ETL Process and Architectural Overview

Business Objects Server

BA Dashboard Pages

The BA Dashboard pages displays information about the Scorecards, Perspectives, Objectives, Key Performance Indicators (KPIs), and Metrics in components with various formats.

Contexts and Analytic Views

The Contexts provide business-oriented views of the data and act as an interface between the data warehouse and the components that display the data. The Contexts contain a schema of the tables and joins in the data warehouse. They also contain analytic views that ensure the consistency and quality of the data.

Contexts and Analytic Views

Web Application Server

BA KPI Engine

The Business Analytics KPI engine calculates the value of a KPI using the formula specified for the KPI and provides the status and score for the KPI using the thresholds specified for the KPI.

BA KPI Engine

Consolidation Layer

The Consolidated Entities page enables you to consolidate an entity that appears in different data sources, to prioritize the data sources, to select the relevant entity dimensions (columns), and to select when to use the consolidation.

Consolidation Table

Consolidation Layer

DW Target Database

Target Tables

The data warehouse target tables contain the data the applications use for reporting purposes. The data resides in standard fact and dimension tables.

Target Fact Tables

AACOST
ACTUALCOST_CSTM
ACTUALCOST
APCOST
ASSET
BACKUP_SESSION
BUSINESS_APP_PERF
BUSINESS_APP_STATUS
BUSPROCSERVICE
BUSTRANSSERVICE
CHANGE
COMPLIANCE
COMPLIANCE_STATUS
DAILY_ROI
DEFECTHIST
EMPLOYEEUTILIZATION
EMPLOYMENTSTAT
INCIDENT
INTERACTION
NODE_CONSUMPTION
NODE_METRIC
PERSONCOMPETENCY
PLANNEDCOST_CSTM
PLANNEDCOST
POLICY_REMEDIATION
POLICY_RESULT
PROJECTISSUE
PROJECTSNP
PROJECTTASK
PROJECT
REQUIREMENTHIST
REQUIREMENT_BRIDGE
RESTORE_SESSION
REVENUE
SERVICESTATUS
SERVICEUTILIZATION
SERVICE_PROVIDER_BILLING
SERVICE_PROVIDER_UTILIZATION
SLAOUTAGE
SLASTATUS
SOFTWARECOUNTSNP
SOFTWARECOUNT
STORAGE_EVENTS
STORAGE_PATH_COMPLIANCE
STORAGE_UTILIZATION
SURVEYSCORE
TESTHIST
TESTINSTANCE
TIME_TO_MARKET

Target Dimension Tables

APPLICATION
ASSET
BACKUP_POLICY
BUDGETLINE
BUDGET
BUSPROCESS
BUSTRANSACTION
CFGITEM
CHANGE
CI
CI_TOPOLOGY
COMPETENCY
COMPONENT
CONFIGPOLICYRESULT
CONFIGPOLICY
CONTRACT
COSTCATEGORY
COSTCENTER
CURRENCY
CYCLE
DEFECT
DEFECT_PLHD
DNFLAG
EMPLOYMENT
EXCHANGE
INCIDENT
INTERACTION
ITLFUNCTION
LOCATION
MEASUREITYPE
MODEL
NODE
OCFLAG
ORG
PERIOD
PERSON
POLICY
PROBJECTIVE
PROBLEM
PROGRAM
PROJECTISSUE
PROJECTTASK
PROJECT
REQUEST_DETAIL_PLHD
REQUEST
REQUEST_HEADER_PLHD
REQUEST_TYPE_UD_PLHD
REQUEST_USERDATA_PLHD
REQUIREMINT
REQUIREMENT_PLHD
SCENARIO
SERVICE_BLUEPRINT
SERVICE
SERVICE_INSTANCE
SLA
SOFTWARECOUNT
STATUS
STORAGE_VOLUME
SUBSCRIPTION
SURVEY
TEST
TEST_PLHD

Target Layer

DW Staging Database

Dimensional Transform (XFR)

The Transformation Layer contains load-ready tables in the structure of the target tables. Load-Ready New (XFRN) tables contain new records, Load-Ready Update (XFRU) tables contain records that need to be updated, and Load-Ready Delete (XFRD) tables contain records that have been flagged as deleted.

The transform processes extract data from the MSC tables, convert it into a dimension and fact format that is compatible with the target database, and load it into the load-ready target staging tables.

XFRN
XFRU
XFRD

Transformation Layer

Multiple Source Identity (MSI)

The Multiple Source Identity (MSI) tables contain data from the SSI tables with the addition of the enterprise key.

MSI

Single Source Identity (SSI)

The Single Source Identity (SSI) processes restructure the diverse source system data from the EXT tables into standard structures. The SSI tables contain this conformed data.

Since the data warehouse uses data from multiple source systems that do not use the same business keys for the same entities, the ETL process creates an internal enterprise key to link entity records. The ETL process uses data in the SSI tables to generate these enterprise keys, which are used during the consolidation process.

SSI
XREF

Identification Layer

Cross Reference (XREF)

The Cross Reference (XREF) tables contain descriptions of enterprise key to business key relationships, source entities, consolidated entities, and priority relationships. The ETL process uses this information to combine data records from different source models into a single record in the consolidation model.

Extract Staging Tables (EXT)

The Extract Staging (EXT) tables store data in database form. After validating the data, The ETL process loads data from the EXTRACT_FF and EXTRACT_DELETE_FF flat files and alternate source spreadsheets into the EXT tables. The EXT tables also contain pre-fill records from the Transactional Snapshot (TSNP). For each data source, The files are zipped and stored in an archive file.

EXT
TSNP

Extraction Layer

Transactional Snapshot (TSNP)

The Transactional Snapshot (TSNP) tables contain a version of the source data as it existed at the beginning of the previous ETL cycle. The ETL process uses the TSNP tables to identify deleted records by comparing the business keys in the EXTRACT_DELETE_FF flat files to the business keys in the TSNP tables. If a key exists in the TSNP tables but not in EXTRACT_DELETE_FF, the ETL process identifies the associated record as a deleted record and loads it into the EXT tables. The ETL process also uses the TSNP tables as part of the pre-fill process.

File System

EXTRACT_FF
EXTRACT_DELETE_FF

Flat File Extracts

The extract processes extract delta data from the source systems and spreadsheets and write it to delimited flat files on the file system. The flat files (EXTRACT_FF and EXTRACT_DELETE_FF) decouple the data warehouse from the source systems. EXTRACT_FF files contain records that were updated or inserted since the last time ETL ran. EXTRACT_DELETE_FF files contain the complete list of business keys from the source view and are used to identify delete records.

Data Collection Services (DCS)
Extractor Framework (DBServices)

Data Source Entities

PPM
ACTUALCOST, BUDGETLINE, EXCHANGE, LOCATION, ORG, PERSON, PLANNEDCOST, PROBJECTIVE, PROGRAM, PROJECT, PROJECTISSUE, PROJECTTASK, REQUEST, REQUEST_DETAIL_PLHD, REQUEST_HEADER_PLHD, REQUEST_TYPE_UD_PLHD, REQUEST_USERDATA_PLHD, and more.

SM
CFGITEM, CHANGE, INCIDENT, INTERACTION, LOCATION, MODEL, ORG, PERSON, SERVICE, SLA, SLAOUTAGE, SLASTATUS, and more.

ALM
CYCLE, DEFECT, DEFECTHIST, DEFECT_PLHD, DOMAIN_PROJECT, PROJECT, REQUIREMENT, REQUIREMENTHIST, REQUIREMENT_BRIDGE, REQUIREMENT_PLHD, REQUIREMENT_TYPE, TEST, TESTHIST, TESTINSTANCE, TESTINSTANCEHIST, TEST_PLHD, and more.

AWS
ACTUALCOST, SERVICE_PROVIDER_BILLING, and more.

CSA
BILLING, COMPONENT, OEG, PERSON, SERVICE_BLUEPRINT, SERVICE_INSTANCE, SERVICE_OFFERING, SUBSCRIPTION, and more.

AWSCW
SERVICE_PROVIDER_UTILIZATION, and more.

Source Databases

Data Integration Interface

For all data sources, the data warehouse accesses source data using a Data Collection Services (DCS) Extractor specific to the data source type. The integration is performed using the DCS mechanism.

Source Systems

The source systems contain the current source data. You can extract data from HP Project and Portfolio Management (PPM), HP Service Manager (SM), HP Application LifeCycle Management (ALM), Amazon Web Services (AWS), Amazon Web Services CloudWatch (AWSCW), and HP Cloud Service Automation (CSA). Some of the source systems reside in remote databases managed by an RDBMS.

The data warehouse uses the change data capture (CDC) technique for extracting data from the source tables. The ETL process extracts data if the last modification date is greater than or equal to the maximum last modification date value.

PPM
SM
ALM
AWS
CSA
AWSCW

© Copyright 2015. Hewlett-Packard Development Company, L.P.