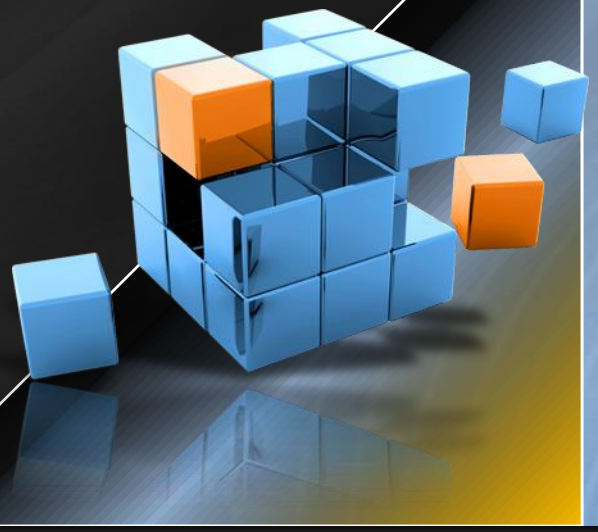
COMFRAME
Robert C. Cain, MVP

COMFRAME
precision software innovation

The Decoder Ring for Data Warehousing / Business Intelligence With SQL Server



Microsoft GOLD CERTIFIED Partner

Data Management Solutions
Information Worker Solutions
Business Intelligence
Custom Development Solutions

BIRMINGHAM • NASHVILLE • BEIJING

Your Presenter

COMFRAME
precision software innovation

- Robert C. Cain
- Senior BI Architect COMFRAME
  - Offices in Birmingham, Nashville & Beijing
  - Microsoft Partner, 1 of 37 Nationally Managed
  - Systems Integrator
  - App Dev, SharePoint, BI, EPM
- 5 years BI at Nuclear Power Co.
- 10 years as a consultant in the B'ham Market
- Wide range of .Net applications, ASP & Win
- SQL Server Data Warehouse
- http://arcanecode.com
- http://twitter.com/arcanecode

Why learn about DW/BI?

COMFRAME
precision software innovation

- DBA
  - Implement new Data Warehouse project
  - Install and configure BI tools (SSIS, SSRS, SSAS)
- DB Designer / Developer
  - Design and script a DW
  - DW significantly different from traditional database designs.
- Software Developer
  - Interact with warehouses for additional data or reports
  - Data mining results into your applications

Business Intelligence is HOT

COMFRAME
precision software innovation

- According to Computerworld, BI is the 5th hottest IT Skill for 2009
- Dice.com over 2,800 job openings

What is a Data Warehouse

COMFRAME
precision software innovation

- A giant storehouse for your data
- ALL of your data
- Aggregation of data from multiple systems

What is Business Intelligence

COMFRAME
precision software innovation

- Leveraging data you already have to convert knowledge into informed actions
- Providing ways to measure the health of your business
- Examining the data in your warehouse to look for three main areas of interest.
- Aggregations
- Trends
- Correlations (Data Mining)

Why Have a Data Warehouse?

COMFRAME
precision software innovation

- Combine data from multiple systems and resolve inconsistencies between those systems
- Make reporting easier
- Reduce the load on production systems
- Provide for long term storage of data
- Provide consistency among system transitions

Some More Reasons for a Data Warehouse

COMFRAME
precision software innovation

- Make the data available for analysis
- Ability to apply advanced data mining tools
- To extract further value from the data you already own

Business Intelligence

What's wrong with reporting from a Transactional System?

COMFRAME
precision software innovation

- OLTP – On Line Transaction Processing
- Designed for working with single record at a time.
- Data is highly "normalized", i.e. duplicate values have been removed.
- Getting all data for a record can involve many table joins
- Can be quite confusing for 'ad-hoc' reporting
- Can also be slow, having an impact on the OLTP system

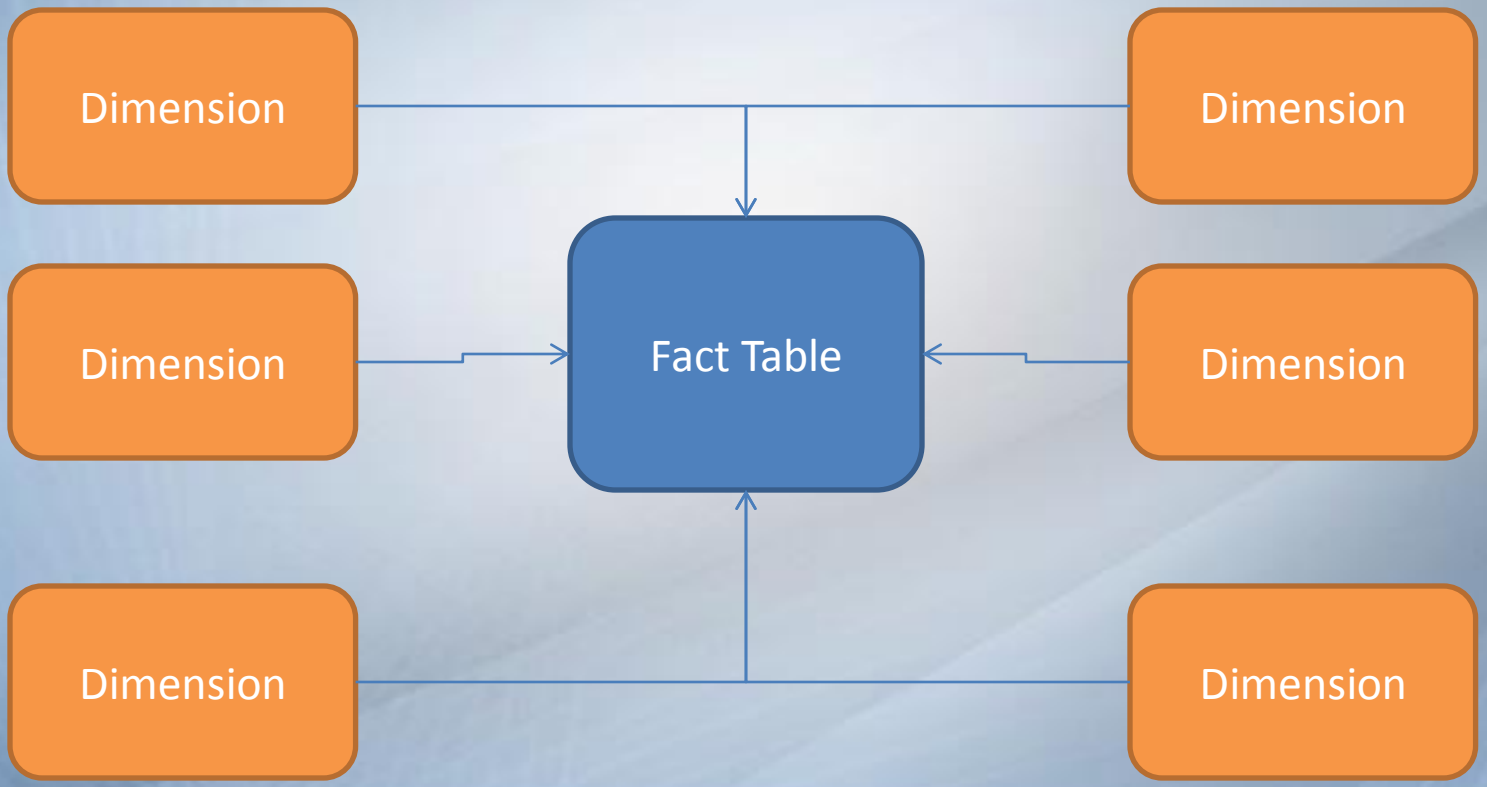
What's different about a Data Warehouse?

COMFRAME
precision software innovation

- Data Warehouses typically use a design called OLAP
- On-Line Analytical Processing
- Data is de-normalized into structures easier to work with.
- Number of tables are reduced, reducing number of joins and increasing simplicity
- Often a Star Schema or Snowflake Schema

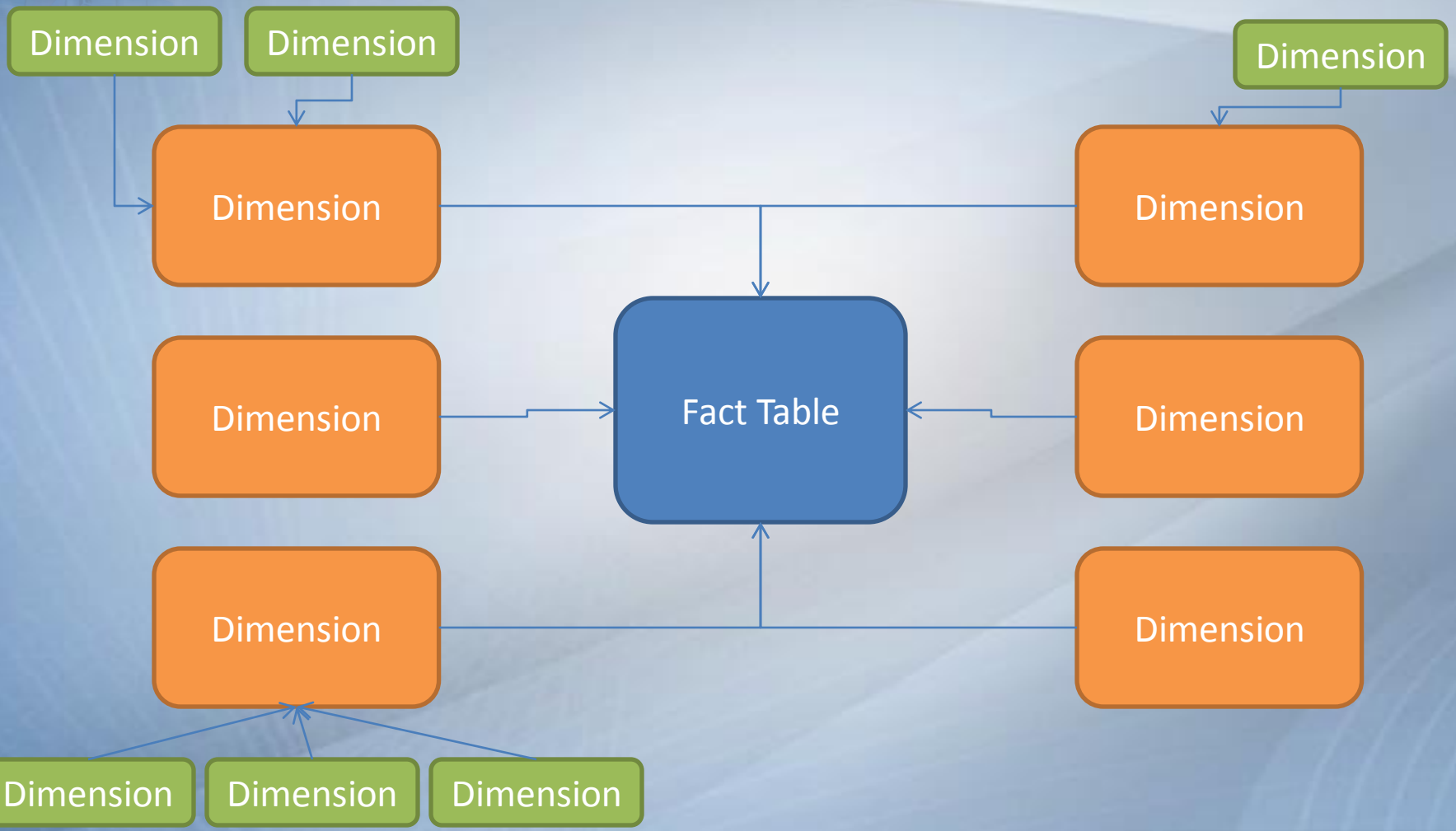
Star Schema

COMFRAME
precision software innovation



Snowflake Schema

COMFRAME
precision software innovation



Types of Tables in a Warehouse

COMFRAME
precision software innovation

- Facts
- Dimensions
- Both require the concept of Surrogate Keys
- A new key, typically some type of INT, that is used in place of any other key as the Primary Key

Reasons for Surrogate Keys

COMFRAME
precision software innovation

- Preserve data in case of source system change
- Combine data from multiple sources into a single table
- Source System keys can be multi-column and complex, slowing response time
- Often the key is not needed for many data warehousing functions such as aggregations

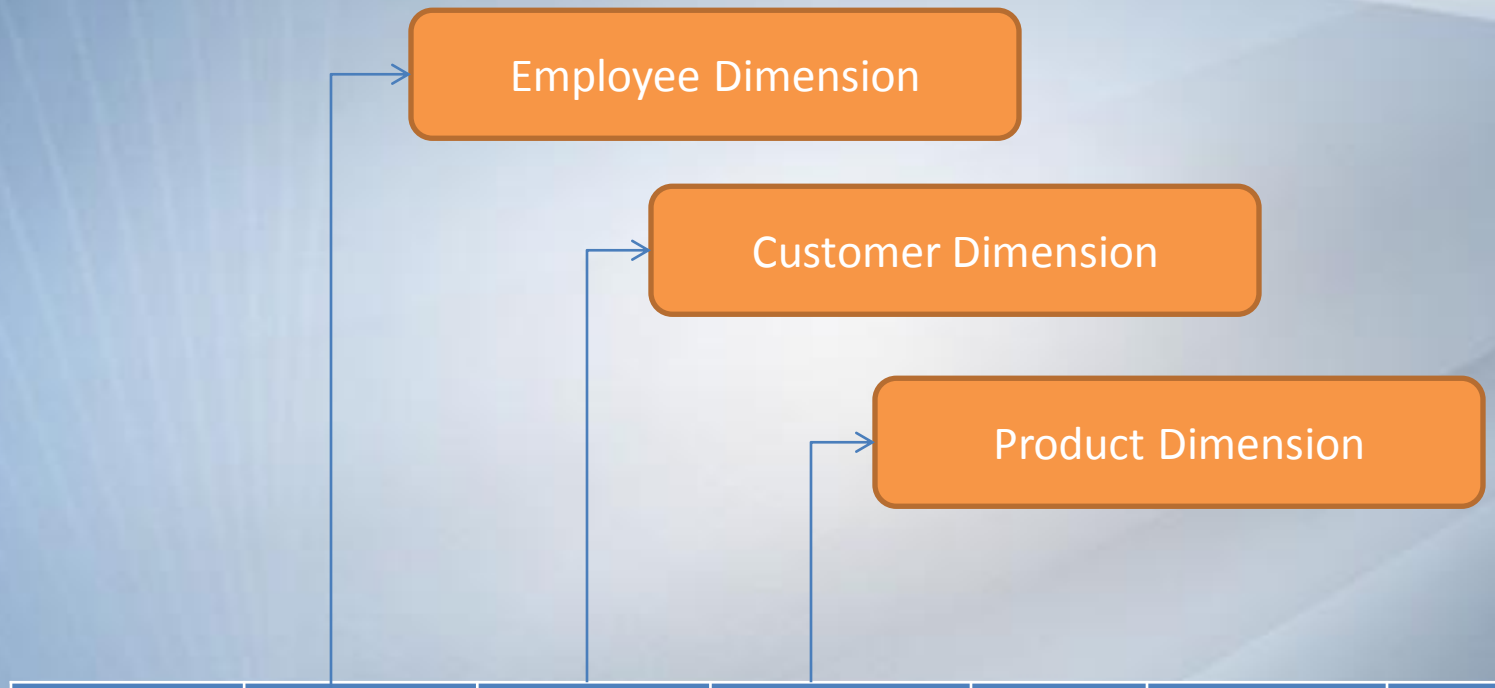
Fact Tables

COMFRAME
precision software innovation

- A Fact marks an event, a discrete happening in time
- Facts join dimensions, “who” did an action (SoldBy, SoldTo) to the “what”, what object was acted upon (Product).
- Facts also hold numeric measures to quantify the fact: Quantity, SaleAmt, SaleDate

Fact Table Example - Sales

COMFRAME
precision software innovation



ID	SoldByID	SoldToID	ProductID	Qty	SaleAmt	SaleDate
3456	1234	6789	987	3	156.00	7/17/2009

Dimensions

COMFRAME
precision software innovation

- Dimensions hold the values that describe facts
- “Look Up Values”
- Some examples: Time, Geography, Employees, Products, Customers
- When a Dimension can change over time, it’s known as a Slowly Changing Dimension
- Dimensions typically apply to the column, not the row.
- Many types of Dimensions

Type 0 Dimension (Fixed)

COMFRAME
precision software innovation

- Type 0 Dimensions are also called Fixed
- For data that will not change. Ever.
- Best used for static data like colors, sizes, etc.

ID	Description
1	Blue
2	Black
3	Green
4	Yellow

Type 1 Dimension

COMFRAME
precision software innovation

- When a dimensions value is updated, the old one is simply overwritten

Original Value

ID	EmployeeID	Last	First
1234	PQ1894958	McGillicutty	Hortence

New Value

ID	EmployeeID	Last	First
1234	PQ1894958	Hollywoger	Hortence

Type 2 Dimension

COMFRAME
precision software innovation

- When a dimension is changed, a new record is inserted and old one dated

Original Value

ID	EmployeeID	Last	First	FromDate	ThruDate
1234	PQ1894958	McGillicuty	Hortence	12/1/1998	<NULL>

New Value

ID	EmployeeID	Last	First	FromDate	ThruDate
2468	PQ1894958	Hollywoger	Hortence	7/6/2008	<NULL>
1234	PQ1894958	McGillicuty	Hortence	12/1/1998	7/5/2008

Type 3 – Just Say NO

- When a dimensions value is updated, a new column is added

Original Value

ID	EmployeeID	Last1	First
1234	PQ1894958	McGillicutty	Hortence

New Value

ID	EmployeeID	Last1	Last2	First
1234	PQ1894958	Hollywoger	McGillicutty	Hortence

- Almost never used

Type 4 Dimension

COMFRAME
precision software innovation

- When a dimension is changed, a old record is copied to history table and current one updated

Original Value in DimEmployee

ID	EmployeeID	Last	First
1234	PQ1894958	McGillicuty	Hortence

New Value in DimEmployee

ID	EmployeeID	Last	First
1234	PQ1894958	Hollywoger	Hortence

New Value in DimEmployee_History

ID	EmployeeID	Last	First	FromDate	ThruDate
1234	PQ1894958	McGillicuty	Hortence	12/1/1998	7/5/2008

Type 4 Dimension (Another Way)

COMFRAME
precision software innovation

- When a dimension is changed, old record is updated in history table, current one copied in

New Value in DimEmployee

ID	EmployeeID	Last	First
1234	PQ1894958	Hollywoger	Hortence

New Value in DimEmployee_History

ID	EmployeeID	Last	First	FromDate	ThruDate
1234	PQ1894958	Hollywoger	Hortence	7/6/2008	<NULL>
1234	PQ1894958	McGillicuty	Hortence	12/1/1998	7/5/2008

Different Dimension Types in a Table

COMFRAME
precision software innovation

- Often a single row holds multiple Dimensional Types
- The business should determine what data is significant enough to track changes on.

Example

ID	EmployeeID	Last	First	HrsLastMo	FromDate	ThruDate
1234	PQ1894958	McGillicuty	Hortence	200	12/1/1998	<NULL>

- Hours Last Month = Type 1
- Last Name = Type 2

Different Dimension Types in a Table

COMFRAME
precision software innovation

Original Value

ID	EmployeeID	Last	First	HrsLastMo	FromDate	ThruDate
1234	PQ1894958	McGillicuty	Hortence	200	12/1/1998	<NULL>

Update to Hours Last Month (Type 1)

ID	EmployeeID	Last	First	HrsLastMo	FromDate	ThruDate
1234	PQ1894958	McGillicuty	Hortence	280	12/1/1998	<NULL>

Update to Last Name (Type 2)

ID	EmployeeID	Last	First	HrsLastMo	FromDate	ThruDate
1234	PQ1894958	McGillicuty	Hortence	200	12/1/1998	4/22/2010
6789	PQ1894958	Hollywoger	Hortence	200	4/23/2010	<NULL>

Conformed Dimensions

COMFRAME
precision software innovation

- When pulling in data from multiple systems, you often have to reconcile different primary keys.
- This process is known as conforming your dimensions.

ID	Product	InventoryID	PurchasingID	WorkMgtID
9876	Widget	459684932	Wid45968	602X56VV1

Dimensions in a Star Schema

COMFRAME
precision software innovation

ID	SoldByID	SoldToID	ProductID	Qty	SaleAmt	SaleDate
3456	1234	6789	987	3	156.00	7/17/2009

Column	Value
ProductID	987
BusinessID	SHBL4X
Description	Knit Shirt
Color	Blue
Size	4XL
Sleeve	Long

- Flattened model easier to link
- Simpler for ad-hoc reporting
- Takes more database space
- Updates take more work (4XL to XXXXL requires 1 update per record)

Dimensions in a Snowflake Schema

COMFRAME
precision software innovation

ID	SoldByID	SoldToID	ProductID	Qty	SaleAmt	SaleDate
3456	1234	6789	987	3	156.00	7/17/2009

Column	Value
ProductID	987
BusinessID	SHBL4X
Description	Knit Shirt
Color	2
Size	7
Sleeve	2

ID	Value
1	Red
2	Blue
3	Green
…	…

ID	Value
6	3XL
7	4XL
8	5XL
…	…

ID	Value
1	Short
2	Long

- Takes less database space. Linked Dimensions reusable .
- Easier to update (change 4XL to XXXXL requires 1 Update)
- More difficult as it requires many links
- More links makes it difficult for ad-hoc reporting (views can help with this)

Getting Data Into A Warehouse

COMFRAME
precision software innovation

- ETL
- Extract
- Transform
- Load
- SSIS – SQL Server Integration Services

Getting Data Out of Your Warehouse

COMFRAME
precision software innovation

- SSRS – SQL Server Reporting Services
- SSAS – SQL Server Analysis Services

KPI

COMFRAME
precision software innovation

- Key Performance Indicators
- Dashboards
- Quick, at a glance indicator of system health

Region	Sales (USD)	Trending	Status
US	482m		
Europe	399m		
Asia	123m		
South America	225m		

Warehousing Methodologies

COMFRAME
precision software innovation

- Inmon – Bill Inmon - Top down
- Kimball – Ralph Kimball - Bottom up

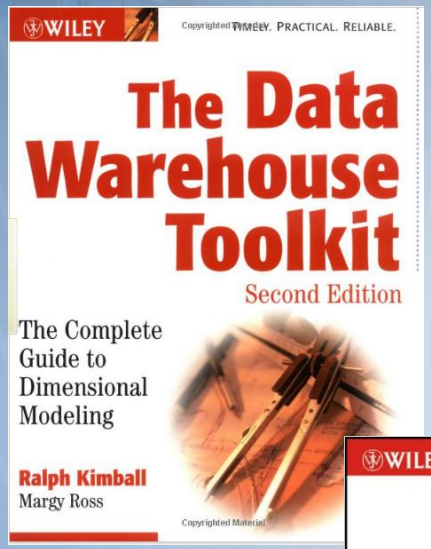
Demos

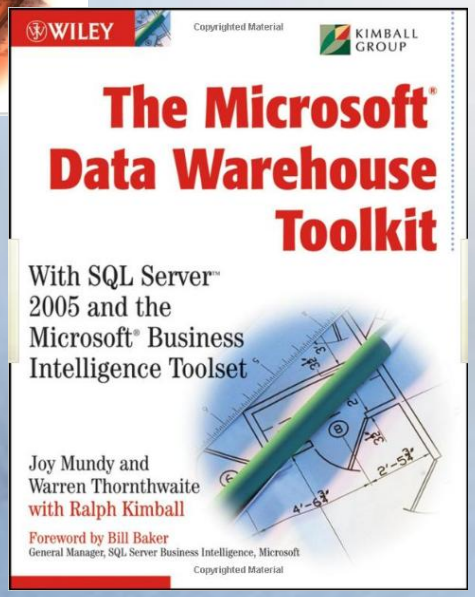
COMFRAME
precision software innovation

- Demo Site - http://www.codeplex.com/Wiki/View.aspx?ProjectName=SqlServerSamples
- Sample DBs - http://www.codeplex.com/MSFTDBProdSamples/
- SSAS - http://msftasprodsamples.codeplex.com/Release/ProjectReleases.aspx?ReleaseId=18652
- SSIS - http://www.codeplex.com/MSFTISProdSamples/
- SSRS - http://www.codeplex.com/MSFTRSProdSamples/

Resources

COMFRAME
precision software innovation

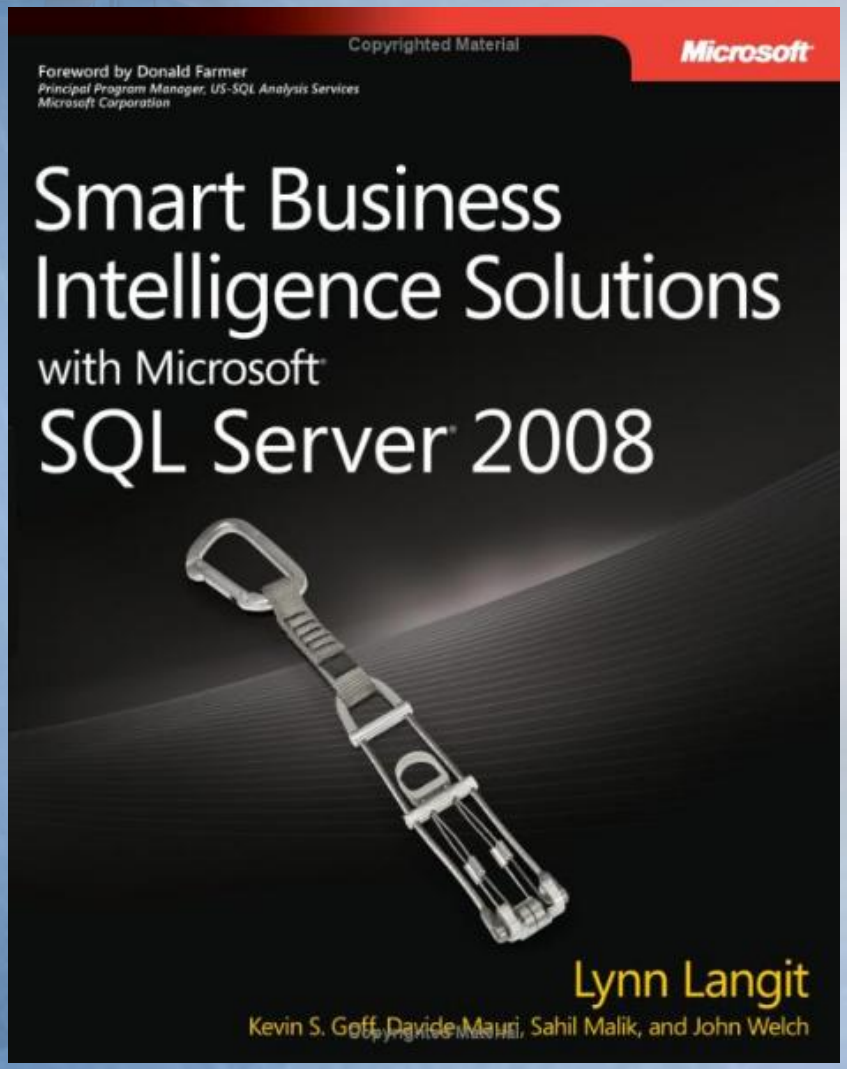




The Data Warehouse Toolkit
and
The Microsoft Data Warehouse Toolkit
by the Kimball Group

- http://www.amazon.com/Data-Warehouse-Toolkit-Complete-Dimensional/dp/0471200247/ref=pd_bbs_sr_1?ie=UTF8&s=books&qid=1239580212&sr=8-1
- http://www.amazon.com/MicrosoftData-Warehouse-Toolkit-MicrosoftBusiness-Intelligence/dp/0471267155/ref=sr_1_fkmr0_1?ie=UTF8&qid=1264636802&sr=8-1-fkmr0

Resources

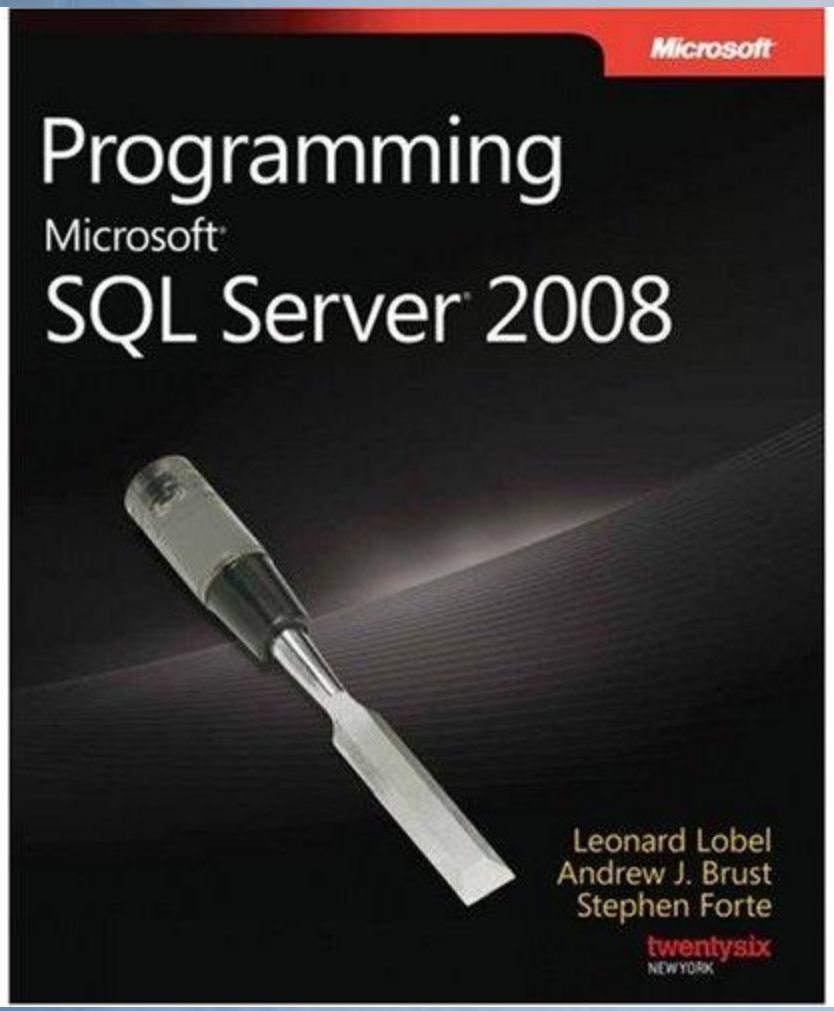


Smart Business Intelligence Solutions with Microsoft SQL Server 2008

- http://www.amazon.com/Business-Intelligence-Solutions-Microsoft%C2%AE-PRO-Developer/dp/0735625808/ref=sr_1_1?ie=UTF8&s=books&qid=1239580654&sr=1-1

Resources

COMFRAME
precision software innovation

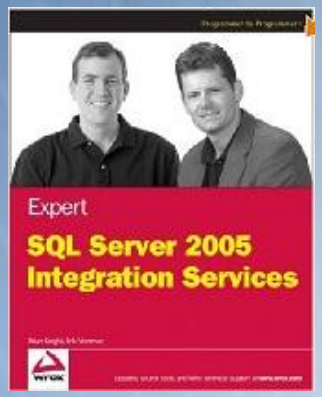


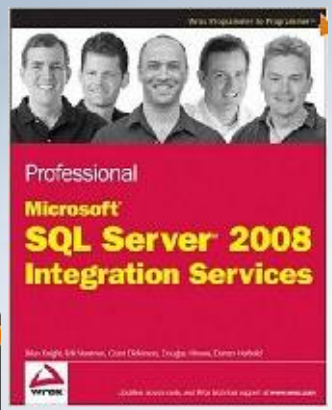
Programming Microsoft SQL Server 2008

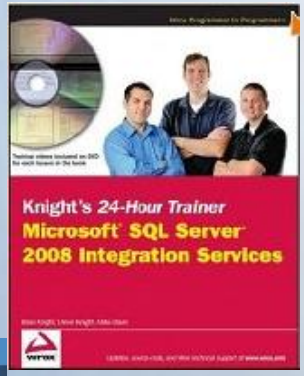
- http://www.amazon.com/Programming-Microsoft-Server-2008-PRO-Developer/dp/0735625999/ref=sr_1_1?ie=UTF8&s=books&qid=1239580376&sr=1-1

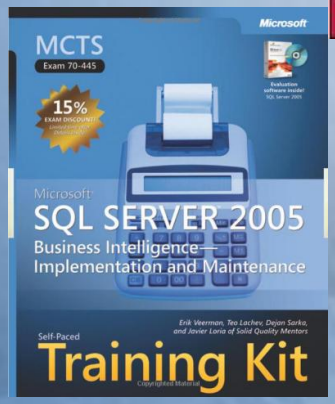
Resources - SSIS

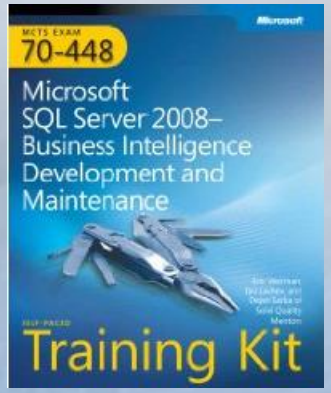
COMFRAME
precision software innovation





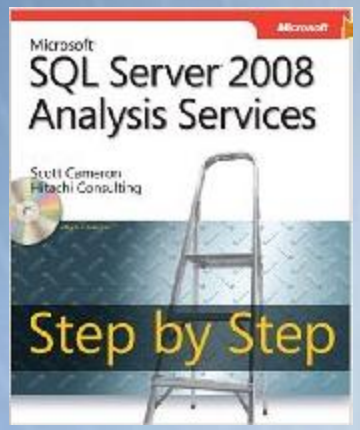


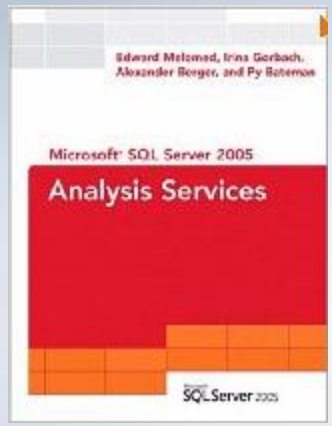


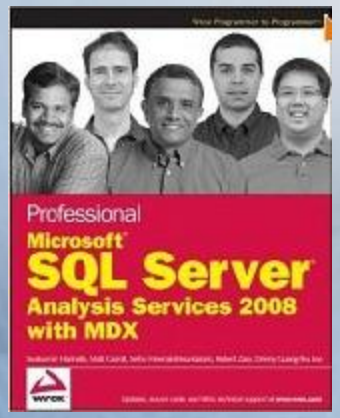


- Erik Veerman / Brian Knight Books
- http://www.amazon.com/Expert-Server-Integration-Services-Programmer/dp/0470134119/ref=sr_1_5?ie=UTF8&s=books&qid=1239833324&sr=8-5
- http://www.amazon.com/Professional-Microsoft-Integration-Services-Programmer/dp/0470247959/ref=sr_1_1?ie=UTF8&s=books&qid=1239833324&sr=8-1
- http://www.amazon.com/MCTS-Self-Paced-Training-Exam-70-445/dp/0735623414/ref=sr_1_7?ie=UTF8&s=books&qid=1239833324&sr=8-7#
- http://www.amazon.com/MCTS-Self-Paced-Training-Exam-70-448/dp/0735626367/ref=sr_1_4?ie=UTF8&s=books&qid=1239833324&sr=8-4
- http://www.amazon.com/reader/0470496924?_encoding=UTF8&ref_=sib_dp_pop_fc&page=1#reader

Resources - SSAS



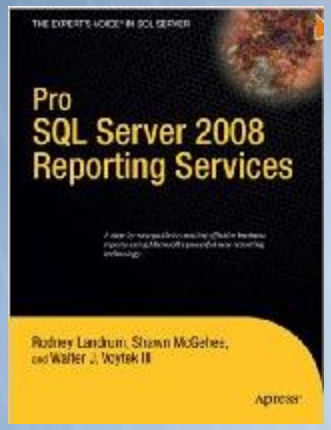


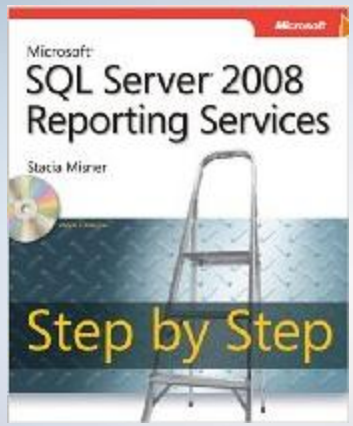


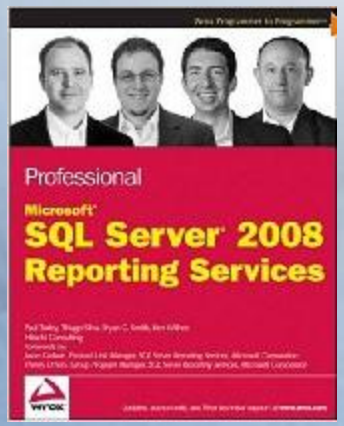
- http://www.amazon.com/Microsoft%C2%AE-Server%C2%AE-Analysis-Services-Microsoft/dp/0735626200/ref=sr_1_5?ie=UTF8&s=books&qid=1252100419&sr=1-5
- http://www.amazon.com/Microsoft-Server-2005-Analysis-Services/dp/0672327821/ref=sr_1_2?ie=UTF8&s=books&qid=1252100419&sr=1-2
- http://www.amazon.com/Professional-Microsoft-Analysis-Services-Programmer/dp/0470247983/ref=sr_1_4?ie=UTF8&s=books&qid=1252100419&sr=1-4

Resources - SSRS

COMFRAME
precision software innovation







- http://www.amazon.com/Pro-Server-2008-Reporting-Services/dp/1590599926/ref=sr_1_11?ie=UTF8&s=books&qid=1252100749&sr=1-11
- http://www.amazon.com/Microsoft%C2%AE-Server%C2%AE-Reporting-Services-Microsoft/dp/0735626472/ref=sr_1_10?ie=UTF8&s=books&qid=1252100749&sr=1-10
- http://www.amazon.com/Professional-Microsoft-Reporting-Services-Programmer/dp/0470242019/ref=sr_1_15?ie=UTF8&s=books&qid=1252100793&sr=1-15

Resources

COMFRAME
precision software innovation

Blogs

SSIS Junkie - http://blogs.conchango.com/jamiethomson/default.aspx
Brian Knight - http://pragmaticworks.com/community/blogs/brianknight/default.aspx

Podcast

SQL Down Under - http://www.sqldownunder.com/PreviousShows/tabid/98/Default.aspx
SQL Share (formerly JumpstartTV) – http://sqlshare.com

Forums

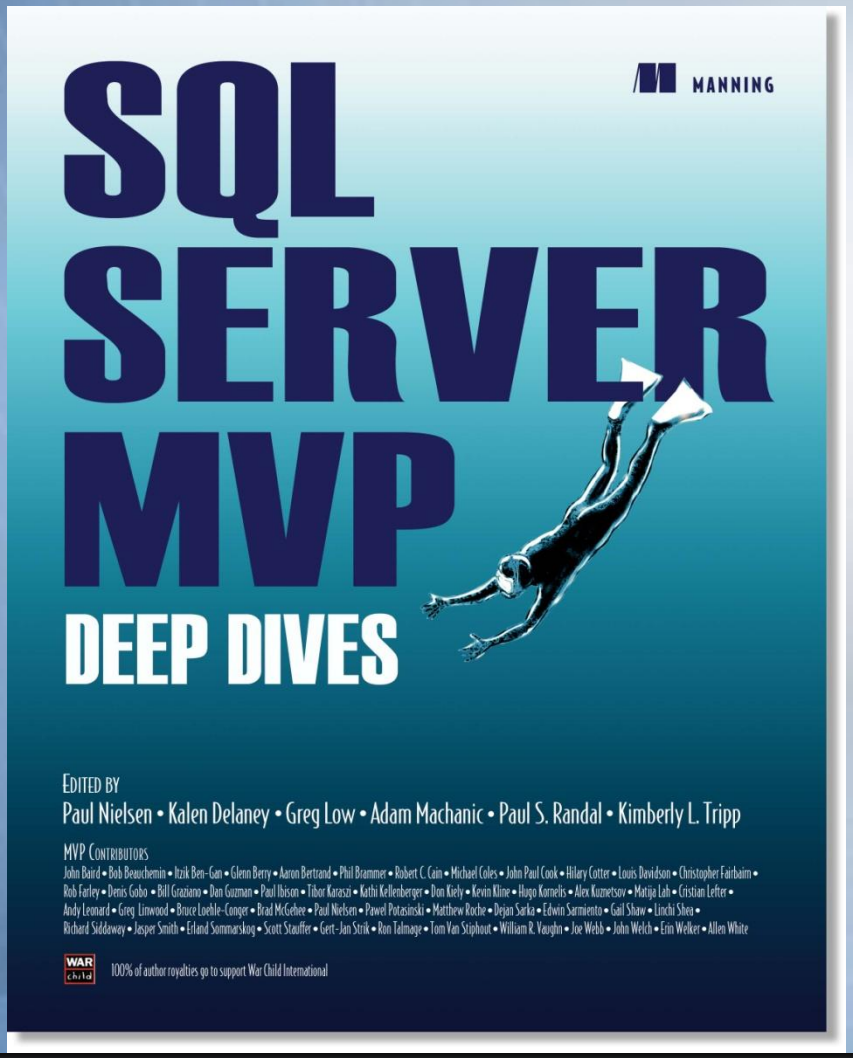
MSDN SSIS Forum - http://forums.microsoft.com/MSDN/ShowForum.aspx?ForumID=80&SiteID=1

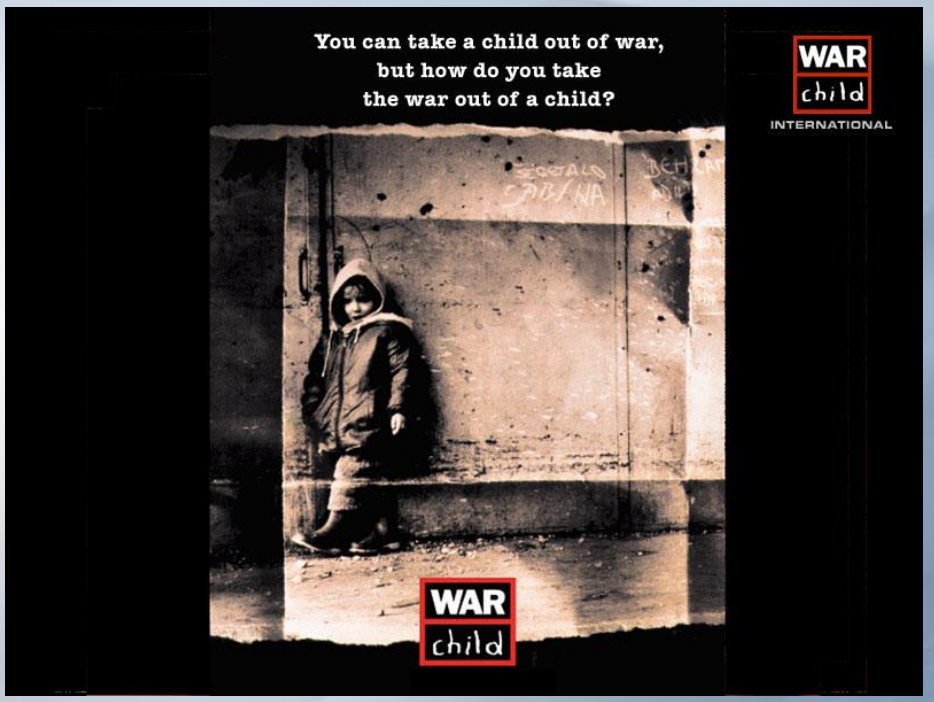
Other

Microsoft BI Site - http://www.microsoft.com/events/series/bi.aspx
Wikipedia Article - http://en.wikipedia.org/wiki/Data_Warehouse
SQL Serverpedia - http://sqlserverpedia.com/
COMFRAME BI – http://www.comframe.com/bi

War Child International

COMFRAME
precision software innovation





http://www.warchild.org
http://www.SQLServerMVPDeepDives.com

Thanks Again!

COMFRAME
precision software innovation

- Questions?
- All material available at http://arcanecode.com
- rcain@comframe.com