

Volume 2005-06

Published: June 23, 2005

www.dbaOnline.org

Email: admin@dbaOnline.org

Current Vol. Editor: Carol Sheng

DBA ONLINE

美國數據管理協會

Newsletter

*DBA Online - On the front line of database administration
DBA Online powers Oracle DBA*

DBA Online

Contents

Summery of DBAOnline SIG Seminar in June 2005	1
Summer Picnic News.....	3
Star Schema in Data Warehouse	4

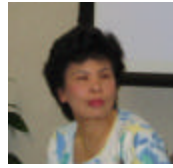
DBA Online SIG Seminar in June 2005



The DBAOnline has successfully held a Special Interest Group (SIG) seminar in new MetLife conference room in Edison, NJ on Saturday June 18, 2005. This SIG Seminar focused on Oracle data warehousing and was composed on two sessions, technical introductions and panel discussion. The SIG attracted more than 70 IT professionals. Four Oracle data warehousing experts were invited as our speakers and panelists to share with us there knowledge and real world experience on each aspect of the life cycle of data warehousing project. The SIG attracted more than 70 IT professionals



Mr. Jiangang Luo, president of New Jersey Chinese Computer Professionals Society and a Principal Consultant of Oracle Corporation with over 10 years of IT experience, specialized in Decision Support Systems and Business Intelligence applications, was our first guest panelist. In his 15-min introduction presentation, Jiangang give us the whole picture on data warehouse architecture so we could see the project from an architect's point of view.



Mrs. Jayne Chu, technical manager of the Advanced Technology Services consulting division of the Oracle Corporation, has led and managed data warehouse and data mart project teams with the responsibilities for all phases of Data Warehouse development life cycle. She is currently a project manager and the lead data modeler for the data warehouse project for one of the largest public school systems in California. She has a very extensive background in data modeling for large-scale data warehouses. During her brief presentation she shared with us about the considerations for Physical Models in Data Warehouses, explaining the advanced data warehouse models and justification of their application, the challenges of managing dimensions with changing data and some alternating approaches of managing large dimensions and they way to de-normalize a star model and benefits gained.



Mr. Steven Liu, Principle Consultant of Oracle Corporation, has over ten years of IT experiences and specializes in data warehouse & business intelligence. His expertise includes ETL, data warehouse architecture, VLDB and performance architecture. During his presentation, he gave us the overall introduction on the database data warehouse implementation, audience was really impressed by his concise slide which covers almost every part of the data warehouse implementation.



Mr. Qi Wang, Consultant Technical Manager of Oracle East Technology Service's Business Intelligence and Data Warehouse consulting practice, is responsible for implementing Oracle's Business Intelligence products, with specific emphasis on Express Server, Sales Analyzer, Financial Analyzer, and the 10g OLAP suite. For past 5 years with Oracle and 10 year prior to that, he has participated in all phases of consulting projects, and has extensive experience developing OLAP solutions. That is why he gave us the presentation on 'ROLAP' and 'MOLAP'. For those people only heard of the name but not really know much detail about the 'OLAP' implementation, Qi's presentation shared with us the knowledge and some practical tools they could practically us if they have to work on a data warehouse project in real life.

During the panel discussion session, the audience raised questions and the speakers answer them and explained in detail. Our panelists also collected many questions, which most members would commonly interested in, before hand and gave a detail answer in this SIG session to make the SIG more efficiency and informative. This

format received a rave review from our members.

Our president David Wang also joined our guest speakers to answer some questions, especially regarding the question on career direction of a current DBA, David gave some very insightful answer which we believe will clear many DBA's doubt.



Although the SIG was our first try this year to explore in-depth knowledge of a particular technical topic and open dialogues between our members and speakers, the result turned out to be really good and exceed our expectation. There were a lot of back-end jobs done by our committee members. Pauline and April wrote Chinese and English news long before the event. Wei prepared all the supply and water for our members. Warren spent a lot of his sleep time designed and prepared the professional-looking program brochure. Especially our host and speaker Mr. Qi Wang organized all the details of the event and spend numerous hours to co-ordinate it. It is a truly team effort under Qi's leadership. Of course all these won't happen if we don't have David as our default support for all the tasks involved. And also our thanks go to MetLife in Edison and for all the great help from Grace, Ana and Maggie along the event. Mostly, our thanks go to all members and guests who come to the SIG to learn and share on a wonderful weekend.

DBA Online Committee



✚ Summer Picnic, July 16

DBAOnline is going to organize another summer picnic on July 16, 2005, 10am-4pm, at Donaldson Park at Highland Park, New Jersey. It will be a Potluck style picnic. Please bring your favorite foods and share with others. DBAOnline will set up many entertainment activities like volleyball, poker game, playing cards game, Karaoke (if possible), chatting and networking, teamwork game, etc.

Please reserve July 16 (Saturday) for the fun time

Please go <http://co.middlesex.nj.us/parksrecreation/donaldson.asp> for more info about the Donaldson Park.

Please send email to admin@dbaonline.org for any suggestion and question you have for the picnic.

Please see the detail and register the picnic online at www.dbaonline.org when it is available.



✚ Star Schema in Data Warehouse

Provided by Carol Sheng

What is Star Schema?

Star Schema is a relational database schema for representing multi-dimensional data. It is the simplest form of data warehouse schema that contains one or more dimensions and fact tables. It is called a star schema because the entity-relationship diagram between dimensions and fact tables resembles a star where one fact table is connected to multiple dimensions. The center of the star schema consists of a large fact table and it points towards the dimension tables. The advantages of star schema are slicing down, performance increase and easy understanding of data.

Steps in designing Star Schema

- Identify a business process for analysis (like sales).
- Identify measures or facts (sales dollar).
- Identify dimensions for facts (product dimension, location dimension, time dimension, organization dimension).
- List the columns that describe each dimension. (region name, branch name, region name).
- Determine the lowest level of summary in a fact table (sales dollar).

Important aspects of Star Schema & Snow Flake Schema

- In a star schema every dimension will have a primary key.
- In a star schema, a dimension table will not have any parent table.
- Whereas in a snow flake schema, a dimension table will have one or more parent tables.
- Hierarchies for the dimensions are stored in the dimensional table itself in star schema.
- Whereas hierarchies are broken into separate tables in snow flake schema. These hierarchies help to drill down the data from topmost hierarchies to the lowermost hierarchies.

Glossary:

Hierarchy

A logical structure that uses ordered levels as a means of organizing data. A hierarchy can be used to define data aggregation; for example, in a time dimension, a hierarchy might be used to aggregate data from the Month level to the Quarter level, from the Quarter level to the Year level. A hierarchy can also be used to define a navigational drill path, regardless of whether the levels in the hierarchy represent aggregated totals or not.

Level

A position in a hierarchy. For example, a time dimension might have a hierarchy that represents data at the Month, Quarter, and Year levels.

Fact Table

A table in a star schema that contains facts and connected to dimensions. A fact table typically has two types of columns: those that contain facts and those that are foreign keys to dimension tables. The primary key of a fact table is usually a composite key that is made up of all of its foreign keys.

A fact table might contain either detail level facts or facts that have been aggregated (fact tables that contain aggregated facts are often instead called summary tables). A fact table usually contains facts with the same level of aggregation.

In the example figure 1.6, sales fact table is connected to dimensions location, product, time and organization. It shows that data can be sliced across all dimensions and again it is possible for the data to be aggregated across multiple dimensions. "Sales Dollar" in sales fact table can be calculated across all dimensions independently or in a combined manner, which is explained below.

- Sales Dollar value for a particular product
- Sales Dollar value for a product in a location
- Sales Dollar value for a product in a year within a location

Sales Dollar value for a product in a year within a location sold or serviced by an employee

Example of Star Schema: Figure 1.6

