BIXCHANGE

OLAP & A Show

Designing Highly Effective OLAP Solutions

DAMAN CONSULTING

BIXCHANGE

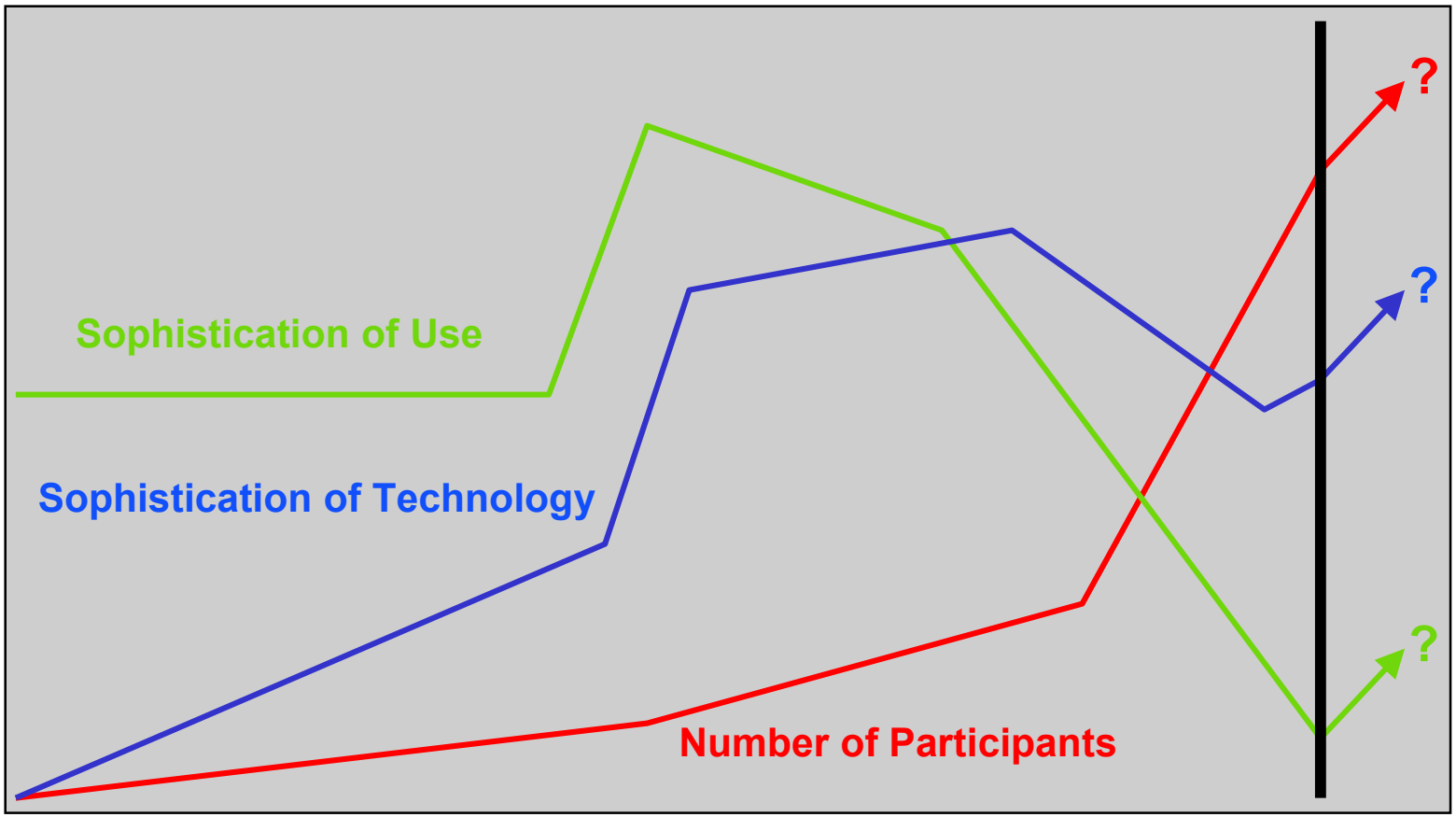
Contents

- ✔ **BI Trends**
- ✔ **Seven Habits**
- ✔ **Trivial Pursuit**
- ✔ **Characteristics of a Good Show**
- ✔ **Q&A**

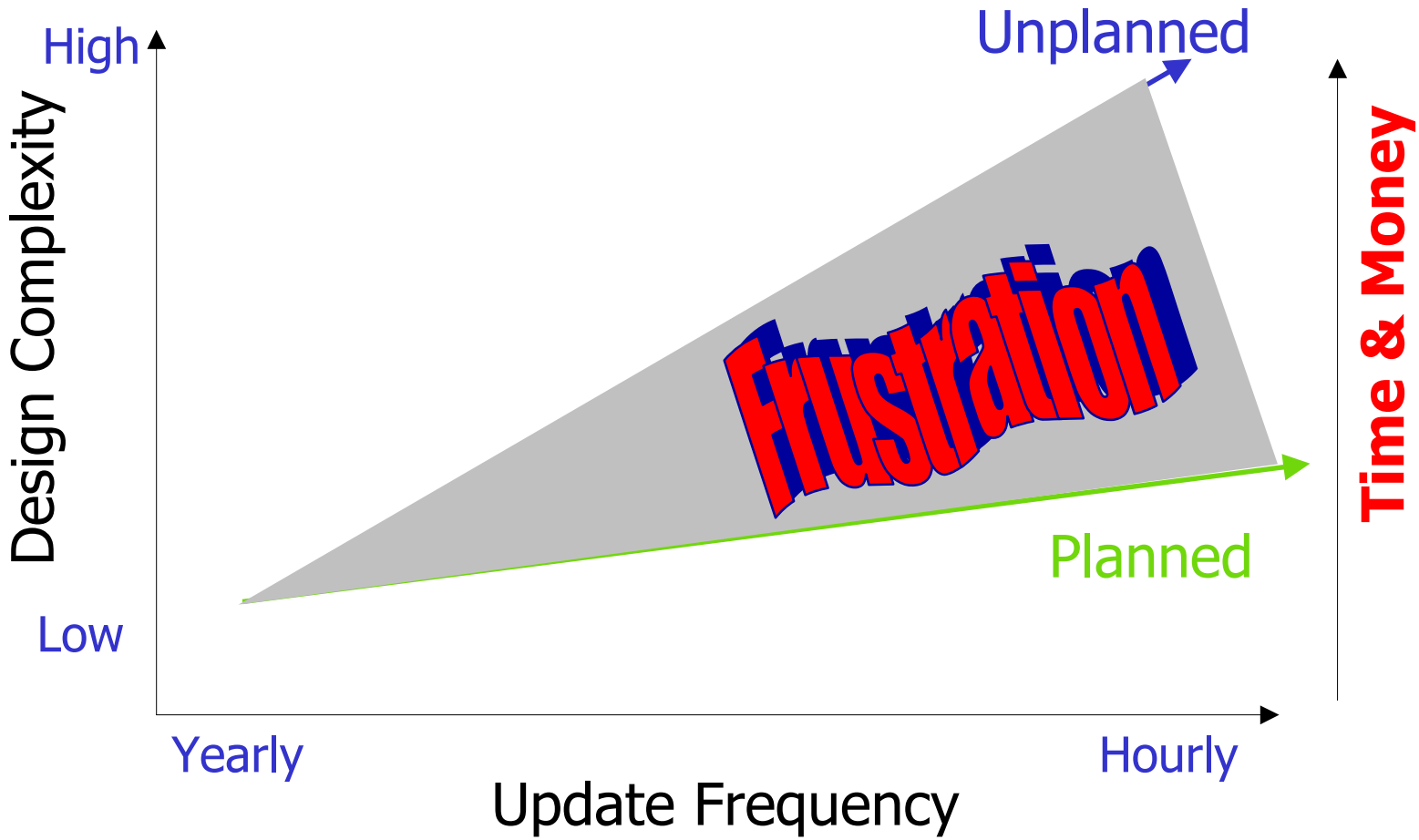
"A new factor, that of rapid change, has come into the world. We have not yet learned how to adjust ourselves to its economic and social consequences."

Harvard Business Review
October 1932

Business Intelligence Trends



Highly Effective OLAP Solutions



Contents

- ✔ BI Trends
- ✔ Seven Habits
- ✔ Trivial Pursuit
- ✔ Characteristics of a Good Show
- ✔ Q&A

Seven Habits for Designing Highly Effective OLAP Solutions

- Visualize
- Design with the end in mind
- First things first
- Focus on the Customers
- Listen first, then execute
- Collaborate
- Review, Analyze, Iterate

Step-by-Step™ DW Development Lifecycle

Establish the Program	1. Develop the Rationale 2. Target the Opportunities 3. Define the Architecture 4. Create the DW Program
Prepare the Project	5. Plot the Project Tasks 6. Establish the Infrastructure 7. Procure the Tool Kit 8. Assemble the Team
Initiate the Database	9. Investigate Consumer Needs 10. Triage the Source Elements 11. Model States and Dimensions 12. Design the Database 13. Acquire Source Data 14. Populate the Database
Explore the Data	15. Iterate Base Table Design 16. Explore Consumer Usage Interactively 17. Tune Collection Design 18. Plan Cycles and Production Migration
Implement the Deliverables	19. Prepare for Release 20. Train the Consumers 21. Initiate Support Processes 22. Migrate to Production
Expand the Environment	23. Manage the DW Inventory 24. Synchronize with Evolving Business Needs 25. Evangelize Endlessly 26. Do it Again!

BIXCHANGE

Contents

- ✔ BI Trends
- ✔ Seven Habits
- ✔ Trivial Pursuit
- ✔ Characteristics of a Good Show
- ✔ Q&A

Trivial Pursuit
MILLENNIUM
EDITION
History

BIXCHANGE

Trivial Pursuit

Millions

RANK	MOVIE TITLES (1900-2001)	DOMESTIC	WORLD	Budget	Return on Investment DOMESTIC	Return on Investment WORLD
1	Titanic (1997)	$600.80	$1,835.40	$200	300%	918%
2	Star Wars: Episode I - The Phantom Menace (1999)	$431.10	$922.60	$115	375%	802%
3	Jurassic Park (1993)	$357.10	$920.10	$63	567%	1,460%
4	Independence Day (1996)	$306.20	$811.20	$75	408%	1,082%
5	Star Wars (1977)	$461.00	$798.00	$11	4,191%	7,255%
124	**Waterworld (1995)**	**$88.20**	**$255.20**	**$175**	**50%**	**146%**
137	**Blair Witch Project, The (1999)**	**$140.50**	**$240.50**	**$0**	**401,428,571%**	**687,142,857%**

Source: IMDB.com & WorldwideBoxOffice.com

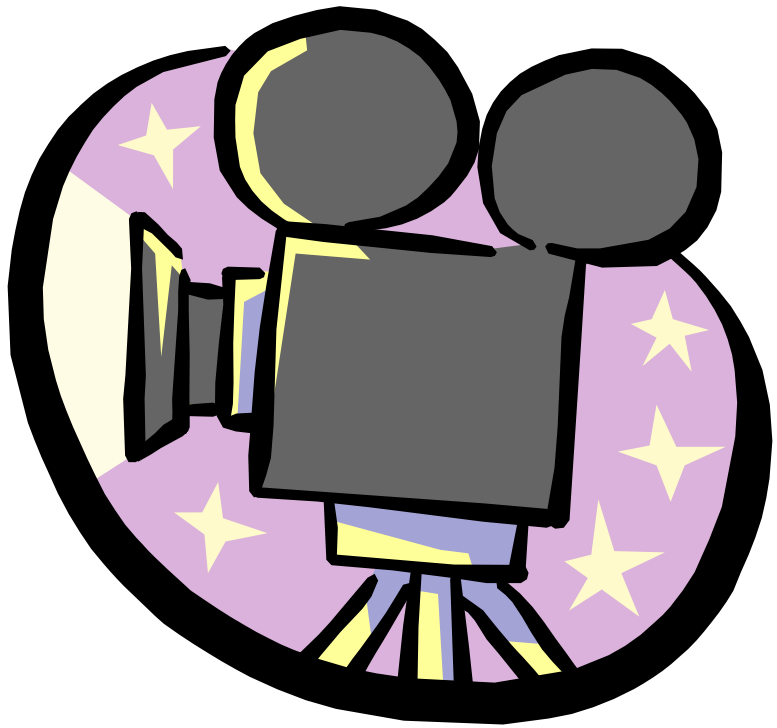
BIXCHANGE

Contents

- ✔ BI Trends
- ✔ Seven Habits
- ✔ Trivial Pursuit
- ✔ Characteristics of a Good Show
- ✔ Q&A

What defines a good show?

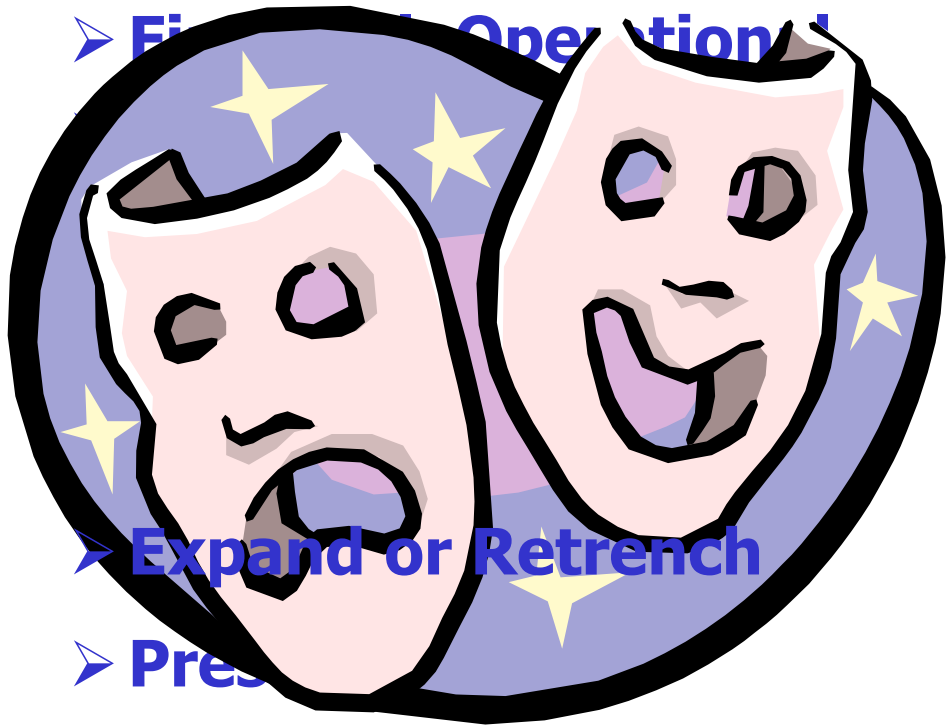
- Actors?
- Directors?
- Editors?
- Producers?
- Story?
- Setting?
- Special Effects?
- Budget?



BIXCHANGE

Characteristics of a good OLAP Show

- PLOT
- SETTING
- CHARACTERS
- POINT OF VIEW
- SYMBOLISM
- THEME
- STYLE/VOICE



- Fi… Oper…tion…
- Expand or Retrench
- Pres…

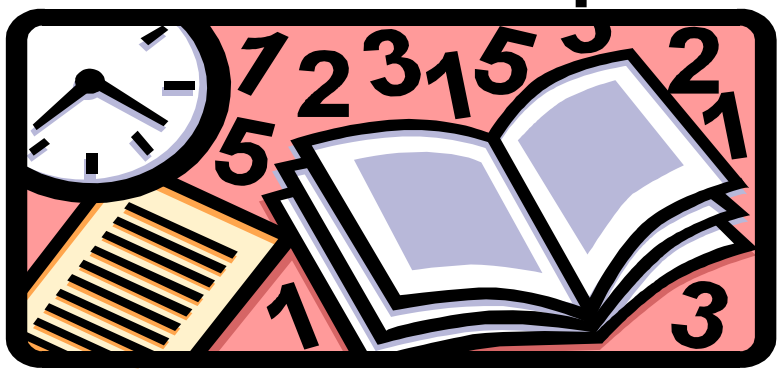
PLOT

- **What *exactly* happens in the presentation?**
- **Is there revelation, resolution, or both?**

If conclusions are drawn, what recommendations made?

SETTING

- What is the time period of the presentation?
- Does the setting make a difference, or could this story take place anytime, anywhere?
- How might a different time period affect the story?



Design around time.

CHARACTERS

- **What characters are multi-dimensional/ dynamic?**
- **What characters are flat/static? Why?**
- **How is their character revealed (hierarchy, comments, member properties, etc.)?**

Understand how your customers look at dimensions.

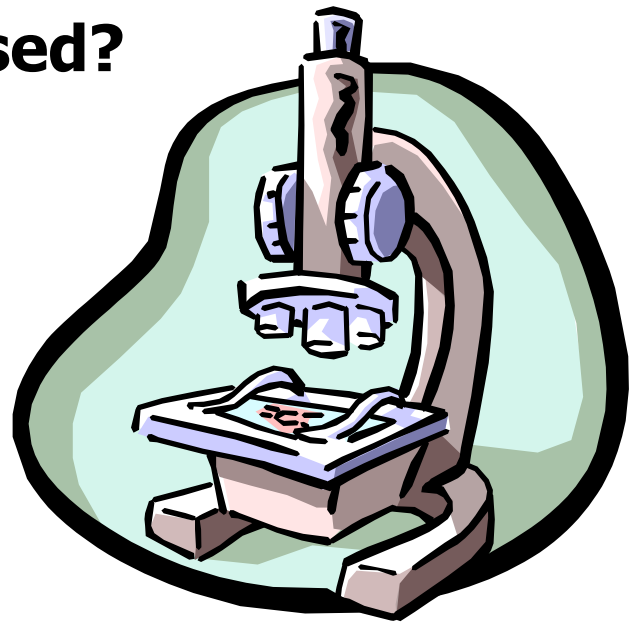
POINT OF VIEW

- **Who tells the story?**
- **What about the narrator makes a difference in the story (participant/nonparticipant, function, character, limitations, etc.)?**

Is there enough depth to satisfy the info consumer?

SYMBOLISM

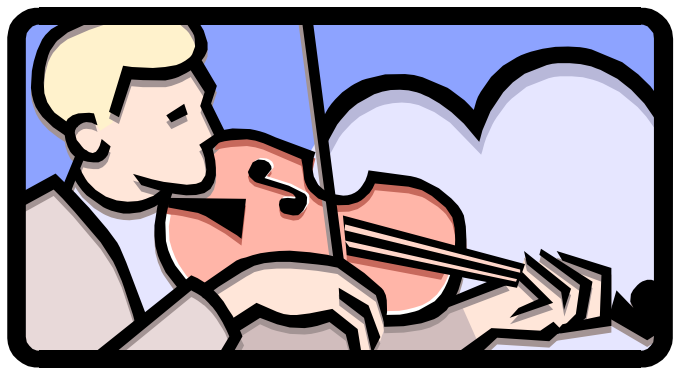
- **What are the symbolic elements?**
- **How and why are they used?**



What about icons, color schemes and tag lines?

THEME

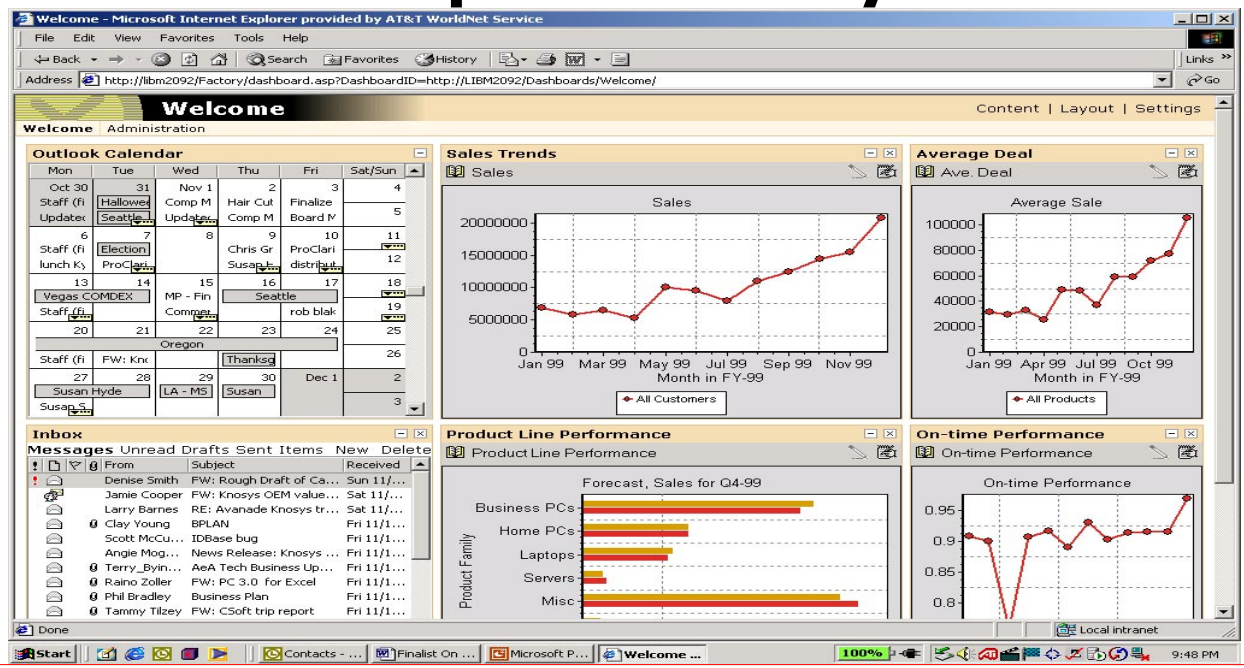
- **What repetitions of words or details do you want people to find?**
- **What does the story say about the business? How loudly? How well?**



To be or not to be subtle???

STYLE/VOICE

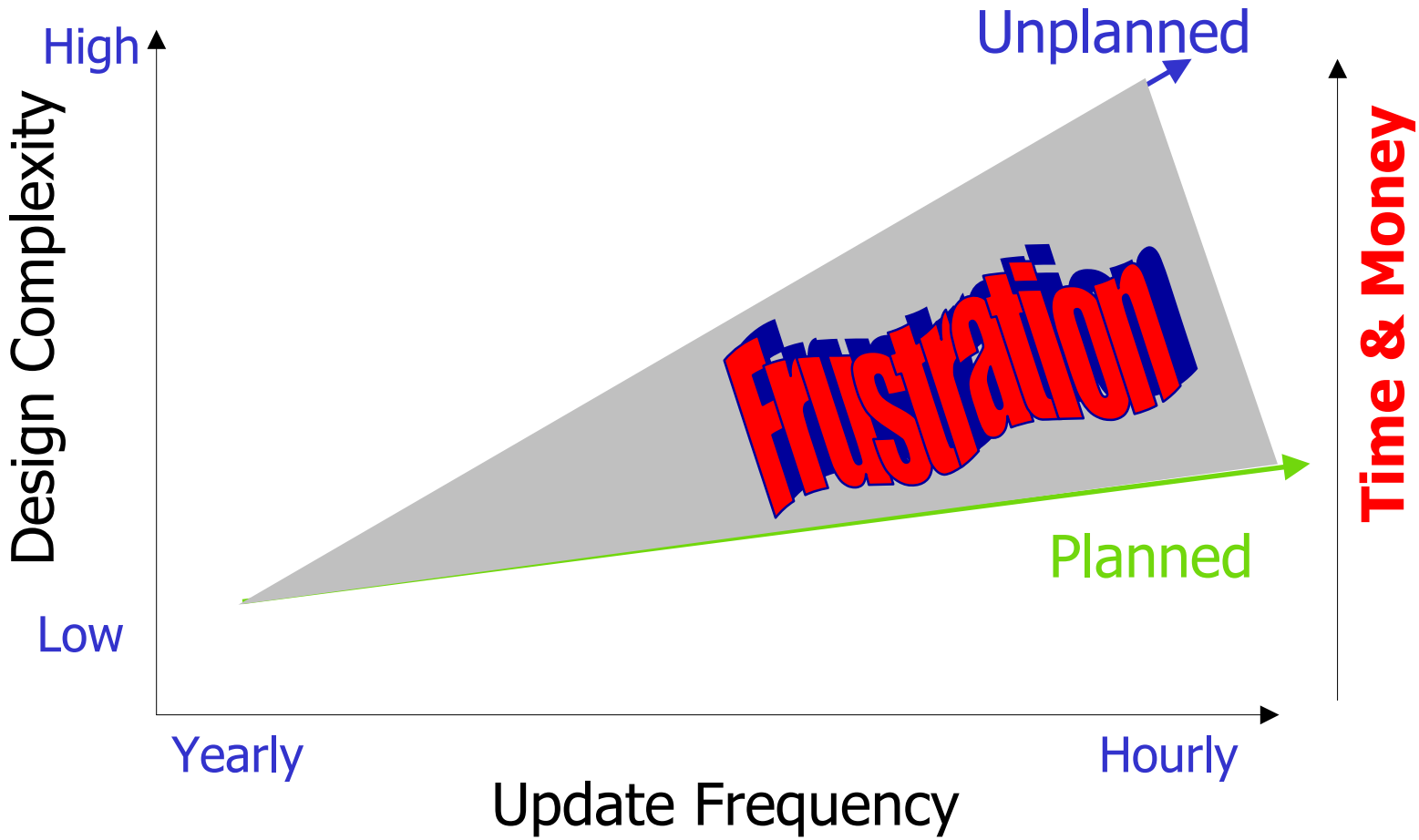
- **What presentation tools are you using?**
- **How does it impact the story?**



One size does not fit all!!!

Back End vs. Front End

Highly Effective OLAP Solutions



Contents

- ✔ BI Trends
- ✔ Seven Habits
- ✔ Trivial Pursuit
- ✔ Characteristics of a Good Show
- ✔ Q&A

BIXCHANGE

"Avoid DRIP through SIT"

DAMAN
CONSULTING

BIXCHANGE

Break Time!

DAMAN
CONSULTING

Introduction to SQL Server 2000 Analysis Services

Agenda

- **Microsoft's Market Perspective and Fit Within Business Intelligence (BI)**
- **Overview of key OLAP & BI concepts**
- **Overview of Analysis Services**
  - Intuitive Design
  - Storage modes & aggregation techniques
  - Use of partitions
  - Data Access

Vision - BI For Everyone

- **Affordable**
- **Scalable**
- **Fast "Go To Market"**
- **Ubiquitously accessible**
- **Embeddable**
  - Great Plains
  - Manugistics
  - Web Trends
  - Business Objects
  - Cognos
  - etc.

BIXCHANGE

Solution – BI Tools In The Box

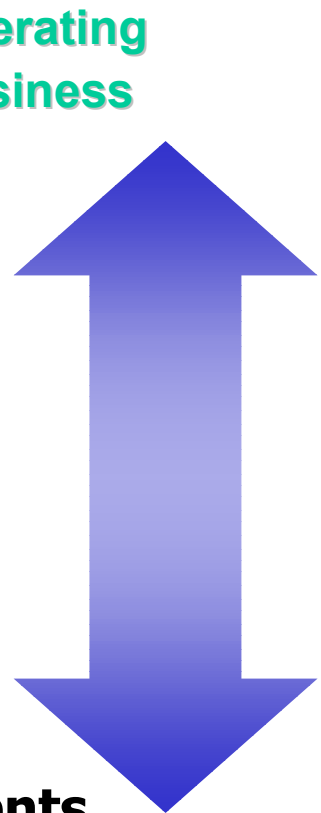
- **Design the data warehouse / data mart**
  - SQL 2000 – Enterprise Manager
  - **Visio 2000 – Enterprise Manager**
- **Populate the data warehouse / data mart**
  - SQL 2000 – Data Transformation Services
- **Create OLAP Cubes**
  - SQL 2000 – Analysis Services
- **Query The Data**
  - SQL 2000 – English Query
  - **Microsoft Office**
    - **Excel & Access**
    - **Office Web Components**

Agenda

- Microsoft's Market Perspective and Fit Within Business Intelligence (BI)
- **Overview of key BI concepts**
- Overview of Analysis Services
  - Intuitive Design
  - Storage modes & aggregation techniques
  - Use of partitions
  - Data Access

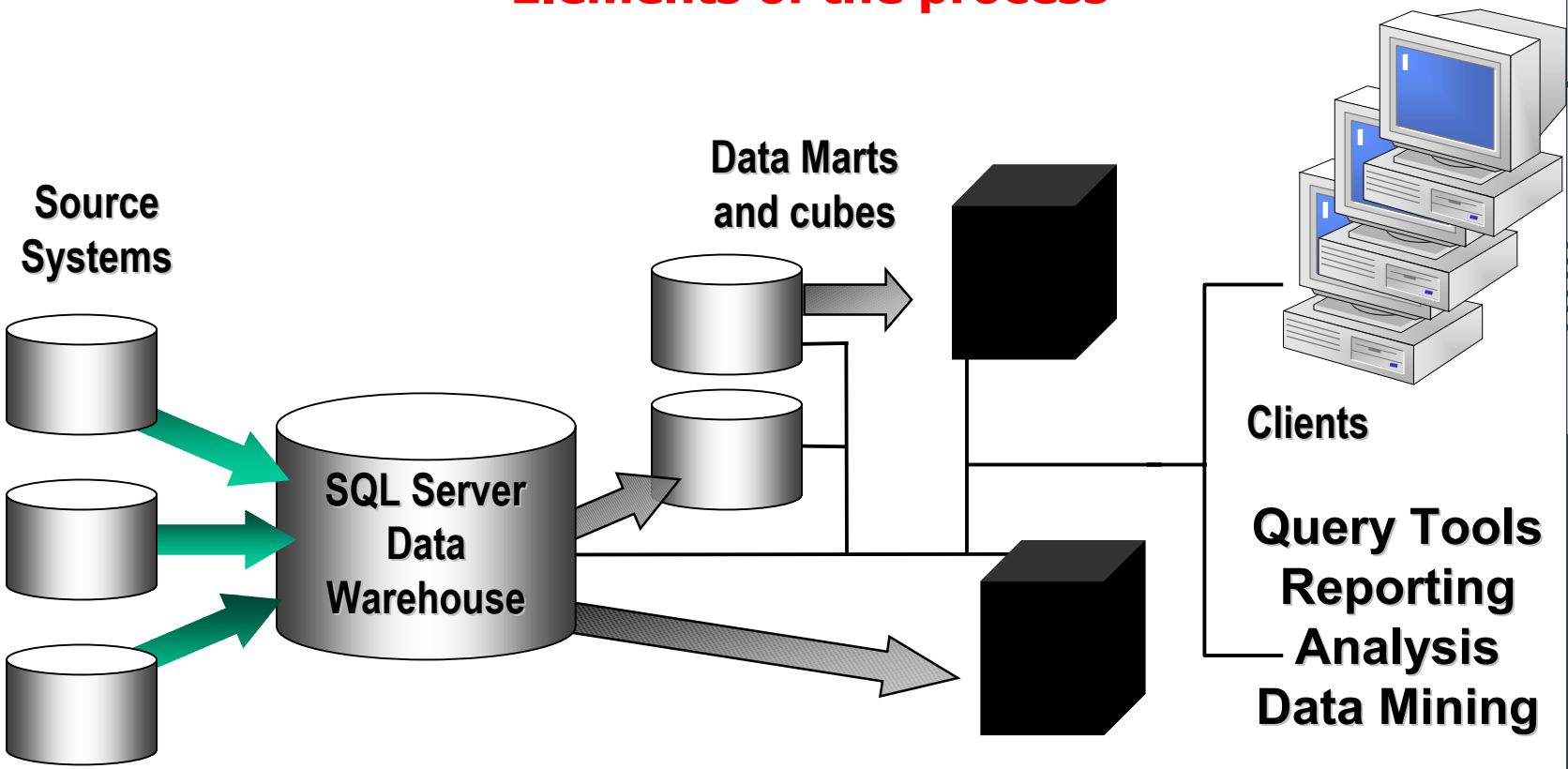
OLTP Differs From BI

- **OLTP supports**
  - Streamlining operations
  - Real time production systems
  - Current, changing data
  - Granular – Transactional
- **BI supports**
  - Analyzing Operations / Improving Decision Making
  - Consistent, heterogeneous data
  - Voluminous, historical, stable data
  - Summarized data
- **BI has different design & storage requirements from OLTP**



Classical BI Architecture

Elements of the process



Design the Data Warehouse → Populate Data Warehouse → Create OLAP Cubes → Query The Data

BI Architectural Goals

- **Provide platform to deliver a great user-analyst experience**
  - With data that is consistent, centralized and easily accessible
  - Without getting in the way of OLTP systems
- **Ability to incorporate data from internal or external sources - regardless of format or platform**
- **Agile so that it can adapt to changes in the business**

BI User Requirements

- **Support decision making – about managing & planning**
  - How/what/when/why/where of a business
  - Facilitate queries without hindering operational systems performance or having to change the design
  - Provide centralized repository of consistent data
  - Answer complex queries quickly
  - Enable data mining to discover new relationships in data
- **Provide different levels of analysis**
  - View data from many perspectives
  - Easily navigate from summary to detail
- **End user acceptance and usage is the true measure of success**

What is OLAP?

- **OnLine Analytical Processing**
  - It's a cool way of cheating that enables you to get queries answered incredibly fast
  - OLAP aggregates data (it pre-summarizes data) across all dimensions

    Example:

    by MO, QTR, YR

    by Country, State, City …

    etc…
  - Basic argument:

    "Why read through each and every detailed transaction to get an answer when the question can be answered more quickly using summary level data"

Why Use OLAP With DWHS?

- **OLAP is an enabling technology that supports dynamic analysis**
  - Intuitive multidimensional model
  - Fast response times against huge databases
  - Offers complete syntax for expressing analytical queries and business logic
  - Optimizes the use of network resources as well as Internet/Intranet deployments

Agenda

- Microsoft's Market Perspective and Fit Within Business Intelligence (BI)
- Overview of key BI concepts
- **Overview of Analysis Services**
  - Intuitive Design
  - Storage modes & aggregation techniques
  - Use of partitions
  - Data Access

OLAP – Enabling Analysis

Data Sources

- Analysis Services works as well with data coming from Oracle, DB/2 and others as it does with SQL
- Analysis Services supports all data sources accessible via ODBC (including flat files)
- Analysis Services also supports newer generation OLE-DB data sources
- DTS and SQL capable of accepting XML

OLAP – Enabling Analysis

Mapping warehouse to cube

- **Dimension tables used to build OLAP dimensions**
  - Shared dimensions enable analysis to take place across cubes
  - Private dimensions support different hierarchies and aggregation rollups
- **Virtual cubes join two or more physical cubes into one logical cube**
  - Requires one or more shared dimensions

OLAP – Enabling Analysis

Dimensional design

- **Dimensional modeling easy to accomplish in Analysis Services**
  - Intuitive design palette enables users to model their data and view the results from the same pane
  - Wizards exist to help guide users through common and complex tasks
- **On-line tutorial exists within product to guide new users through basic and advanced concepts**

Designing dimensions

DAMAN
CONSULTING

OLAP – Enabling Analysis

Cube design

- **Public / Private Dimensions available**
- **Quantitative columns in fact table become measures**
- **Calculated members support creation of:**
  - Derived entities in a dimension
    - New products or regions
  - Derived measures
    - Income = revenue-expense

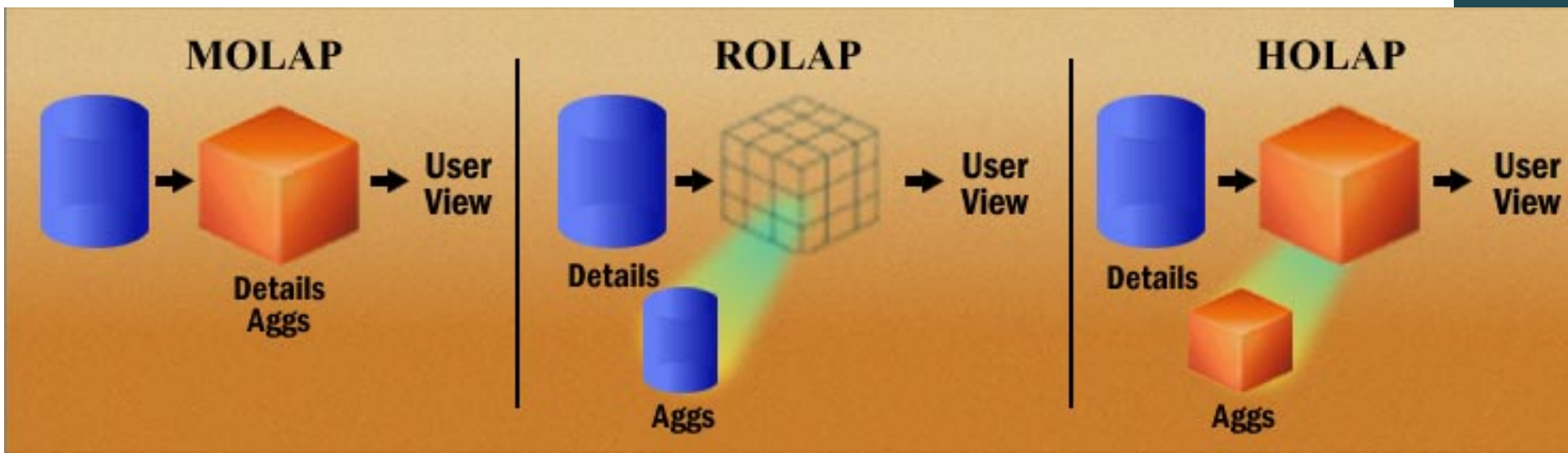
Designing Cubes

DAMAN CONSULTING

OLAP – Enabling Analysis

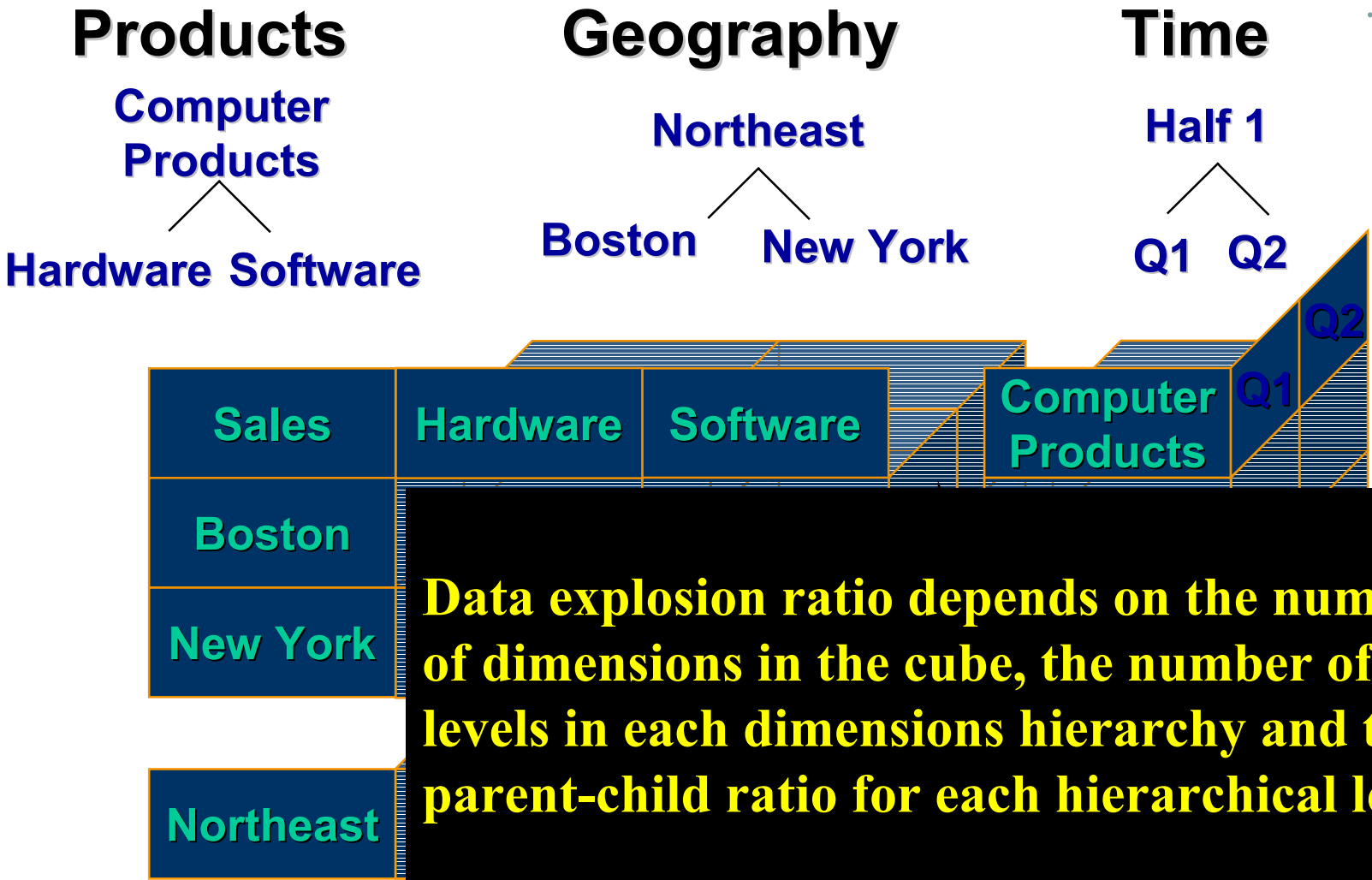
Flexible Storage

- Debates between MOLAP and ROLAP vendors obscure customer needs
- Analysis Services supports MOLAP, ROLAP, and HOLAP and offers seamless integration of all three
- Users & applications see only cubes



OLAP – Enabling Analysis

Dealing with data EXPLOSION!



OLAP – Enabling Analysis

Dealing with data EXPLOSION!

Persist only base aggregates

- Min
- Max
- Sum
- Count
- Distinct Count

*** other, more complex aggregates are derived from base aggregates**

OLAP – Enabling Analysis

Dealing with data EXPLOSION!

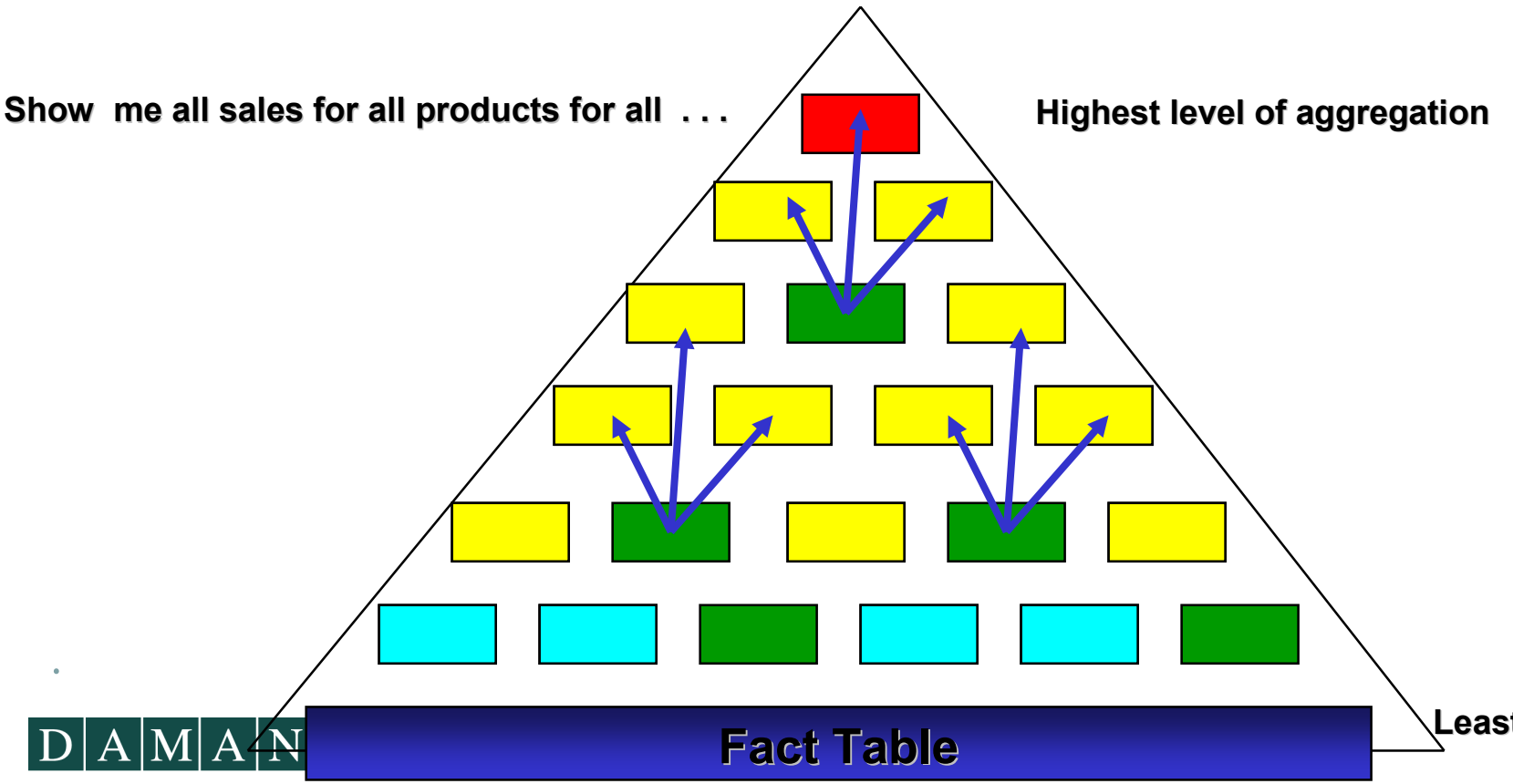
Automagically eliminate all sparse data from the cube

	Furnaces	Air Cond.	HVAC Products
Phoenix		150	250
Anchorage	250		350
USA	350	250	600

OLAP – Enabling Analysis

Dealing with data EXPLOSION!

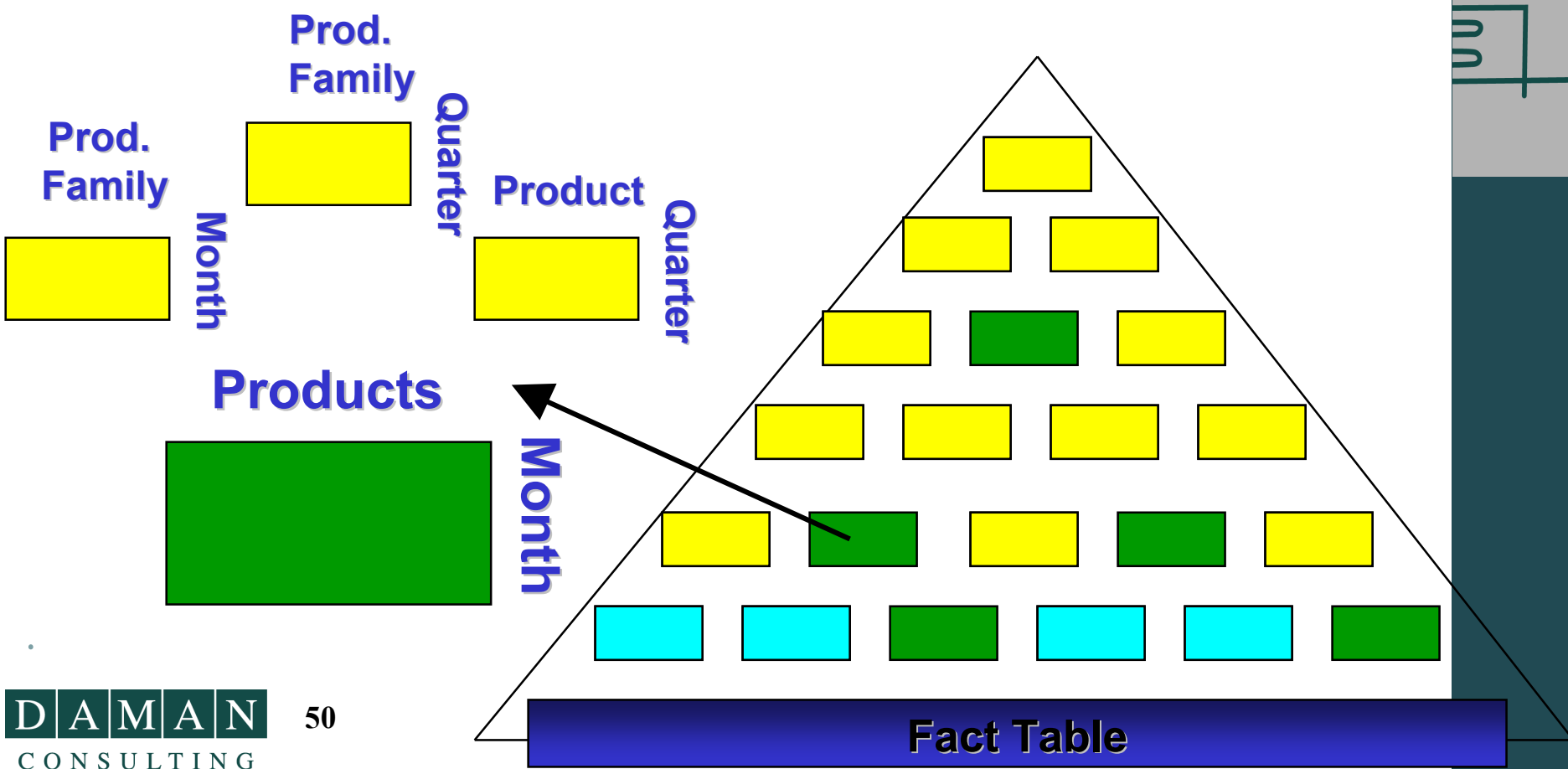
Do only partial pre-aggregation



OLAP – Enabling Analysis

Dealing with data EXPLOSION!

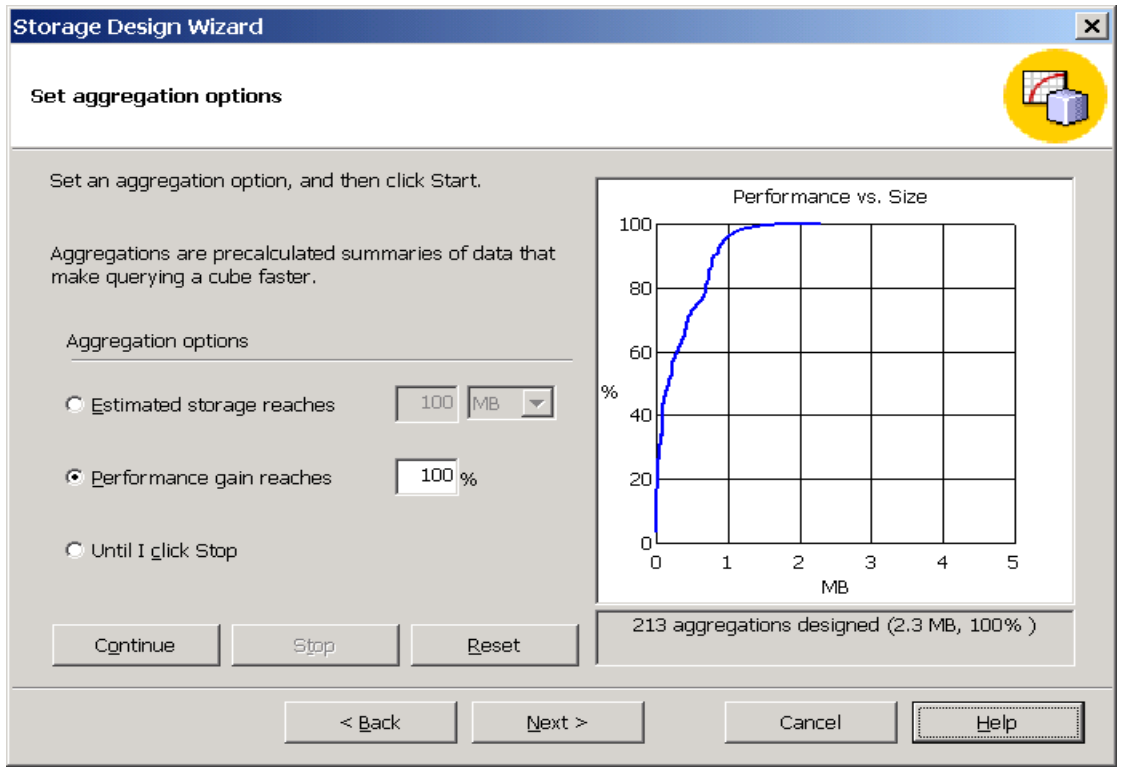
Do only partial pre-aggregation



OLAP – Enabling Analysis

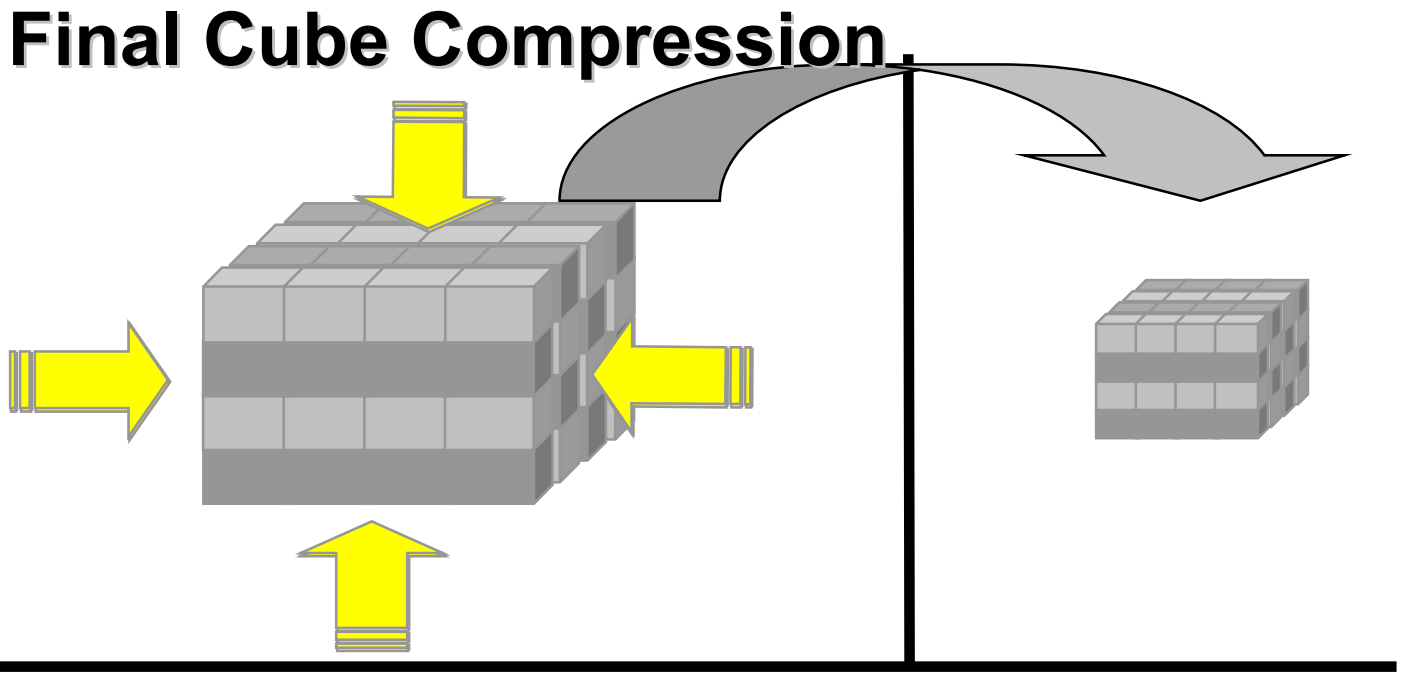
Dealing with data EXPLOSION!

- **Storage Design wizard finds the "80-20" rule in the data**
  - The 20 percent of all possible pre-aggregations that provide 80 percent of the performance gain
  - Analyses the level counts for each dimensions as well as the parent-child ratios for each level



OLAP – Enabling Analysis

Dealing with data EXPLOSION!



When using either MOLAP or HOLAP storage modes, Analysis Services is capable of compressing the final cube size

BIXCHANGE

Redefining OLAP Scalability

T^3 (pronounced T-Cubed)

Goal:

- **Demonstrate the scalability of Analysis Services**
  - Build cube from 1TB+ of source data
  - Describe the techniques used to operate at that scale
- **Use the cube to demonstrate high speed queries**
- **Proof-of-concept system**

Address real business issues:

schema, data, goals

Results audited by Winter Corp.

see http://www.microsoft.com/SQL/techinfo/terabytecube.htm

T³ Storage Requirements

Table storage (relational)

Table	Million Rows	GBytes
detail_brand_*	1,030	163.4
detail_prodmod_*	20	3.2
detail_subcat_*	11	1.7
detail_upc_*	4,881	793.8
month_brand_*	295	47.1
month_prodmod_*	5	0.8
month_subcat_*	3	0.4
month_upc_*	1,429	225.3
Total	7,674	1,236
	7.7 Billion	**1.2 TB**

Cube storage

Cubes	GBytes
Week_Brand	17.8
Week_Class	0.2
Week_Subgroup	0.0
Week_Item	434.7
Month_Brand	4.9
Month_Class	0.1
Month_Subgroup	0.0
Month_Item	24.5
Market Research	482
	471 GB

T³ Performance

- **Processing**
  - 7.7 billion rows, 50 hours
  - 153 million rows/hr
  - 42K rows/sec
  - 60-70% CPU utilization
- **Querying**
  - 50-user workload, 1350 queries, 30-sec think time
  - Cold cache
  - Median response 0.08 sec, mean 1.2 sec
  - Low CPU load - didn't have enough queries running simultaneously!

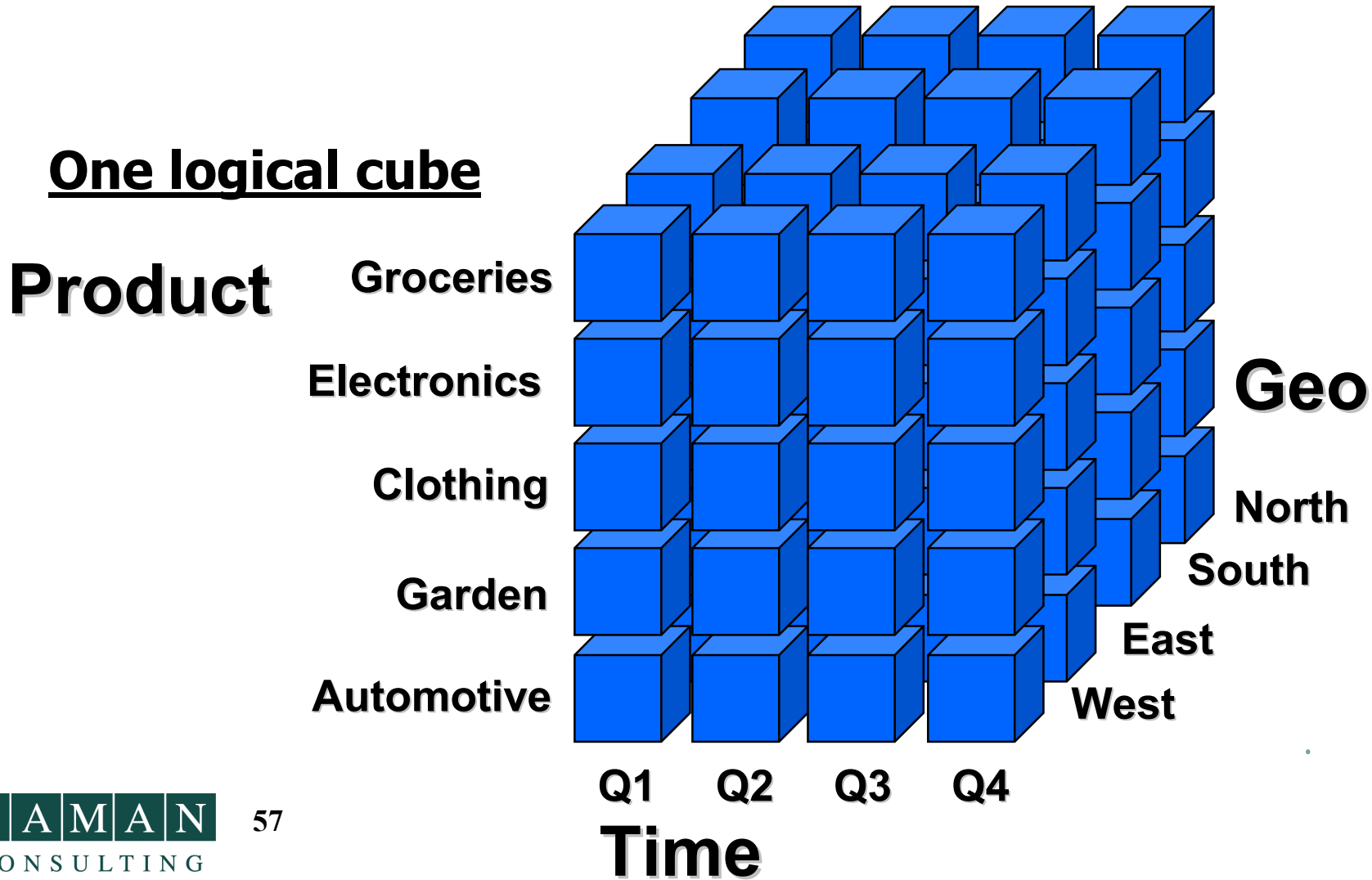
Designing Storage

DAMAN CONSULTING

OLAP – Enabling Analysis

Cube Partitioning

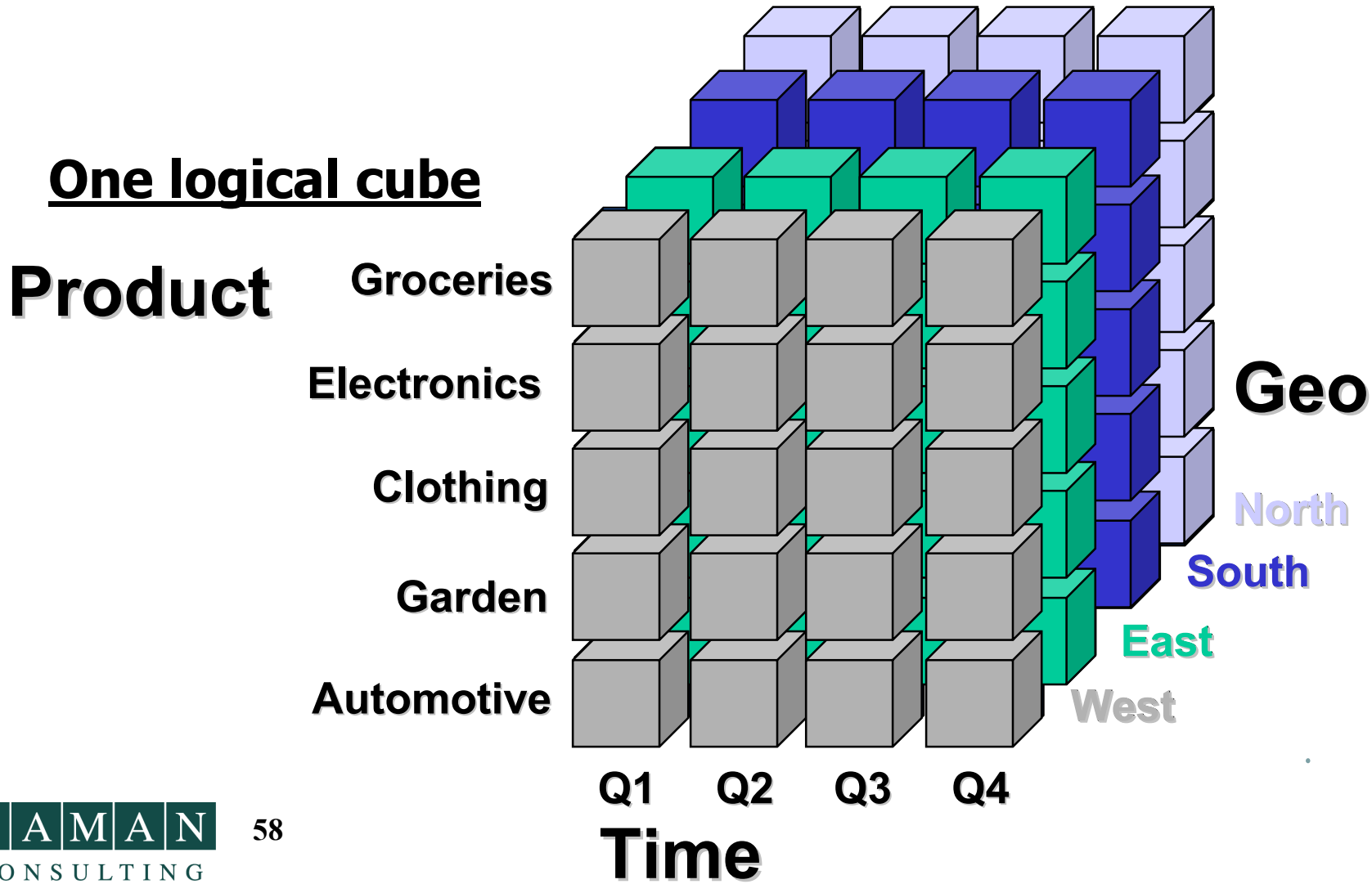
One logical cube



OLAP – Enabling Analysis

Cube Partitioning

One logical cube



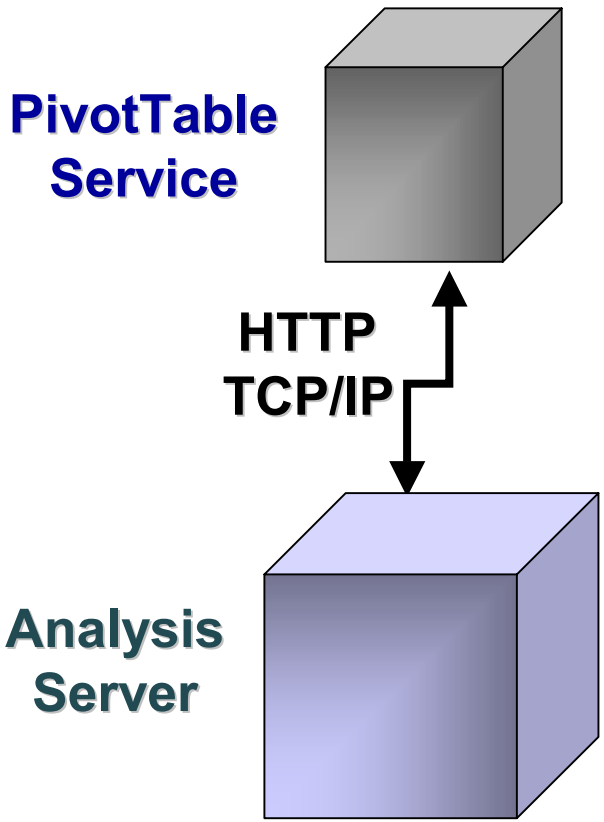
OLAP – Enabling Analysis

Importance of Partitions

- **Central to Analysis Services**
- **Data Management**
  - Incremental updates
  - Adding/dropping time periods
  - Granular control over aggregation design
  - Enable hybrid cube design (e.g. combination of MOLAP and ROLAP partitions in the same logical cube)
- **Basis for advanced features like Write-back**
- **Important scalability & performance implications**
  - Single server parallelism (partition processing)
  - Multi-server parallel processing
  - More efficient query processing

OLAP – Enabling Analysis

Efficient Network Utilization



Client Features

- multidimensional calc engine
- data & metadata caching (session)
- query management
- client OLEDB for OLAP interfaces

Server Features

- multidimensional calc engine
- data & metadata caching (multi-user)
- query management
- server DSO admin interfaces
- security
- data refresh
- aggregation management

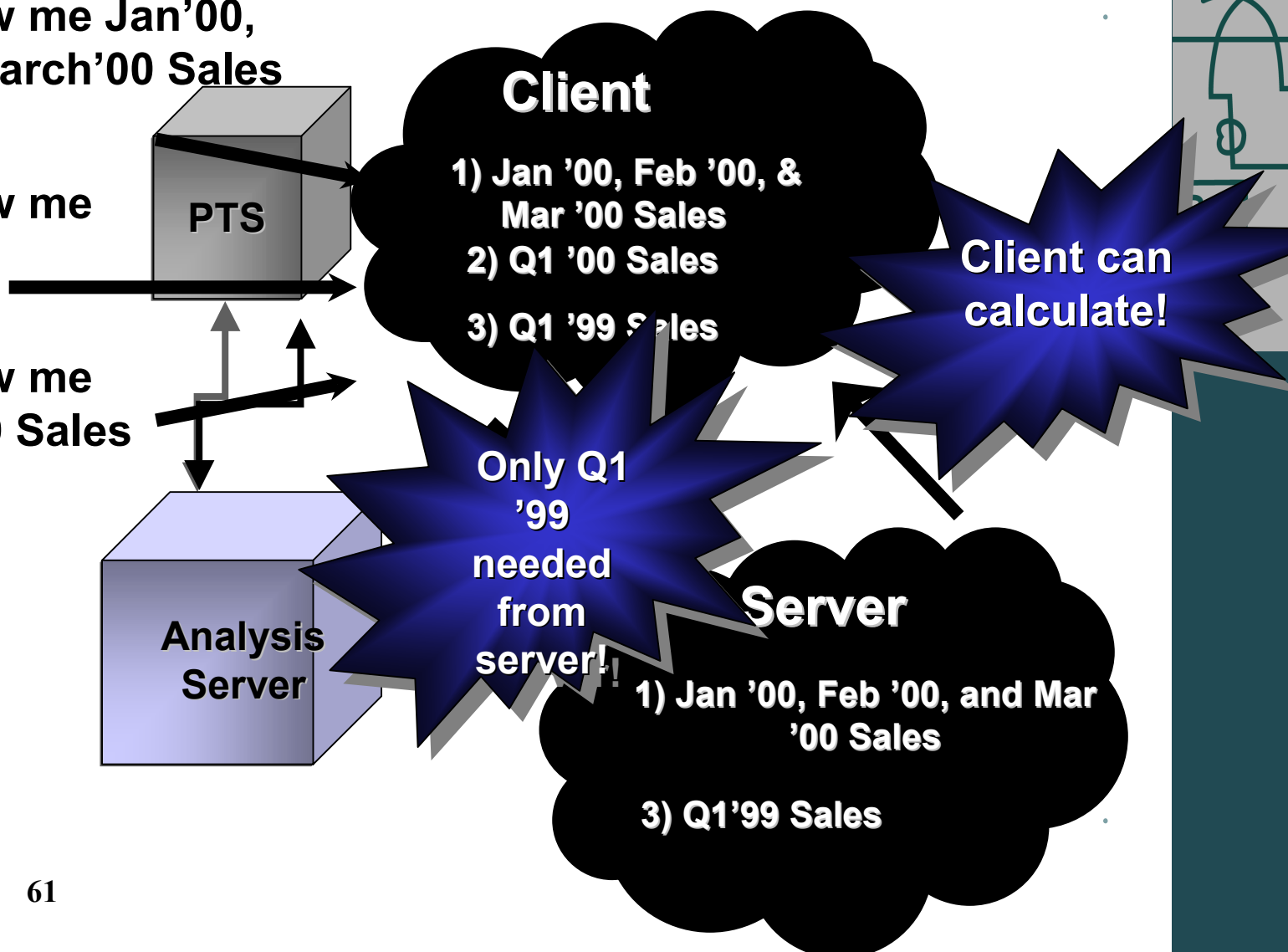
OLAP – Enabling Analysis

Cooperative Caching

Query 1) Show me Jan'00, Feb'00, and March'00 Sales

Query 2) Show me Q1'00 Sales

Query 3) Show me Q1'00 & Q1'99 Sales



Office &
Office Web Components

DAMAN
CONSULTING

Features we didn't have time to cover …

- Member properties & virtual dimensions (same as attributes in other products)
- Write-back & drill through
- Virtual cubes
- Data Mining
- Actions
- Custom rollups
- Calculated cells
- Dimension architecture
- Usage based optimization

Solve Prickly OLAP Problems That Plague Other Vendors

- Cube explosion – not uncommon to experience exponential data explosion
- Time needed to reprocess/refresh cube(s)
- Confusion over storage options (MOLAP, ROLAP, HOLAP)
- Inability to handle really complex calculations
- Cost prohibitive to roll out across enterprise

Analysis Services effectively solves all these problems

BIXCHANGE

OLAP Advantages In Summary

Intelligent Aggregations	◆ Significantly smaller databases ◆ Faster initial and incremental processing
Flexible Storage Architecture	◆ Supports MOLAP, ROLAP, and HOLAP equally well ◆ Application requirements determine storage, not vendor
Ease-of-Use Ease-of-Mgmt	◆ Fast "To Market" Solutions ◆ Lower TCO ◆ Broad accessibility to data
PivotTable Service	◆ Client-side cache -- improved performance / efficiency ◆ Mobile/Disconnected analytical support ◆ Web enabled access
Integration	◆ Office, SQL Server, .Net Servers, 3rd Parties ◆ Compelling bridge between Excel and SQL Server ◆ End-to-End data storage, transformation, and analysis

Coming Soon…
Quick Start for Analysis Services

Microsoft Software for an Agile Business

Fast. Scalable. Flexible. Powerful. Reliable. Enterprise-ready.

BIXCHANGE

Break Time!

DAMAN
CONSULTING