Proper Architecture Considerations For Data Warehouse Implementation

Mark Nilan
Data Warehouse Product Manager
SAS Asia Pacific
Explaining The Intelligence Layer

Why Do we Need This?

How Do We Do This?
Agenda

How Do We Architect A Solution Properly?
- The Packaged DW Solutions That SAS Provides
- The Building Products For The Solution
- Architectural Blueprints For Data Quality & Transformation
- Layers Of Construction

Why Do We Need An Architectural Approach?
- Risk Reduction
- Simplicity & Efficiency
- Approach To Project Deliverables
- Metadata Management
Explaining The Intelligence Layer

Why Do We Need This?

How Do We Do This?
Packaged Architecture

Intelligence Value Chain

DATA

INTELLIGENCE
Intelligence Value Chain

DATA

Plan  ETL Q  IS  BI  AI

Plan  Quality ETL  Intelligent Storage  Business Intelligence  Analytic Intelligence
Intelligence Value Chain

Architecture Components

Rapid Warehouse Methodology & TIS, BIS, IIS

SAS MVA Server, MEA DBMS SPD Server

BASE SAS Or DWA Generated Code & SAS Cleanse

Enterprise Miner, ETS, OR, Insight & SAS Stat

Enterprise Guide WEB AF/WEB EIS SAS DA

DATA

INTELLIGENCE
Where Do We Start?

Use SAS Rapid Warehousing Methodology.

- Created from:
  - The collected experience of hundreds of SAS projects
  - Industry ‘best practice’.

- Successful because:
  - Business focused
  - Breaks project down into small, defined, manageable pieces rather than large chunks.
  - Iterative, flexible and dynamic
  - Uses Joint Application Development and Rapid Application Development delivery techniques
SAS® Intelligence Models

Vision

Rapid Warehouse Methodology & TIM, BIM

PLUG-IN COMPONENTS

Based on proven experience
Reduced risk of failure
Intelligence Solutions

There are currently three available:

- SAS® Telecommunications Intelligence Solution
- SAS® Banking Intelligence Solution
- SAS® Insurance Intelligence Solution
Assemble To Order

- Industry specific solutions - customizable to fit customer’s business environment
  - Analytical models
  - Dimensional models
  - Business rules
- Data Models - Logical and physical data models
- Methodologies - best practice guidelines
- System Integration Services - experienced consulting and implementation services
Architecture Components

DATA

- Rapid Warehouse Methodology
- BASE SAS Or DWA Generated Code

Plan

ETLQ

IS

BI

AI

SAS MVA Server, MEA DBMS SPDS

INTERNELIGENCE

Enterprise Miner, ETS, OR, Insight & SAS Stat

Enterprise Guide WEB AF/WEB EIS SAS Intrnet
Mart Driven Enterprise Data Warehouse

Operational Environment

Warehouse Environment

Enterprise Data Warehouse

Finance

Marketing

Sales

Personal Data Marts

Business Unit-Specific Data Marts

Operational Data Store

Inventory

Product

Orders

Contacts

Billing

ERP & Business Systems

External Unstructured Data

Operational Environment

Warehouse Environment

Business & Technical Metadata

Copyright © 2002, SAS Institute Inc. All rights reserved.
Data is extracted from the source systems using the appropriate extraction adaptor. Some data will be operational system dumps and will reside in a RAW data area.

The ODD Area contains SAS views of RBDMS tables or SAS tables. CDC contains SAS tables of the previous loads data for comparison in change data capture. Or EAI queues, log captures or whatever method…

Exceptions area will contain all erroneous records in related tables from any of the validation processes.

STAGING Area contains Tables that are waiting for other Tables to complete. The STAGING Area contains Tables that are waiting for other Tables to complete.

Enterprise Area contains the Fact and Dimension tables of the Historic Normal Form Schema. The tables at this point are at source granularity.

The SUMMARY Mart area stores system summaries of Enterprise layer tables. It is the key part of the Intelligence Architecture.
So a real life Architecture may be:
But Why Use An Architected Approach?

Why Do we Need This?

How Do We Do This?
Lower Risk

The SAS Intelligence Architecture is a low-risk framework that integrates our industry-leading capabilities in data warehousing, advanced analytics, and business intelligence to provide...

an intelligence framework that can grow and adjust with your enterprise.
Mart Driven Enterprise Data Warehouse

Operational Environment

Warehouse Environment

ERP & Business Systems
External Unstructured Data

Operational Data Store
Enterprise Data Warehouse

Business Unit-Specific Data Marts
Personal Data Marts

Finance
Marketing
Sales

Operational Systems Departmental data
Metadata

Personal data marts
External data
Metadata

WEB
WAP telephone
PDA
Workstation
WWW browser

Finance
Marketing
Sales
Proper Intelligence Architecture

WHY
Because it is the foundation –
solidarity is required in its nature…

A gateway to all the functionality of
the intelligence layers above it

The foundation for quality

Repeatable, “assembled to order”
yet unique

ROI measurement as a base is
easier to compare to other packaged
DW and Datamarts
Aspects of Warehousing Projects

- Architecture
- Methodologies
- Tools
- People/Skills
Aspects of Warehousing Projects
Metadata Management Architecture
The Metadata Lifesaver?

- Common metadata architecture is also of paramount importance…

- SAS V8 Hub & Spoke is common across the Blueprint at all levels

- SAS V9 supports CWM standard XML based metadata

- Without common metadata, the successful data warehouse cause is lost!
Two questions for consideration…

**How Do I Properly Architect A Solution?**

- Use SAS. SAS is moving from a ‘build to order’ mentality to an assemble to order architectural approach.
- SAS offers proven rapid DW methodology & Intelligence Solutions in Banking, Telco, Insurance. Other industries coming.
- Architecture includes not just design, it is an instruction list of how to construct & extract source data, assemble changed data, staged data, detail data, summary data and data marts.
- The design & instruction includes the ‘plumbing’ too. That is Enterprise Application Integration, like MQ series, DCOM & CORBA agent technologies and traditional TCP.
Why Do I Need An Architectural Approach?

- Given that the ‘HOW’ consisted of a blueprint design and process instructions to deliver, what benefit does this give me?
- Lower risk – follow the instructions in the methodology and industry solutions. This is an ‘on-ramp’ to success.
- Complex data warehousing is assembling pieces like design, processes, people & tools in an iterative fashion. Done once – done many times. This eases project implementation pains.
- In a sea of diversity, the metadata manager is a life preserver for the warehouse architect – it provides centralised clarity and control to provide management of the architectural approach.
Intelligence Solutions

Blueprint Design

Assemble To Order