case



- Information from a System of Record (SOR) is duplicated throughout an organization
- Over time, the data becomes inconsistent and complex multi system logic chains begin to fail and cause real problems
- Warehouse exists (or is in process of being built) for Business Intelligence that pulls from multiple data sources
- Opportunity! Leverage existing warehouse to pro-actively reconcile SOR to other data sources
- Data is corrected in the source systems
- Over time, identify repeating data inconsistencies and remediate if appropriate

4

The RBDR Solution Overview



- Fact Table highlights data inconsistencies across disparate data sources
- Non-Conformed Dimensions are only used within this subject area
- Data is not cleansed(allow duplication) in the Dimensions and loaded directly from stage tables
- Rules Dimension describes what to reconcile and how
- Standard ETL used for the Dimensions
- PL/SQL fills the Fact Table
- User screens allow rule manipulation

5

What can be Reconciled?



- Find Duplicates - Dimensions are not cleansed prior to load
- Find Mismatches Across Data Sources - Find mismatched attribute values between the System of Record (SOR) and any other data sources
- Find Mismatches Within a Single System - Find mismatched attribute pairs within a given system (Key – Performance)
- Identify Missing Rows - Compare two systems and look for the existence of a row in one system that is not found in another
- Verify Tolerance Limits - Look for out of bounds conditions within data
- Complex Exception Conditions Across Data Sources - Apply user-defined SQL to identify any exception not easily defined within the rules framework

6

To ETL or not to ETL, that is the question



- In some cases, tool-based ETL processing is not the right option:
 - When the complexity of the ETL would lead to an extraordinary development effort
 - When data volumes exceed processing power limitations
 - When business rules change rapidly or need to be switched on or off
 - Bulk data load situation – business rule that are context or time sensitive
- For a Rules-Based Data Reconciliation Model, both the complexity of the ETL and the rapidly changing rules point towards a different solution...PL/SQL

7

PL/SQL Rules Engine - Simplified



- Rules are created by business users through an UI
- Define various rule categories & types
- Complex SQL rules created with support from IT
- PL/SQL Engine processes rules, for example:
 1. SOR Dimension and column identified
 2. Comparative value in separate Dimension identified
 3. Join keys identified between Dimensions
 4. Apply rule (i.e. =, >, <) by creating dynamic SQL and executing:

```

SELECT  *
FROM    DIM_SOR a,
          DIM_OTHER b
WHERE   a.key = b.key
AND     a.column1 = b.column5;

```

- Inconsistencies written as individual rows to the Fact table

8

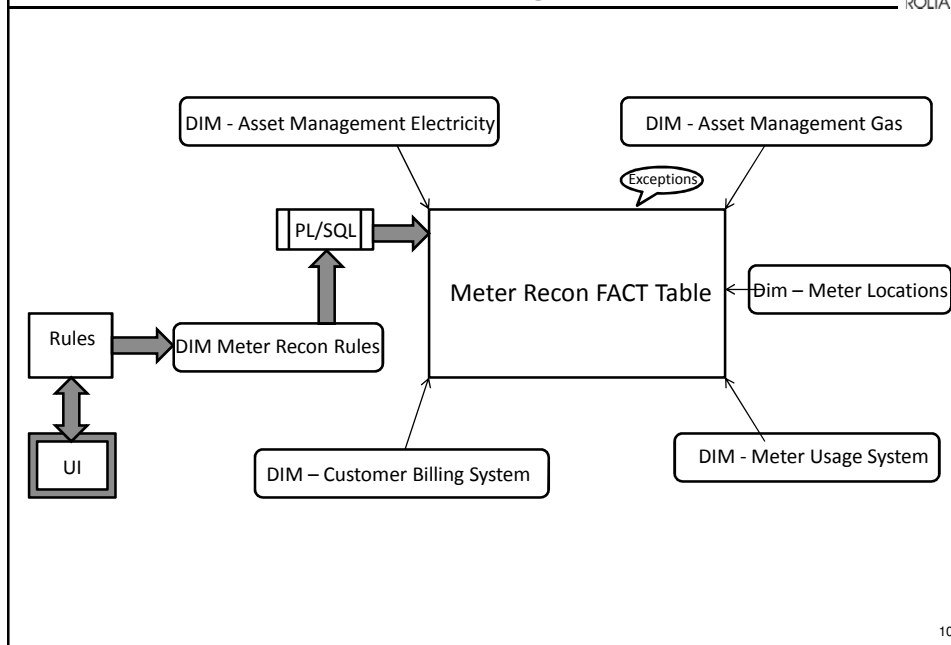
Real-World Example – Meter Recon



- Northeast Utilities serves nearly 3 million customers in New England
- Each customer has at least one meter and/or gas meter
- Information about meters are found throughout Northeast's systems:
 - Asset management – The SOR
 - Collection systems – meter reading
 - Billing systems – electricity usage
 - And many others....
- Meter attribute inconsistencies can cause
 - Meter's to go unread which causes estimated bills
 - Electricity and gas usage to go unbilled entirely
- Northeast previously reacted to these problems, Meter Recon allows NU to proactively correct the data


9

The NU Meter Recon Engine



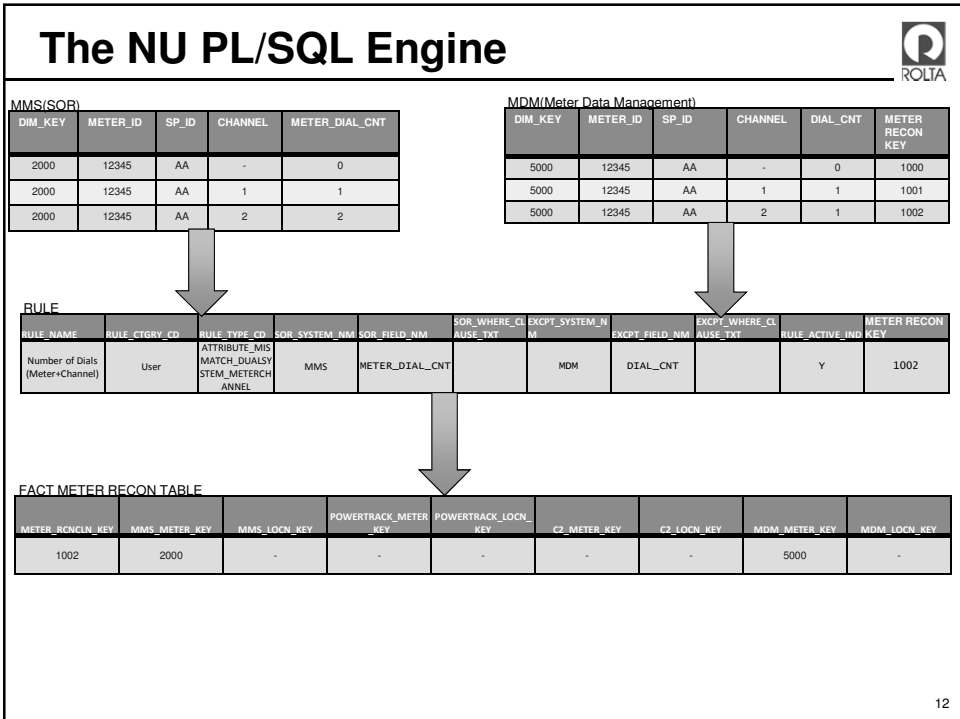
10

Sample NU rules Definition- Examples



RULE_TYPE	Description	Where- Clause Applicability	Entity Type	Rule Type
DUPLICATE_METERCHANNEL	For every meter and channel combination, find all duplicates and report as exceptions. User-type analyses will exclude these duplicates from consideration. This is a system-type rule and cannot be edited or deleted by an end user.	None	Physical Meter and Channel	SYSTEM
ATTRIBUTE_MISMATCH_DUALSYSM_SERVICECHANNEL	At the service point and channel level, compare attributes between two systems and report each mismatch as an exception. This analysis is subject to a possible user supplied 'where clause' for each system. Duplicate service point and channel combinations are excluded from this analysis.	SOR and EXCPT	Service Point and Channel	USER
CONDITION_CHECK_SINGLESYSM	At the service point and channel level, find any records that satisfy the user supplied 'where clause' for the system-of-record, and report all found records as exceptions. The 'where clause' for the exception-system is ignored. Duplicate service point and channel combinations are excluded from this analysis.	SOR only	Service Point and Channel	USER
ENTITY_MISSING_EXCEPTION_METER	A meter is missing from the exception-system but is present in the system-of-record. This analysis is subject to a possible user supplied 'where clause' for each system. Duplicate meters are excluded from this analysis.	SOR and EXCPT	Physical Meter	USER
ENTITY_MISSING_SOR_SERVICECHANNEL	A service point and channel combination is missing from the system-of-record but is present in the exception-system. This analysis is subject to a possible user supplied 'where clause' for each system. Duplicate service point and channel combinations are excluded from this analysis.	SOR and EXCPT	Service Point and Channel	USER

11



Key Benefits



- One truth – clean, reliable and trust worthy data for BI solutions (dashboards, KPIs etc)
- Cost effective approach – no code maintenance for changing rules
- Data profiling tool
- Can be used in new proto type build project(s) to quickly identify SOR related issues
- Maintain History and fix operational issues
- Achieve better customer experience

13



Thanks for Attending!



Ravi Ananthanarayana

Ravishankar.Ananthanarayana@roltasolutions.com

Enterprise Data Architect – BI Practice,
ROLTA Solutions.

333 E Butterfield Rd.,
Lombard, IL – 60148.