



## **A Look at KSU's Progression Tracking System for Support of Retention, Progression, and Graduation**

Erik Bowe & Donna Hutcheson  
Georgia Summit 2006  
September 21, 2006 – 2:00pm

Topic: Data warehouse and Data mining

# [ Agenda ]

- Technically Speaking (How we technically pulled this project off)
  - What is a CIF?
  - What is an ODS?
  - How is the ODS Organized?
  - Possible ODS Uses
  - Technologies Used
  - Data Sources
  - How we handled retention and attrition
- Retention, Progress, and Graduation (RPG): The Progression Tracking System (PTS)
  - Demonstration
- Conclusion

# [ Technically Speaking ]

How we pulled this off:

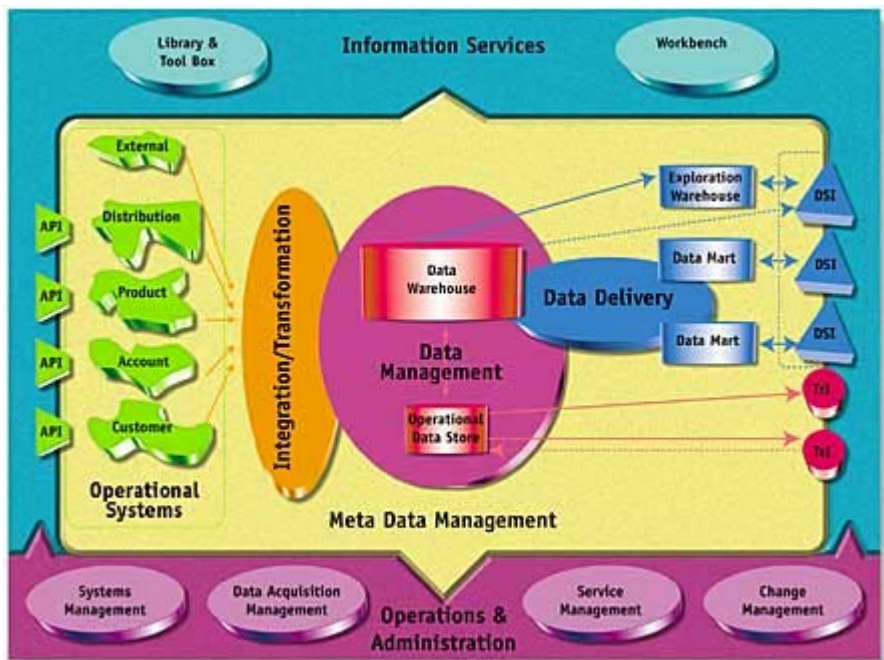
- Used an Operational Data Store (ODS)
  - Contains our Retention, Attrition, and Graduation derived values and raw data
- Used Oracle PL/SQL Packages
  - One for the I&T (integration and transformation) aka ETL (export-transform-load)
  - Another for the web-based interface for the end-user

# [ What is a CIF? ]

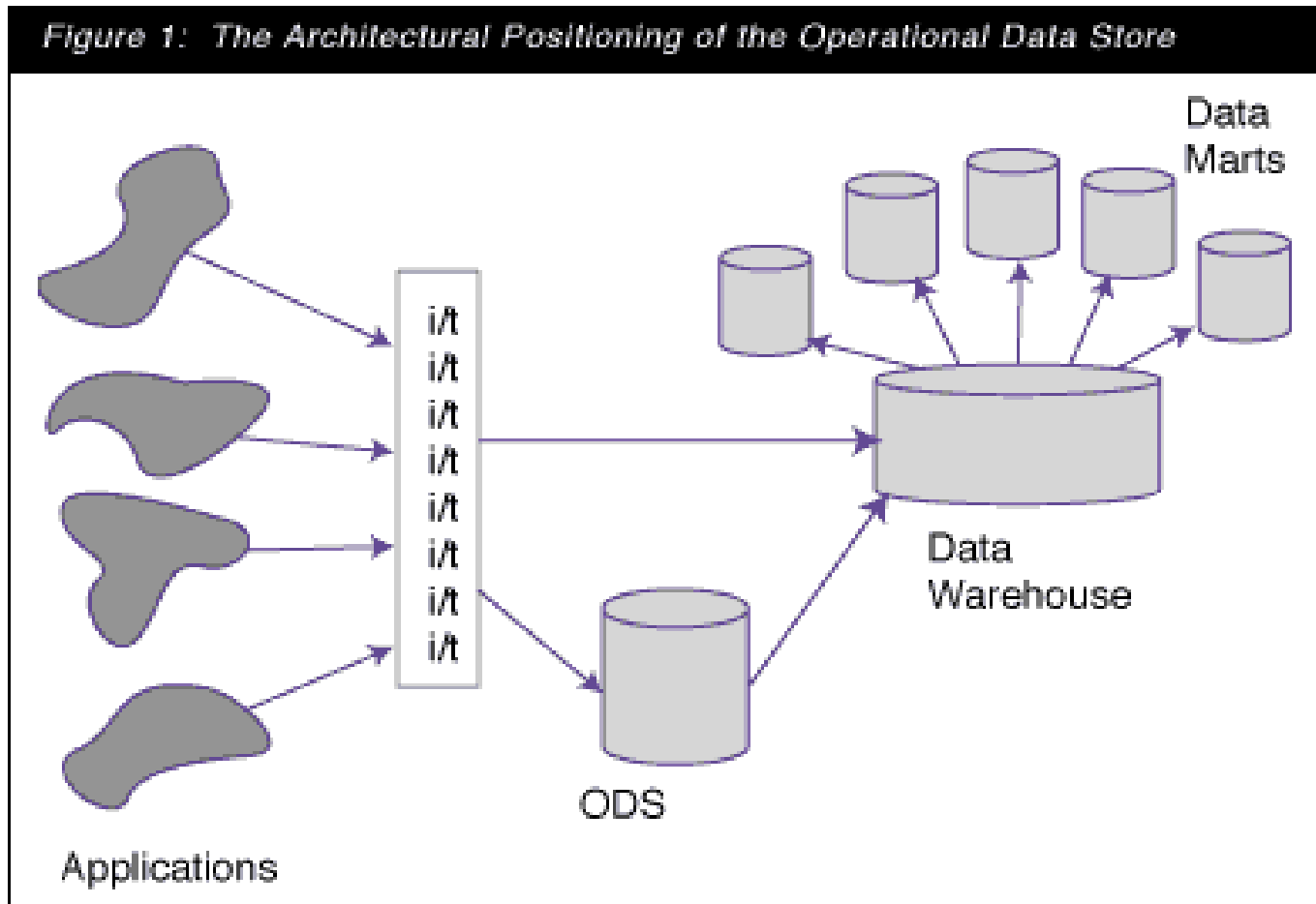
- The Corporate Information Factory (CIF) is a logical architecture whose purpose is to deliver business intelligence and business management capabilities driven by data provided from business operations.
  - The “business” being data about academics (i.e., assessment)
- The CIF has proven to be a stable and enduring technical architecture for any size enterprise desiring to build strategic and tactical decision support systems (DSSs).
- The CIF consists of producers of data and consumers of information.

# What is a CIF?

The CIF architecture:



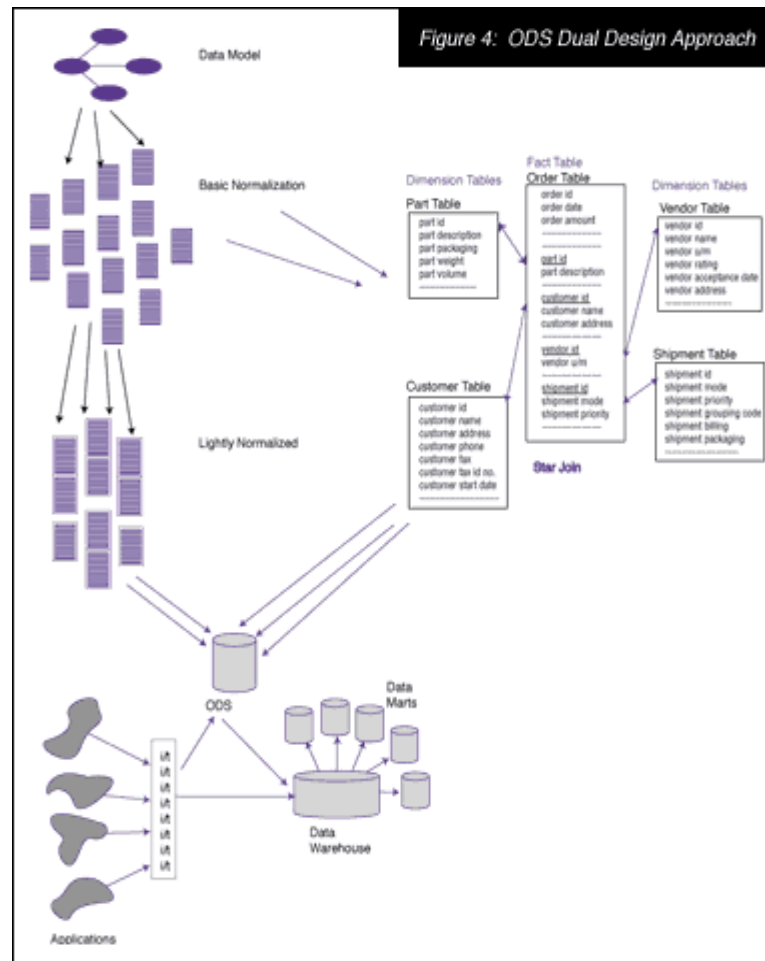
# What is an ODS?



Inmon W., Imhoff, C., and Sousa, R. (2002). *Corporate Information Factory* (2<sup>nd</sup> Ed.).

Inmon B. (1998). The Operational Data Store: Designing the Operational Data Store. *DM Review*.

# How is the ODS Designed?



Inmon W., Imhoff, C., and Sousa, R. (2002). *Corporate Information Factory* (2<sup>nd</sup> Ed.).

Inmon B. (1998). The Operational Data Store: Designing the Operational Data Store. *DM Review*.

# How is the ODS Designed?

*Figure 5: Appropriate Basis for Design*

	Update Farmer	DSS Farmer
Small Amount of Exploration	Lightly Normalized	Star Join
Occasional Exploration	Lightly Normalized	Lightly Normalized
Nearly all Exploration	Normalized	Normalized

The basis of the design of the ODS depends upon who the user is and the type of work the user is doing.

Inmon W., Imhoff, C., and Sousa, R. (2002). *Corporate Information Factory* (2<sup>nd</sup> Ed.).

Inmon B. (1998). The Operational Data Store: Designing the Operational Data Store. *DM Review*..

# Possible ODS Uses

At KSU:

- Enterprise Reporting (continuously)
  - i.e., the Business Intelligence, self-service model
- Fact Book (annually)
- Analytics
- Enterprise Data Warehouse (for elements not covered by Board of Regents')
- Data Collection

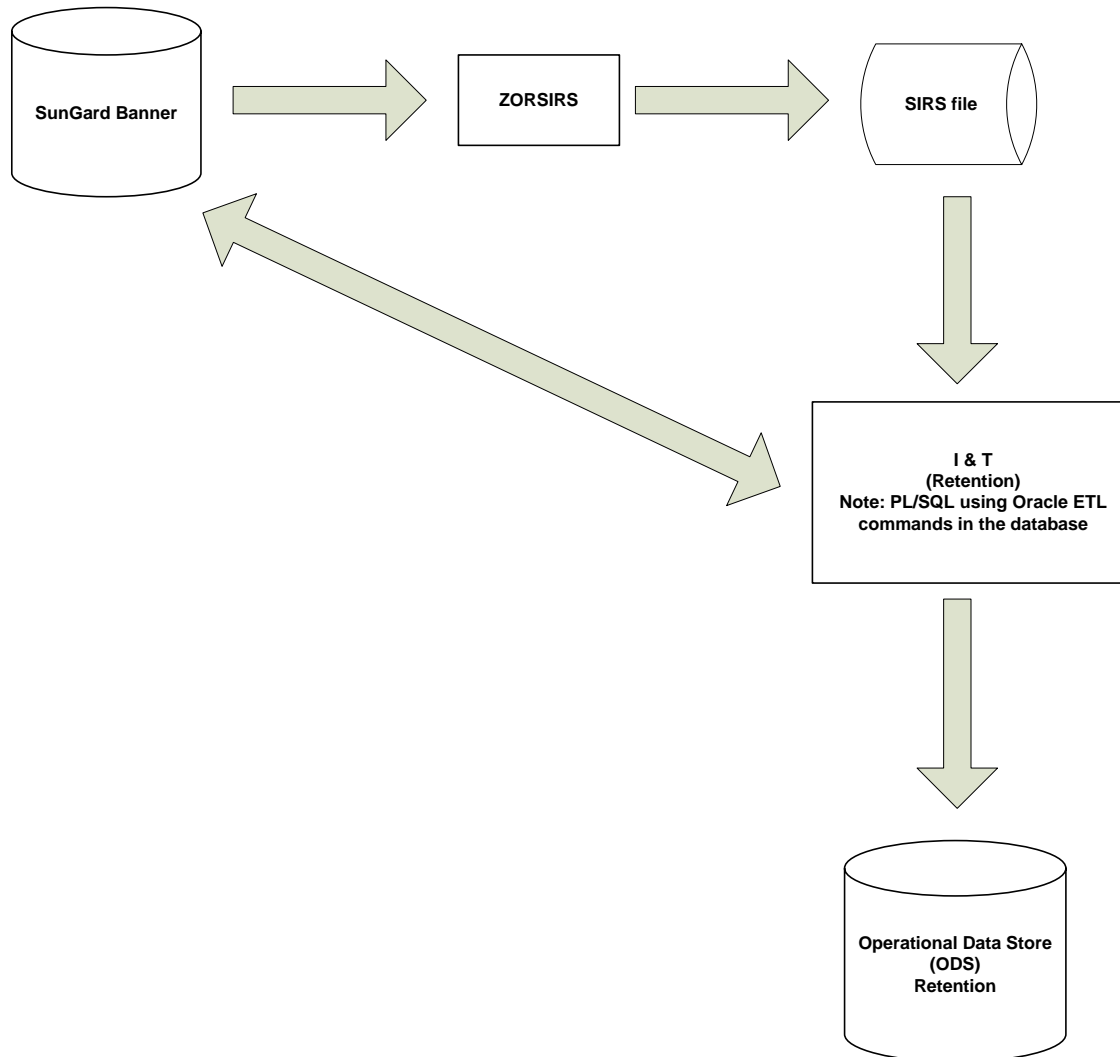
# [ Technologies Used ]

- Oracle Database 10g Enterprise Edition 10.1.4.0.2 With Partitioning, OLAP and Data Mining Options
  - Contains the following schemas:
    - Metadata Repository
    - Portal (with WebDAV)
    - Business Intelligence Discoverer End User Layer (EUL)
- Oracle Application Server 10g Release 2 (10.1.2.0.2)
  - Both infrastructure and middle tier components
- PL/PDF v1.2.4c
- Quest SQL Navigator 5

# [ Data Sources ]

- Enterprise Resource Planning Systems:
  - SunGard Higher Education Banner
    - Student record system
- Student Information Reporting System (SIRS)
  - Flat files reported to the Board of Regents' of the University System of Georgia

# Retention I&T Flow



# [ Handling Retention ]

PIDM/TERM	200308	200401	200508
123	X		X
456	X	X	X
789	X	X	

Assuming 200308 is the beginning cohort year...we followed this logic

1. The first step is to advance student-by-student through our SIRS data mart checking for the existence of a row in the next available term.
2. Repeat the process with #1 until all terms are exhausted.

We built a PL/SQL function to accomplish the task so it could be used in a SQL field list or WHERE clause.

# Handling Attrition

PIDM/TERM	200308	200401	200508
123	X		X
456	X	X	X
789	X	X	

Assuming 200308 is the beginning cohort year...we followed this logic

1. The first step is to advance student-by-student through our SIRS data mart checking for the non-existence of a row in the next available term.
2. Repeat the process with #1 until all terms are exhausted.
3. If data is needed for a given term, for example GPA, for which the student was not retained, we to get the data from the last term attended in our SIRS data mart.

We built a PL/SQL function to accomplish the task so it could be used in a SQL field list or WHERE clause.

# Demonstration

Retention, Progress, and Graduation (RPG):  
The Progression Tracking System (PTS)

The first strategic area addressed by the ODS  
at KSU! Why?

We had high demand for RPG data from  
across campus, and neither an application  
nor data mart with such info information  
existed in our software inventory.

# Improvements

- Move the integration & transformation (i.e., ETL) out of the code (i.e., PL/SQL)
  - Preferably a flexible product that allows the ETL to be expressed as predicates in the metadata
- Create the metadata as expressions
  - Looking currently at SAS products (May 8<sup>th</sup>, 2006)
- Use a 3<sup>rd</sup> party product for the front-end
  - Again, looking at SAS
  - Oracle Data Mining (ODM)
- Use of OLAP cubes for
  - Ethnicity/gender over time
  - Citizenship/gender over time
  - GPA ranges over time

# Conclusion

- The ODS is just one component of the Corporate Information Factory (CIF)
- The ODS data is
  - Current valued,
  - Detailed,
  - Volatile, and
  - Subject-oriented
- There are four different class of ODS
- Implementation is very time consuming



# Questions?

## Contact Information:

Erik Bowe, [ebowe@kennesaw.edu](mailto:ebowe@kennesaw.edu)  
Director, Information Management

Donna Hutcheson, [dhutches@kennesaw.edu](mailto:dhutches@kennesaw.edu)  
Associate Director, Institutional Research and Information  
Management