



# Technical Forum & Executive Briefing

17 al 21  
Octubre  
2011

Imagine PODER Imagine CAPACIDAD

## IBM Smart Analytics System 7700 and 7710 Workload Optimized Systems



# Agenda

Why Analytics

Analytics Infrastructure

IBM Smart Analytics System

# Growing client interest in business intelligence and analytics

- ✓ 83% of CIOs said that analytics is their top priority for gaining a competitive advantage
- ✓ 35% of customers will replace their current data warehouse with a pre-integrated solution in the next 3 years
- ✓ CIOs at high growth companies are **61%** more likely to proactively turn data into actionable information



**BUSINESS INTELLIGENCE  
AND ANALYTICS**

*... driving the need for businesses to evolve and adapt dynamically to achieve agility and growth*

## The POWER of Analytics

# 2.5x increased response rates

U.S. Based High End Retailer drives additional revenue with campaigns while reducing costs of the campaign 30%



### Challenge:

- Limited understanding of product level sales variances reduced effectiveness of vendor negotiations
- Promotions targeted to un-segmented buyers were sub-optimal

### Solution:

- Advertising executives optimized expenses by reducing promotions on losing items and concentrating on marginal products where promotions earn highest ROI
- Customer Segmentation & Market Basket Analysis reduced mailing quantity by 30% and increased response from <2% to almost 5%
- Power Systems + InfoSphere Warehouse

# Analytics are cross industry

“...What is our risk this morning?”



“...Are we using our stimulus funding effectively?”



“...Which treatments are ineffective and should be eliminated to lower costs?”



“...Do we have product issues or fraudulent claims from service?”



“Our prices are lower than others. Is this sustainable given our costs, or a future threat?”

**Fast**  
**Flexible**  
**Affordable**



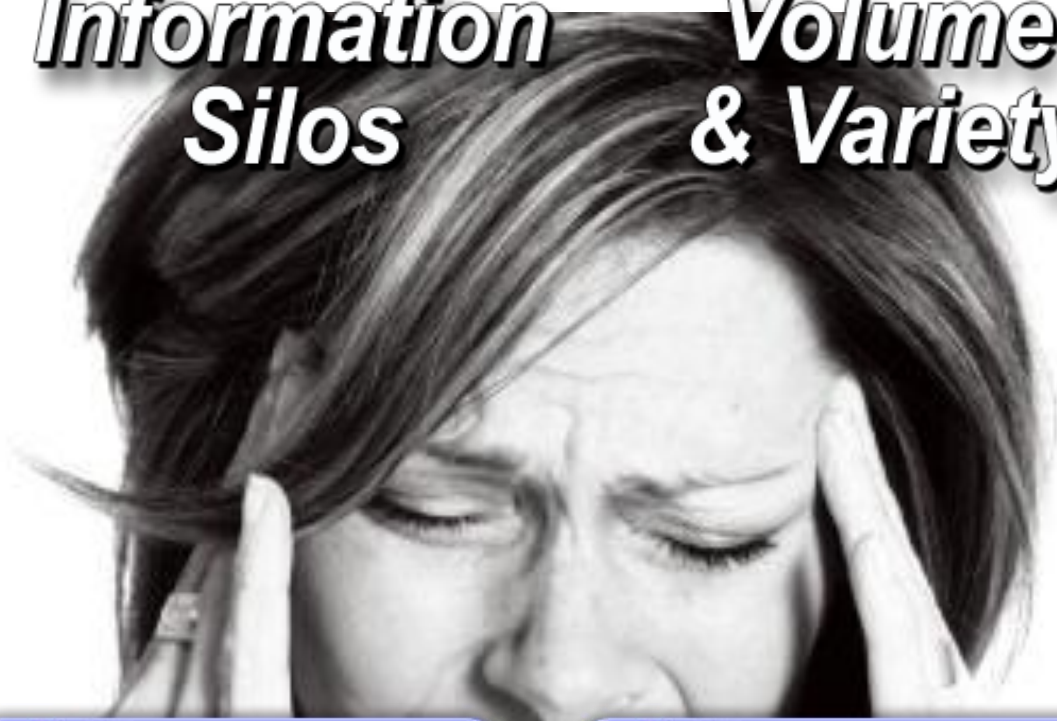
“...How & when should we adjust plans to reduce churn & expand share?”

# Unlocking the Business Value of Information

**Information  
Silos**

**Volume  
& Variety**

**Business  
Velocity**



**59%** of managers  
**miss  
information**  
they should have used

**47%** of users  
**don't have  
confidence**  
in their information

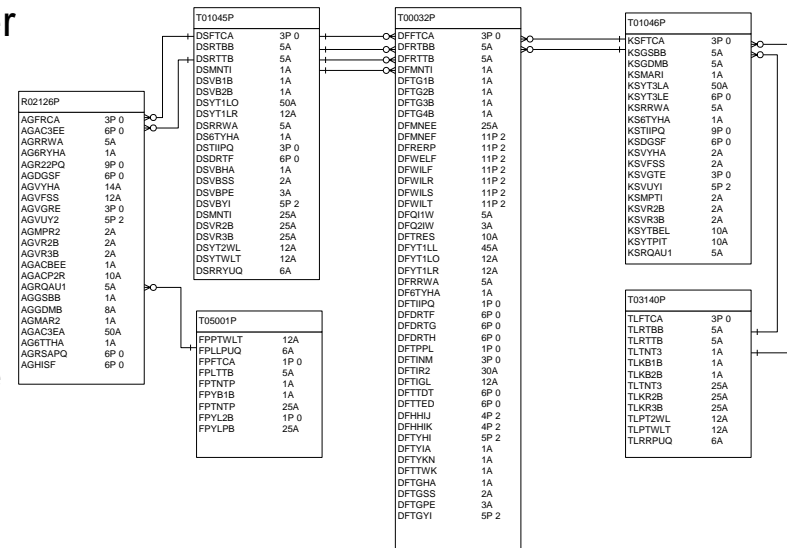
**42%** of managers  
**use wrong  
information**  
at least once a week

# Getting INFORMATION out of operational databases

can be problematic

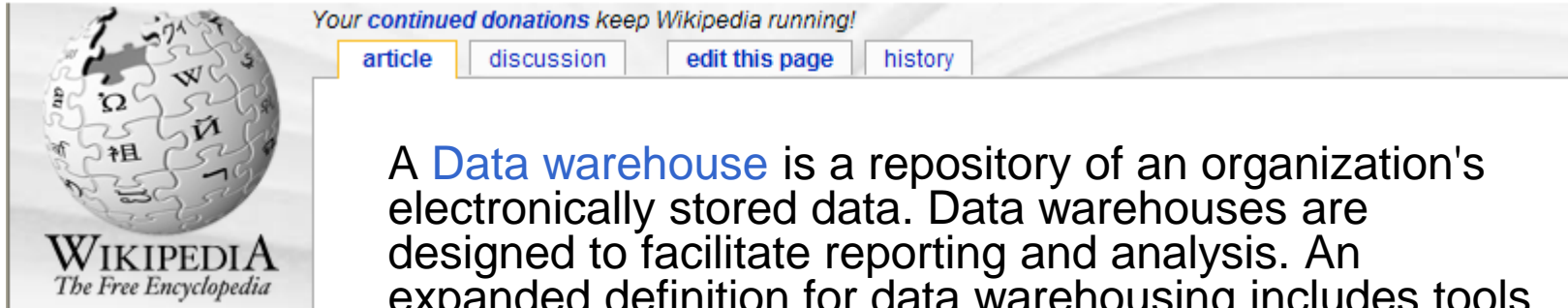
- Databases designed for TRANSACTION processing
- Data elements designed by developers – not meaningful to analysts
- Table/File relationships not well understood – need an expert
- Transactions characterized by
  - Many users
  - Short CPU time required per use
  - Small number of records accessed/updated per transaction
- Analytical Queries characterized by
  - Smaller number of users
  - Intensive resource utilization (I/O)
  - Large number of database records accessed
  - Data loads and queries occurring at same time

## Transaction Schema



- Data may not be consolidated
- Dirty Data

# The Foundation for Analytics includes:



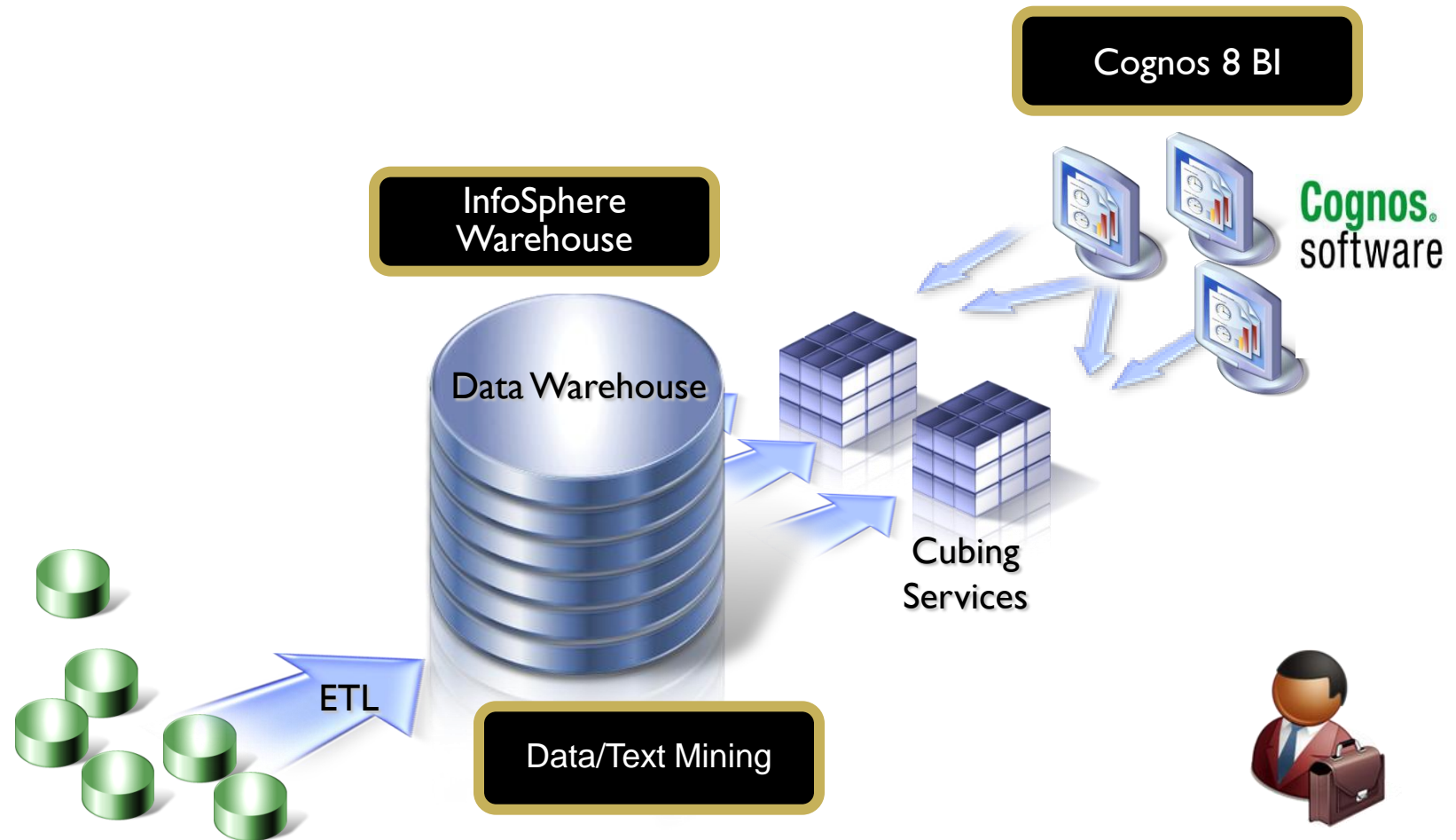
A **Data warehouse** is a repository of an organization's electronically stored data. Data warehouses are designed to facilitate reporting and analysis. An expanded definition for data warehousing includes tools to **extract, transform, and load (ETL)** data into the repository, and tools to manage and retrieve **metadata**.

**Business Intelligence** is the name for the ability to act on information that may come from a data warehouse.

**Data mining (Predictive Analytics)** is a business process based on advanced technology, which finds unknown and complex relationships in data, producing insight into business issues and predictions to improve business decisions.



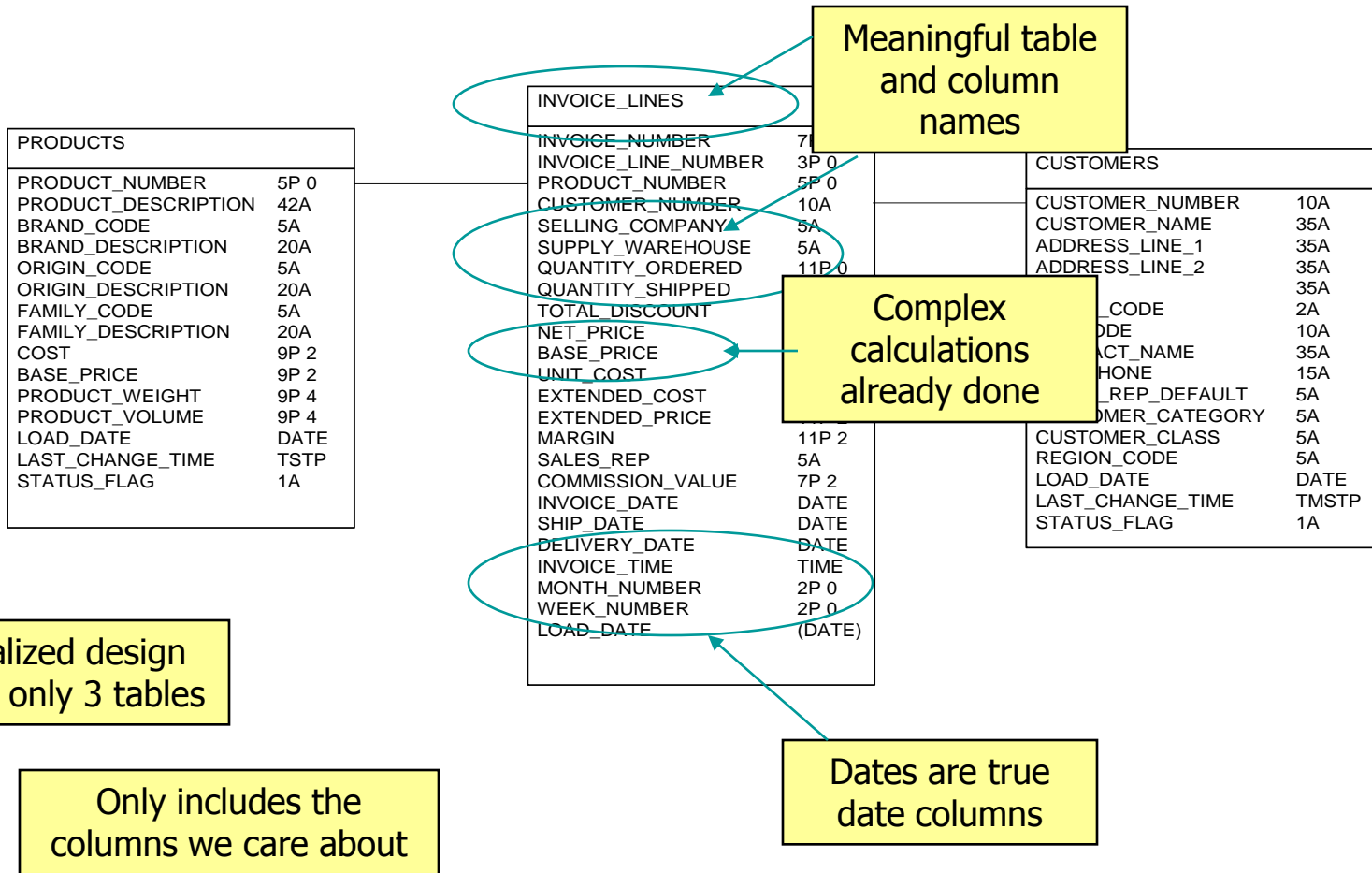
# A Data Warehouse System Schematic



Operational Source Systems  
Structured/ Unstructured Data

ETL – Extract, Transform, and Load

# Analytics Data Model



Data Warehouse Schema

# Traditional approach to Data Warehousing and Business Intelligence

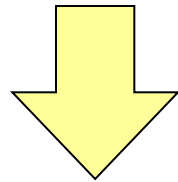
- Evaluate Business Intelligence Tools from several vendors
- Evaluate Database Management Systems from several vendors
- Evaluate Extract/Transform/Load (ETL) tools from several vendors
- Select Vendors (come to contract terms)
- Size System
- Develop Systems Integration Plan
- Develop Operational Skills/Procedures
- Build Multi-vendor Support Structure
- Install, Configure, Tune
- Tune again and again and again



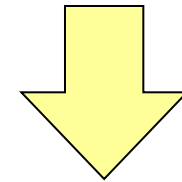
# Workloads Are Different

Business Analytics

OLTP



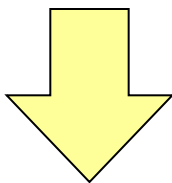
- Complex queries against a data warehouse
- Read only
- Star schema design often used to speed up queries
- Different modes of operation
- Smaller number of users



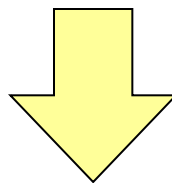
- Transactions against operational data
- Reads and writes
- Normalized schema design to eliminate redundancy
- Multiple user throughput operation
- High number of users

# Modes Of Operation Have Different Performance Objectives

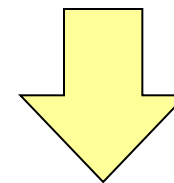
Business Analytics		OLTP
Complex analytic queries	Operational reports with high concurrency	Multiple users running transactions concurrently



Complete queries in the fastest possible time



Achieve maximum throughput



Achieve maximum throughput

Competitor's One-Size-Fits-All Database Machine

versus

IBM Workload Optimized Systems

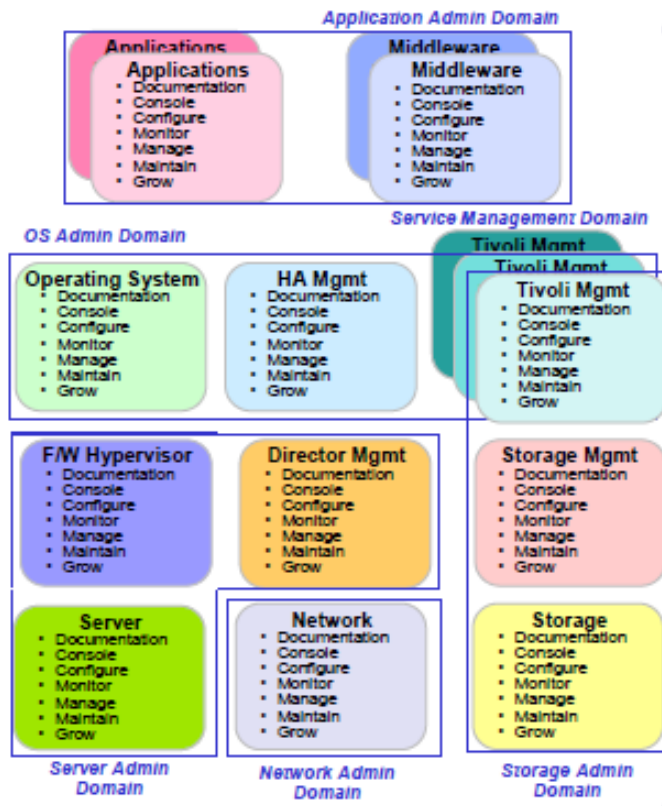
# How Workload Optimized Solutions Deliver Simplification



*Dramatically Reduces Client Understanding, Skill, and Resource needs for the Workload*

## Client view in a general environment

- Many components to understand, manage, and maintain
- Components with broad capabilities, not all of which apply to IM
- Multiple workload administrative domains and boundaries

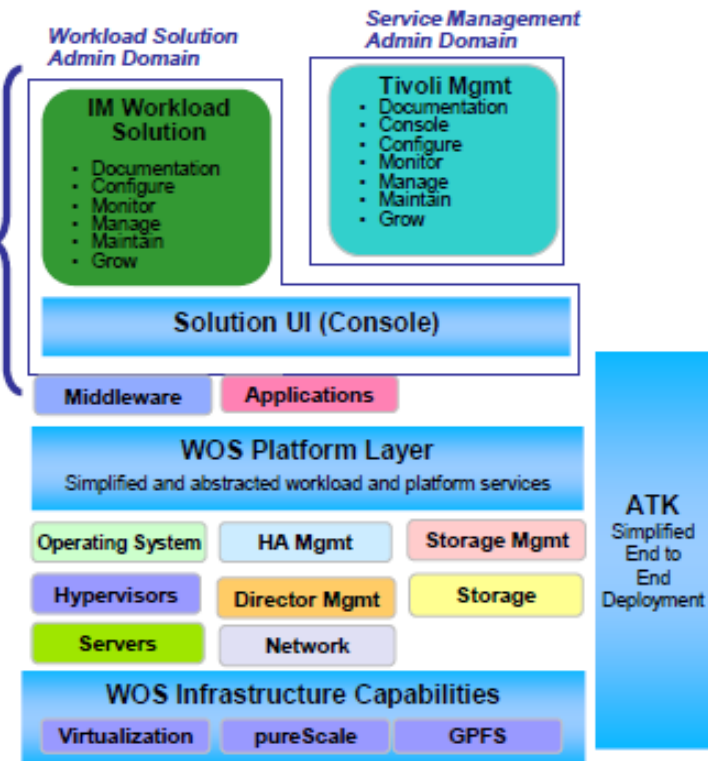


## Client View

*Complexity to be Understood, Staffed, Deployed, and Managed*

## Client View of a Workload as a Solution

- Solution-focused, *abstracted* capabilities for the Client to understand, setup, manage, & use the workload
- One simplified UI for workload management of applications, M/W, platform, and related resources
- Possible reduction/elimination of some administration domains



# So what is the IBM Smart Analytics System?

EVERYTHING you need for Business Analytics –  
not just a data warehouse appliance...

## Analytics Software

- Business Intelligence
- Cubing Services
- Text Analytics & Data Mining

## Powerful Data Warehouse

- Warehousing Platform
- Advanced Workload Management
- System Automation
- Information Integration

## Hardware & Services

- Flexible Server Platform Options
- Modular Storage Capacity
- Build, Deploy, Health Check and Premium Support Services



Transforming  
information into  
business insight

**Faster** - Workload optimized analytics run business intelligence faster

**Less floor space**  
Data compression reduces storage cost

**World record performance**

# IBM Smart Analytics System 7700

## 2x Performance at 1/2 the cost<sup>1</sup>

- **Complete End to End Analytical Solution** shipped in a manner of weeks versus months - reducing risk and improving time to value. **Completely integrated solution based on powerful warehouse infrastructure** with a single point of support
- **Start small and grow big** with **proven, flexible modular design** that preserves your investment and maintains optimized design as you grow
- **Exploit the latest POWER7 architecture**
  - Optimized for Analytics with Massively Parallel Processing design
  - **New IBM System Storage with Solid State Drive (SSD) options**
  - InfoSphere Warehouse 9.7 adds Oracle compatibility



### Compared to POWER6 based 7600

- ☑ 4x more cores per module
- ☑ 50% less space and energy requirements
- ☑ 2x storage capacity per data module

<sup>1</sup> Performance per data module, and cost comparing equivalently sized systems in Raw Data Terabytes



# POWER7 Cognos Module and Control Console

## 20-40% Performance Improvements<sup>1</sup>

- Optional Cognos module is refreshed with Power 740 8 core servers improving performance and supporting higher number of concurrent users per module. Design of the Cognos module includes an active-active, workload balancing architecture that is preconfigured and optimized by IBM before shipping
- The 7700 control console provides another layer above the hardware and software that allows administration over the system as a whole, rather than each component individually. This console allows the user to manage and maintain software for coordinated stack updates (operating system, driver, firmware, and other components).



### Cognos Module

- ☑ 2x performance through optimization at build time
- ☑ 26-40% performance improvement over POWER6
- ☑ High availability for BI built into system

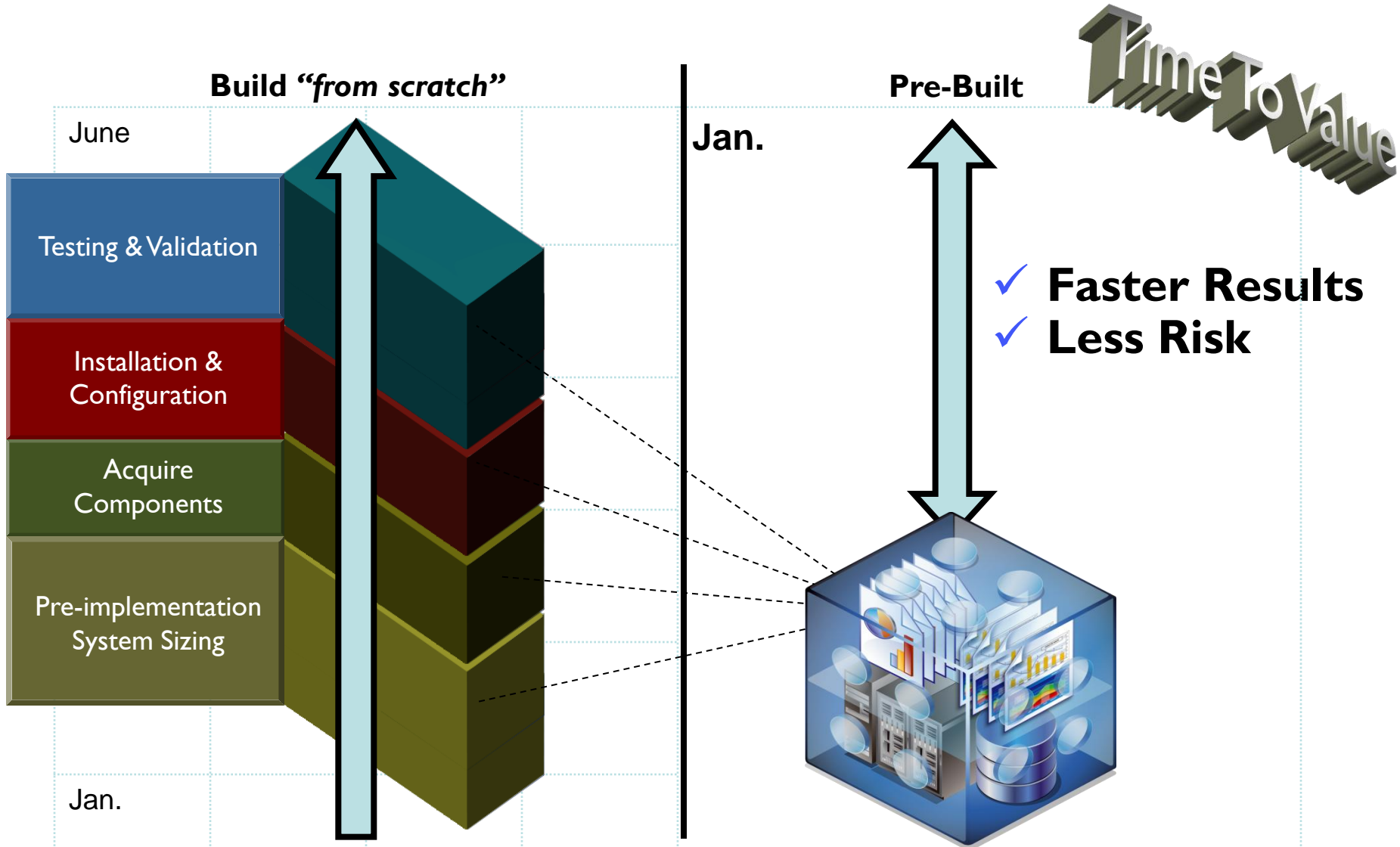
<sup>1</sup> Based on complexity of queries comparing on a core to core basis



# A workload optimized, integrated system

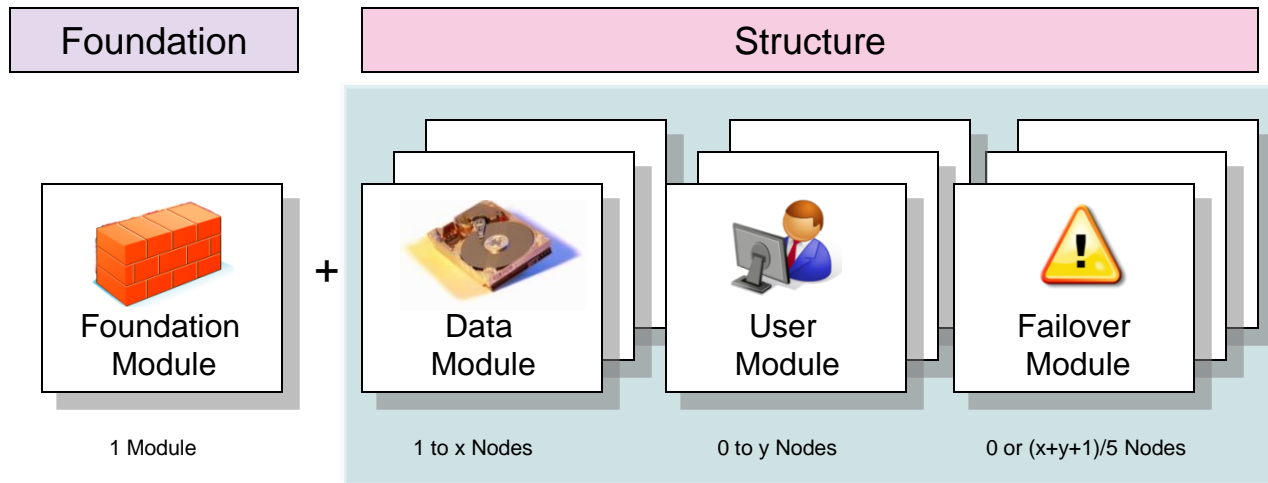
<b>Workload Optimized Analytic System Vs. Custom Deployment</b>		
	<b>Smart Analytics System</b>	<b>Custom</b>
<b>One Call support</b>	<ul style="list-style-type: none"> <li>• <b>Included</b></li> </ul>	<ul style="list-style-type: none"> <li>• <b>Not Available</b></li> </ul>
<b>Coordinated Stack certification (SW, OS, &amp; Firmware)</b>	<ul style="list-style-type: none"> <li>• <b>Included</b></li> </ul>	<ul style="list-style-type: none"> <li>• <b>Not Available</b></li> </ul>
<b>Services Implementation Accelerated Value Prog. Health Check</b>	<ul style="list-style-type: none"> <li>• <b>Included</b></li> <li>• <b>Included</b></li> <li>• <b>1<sup>st</sup> year Included</b></li> </ul>	<ul style="list-style-type: none"> <li>• <b>1 off</b></li> <li>• <b>1 off</b></li> <li>• <b>1 off</b></li> </ul>
<b>Bottom Line</b>	<b>All part of the System</b>	<b>\$ Significantly more expensive</b>

# Months versus weeks



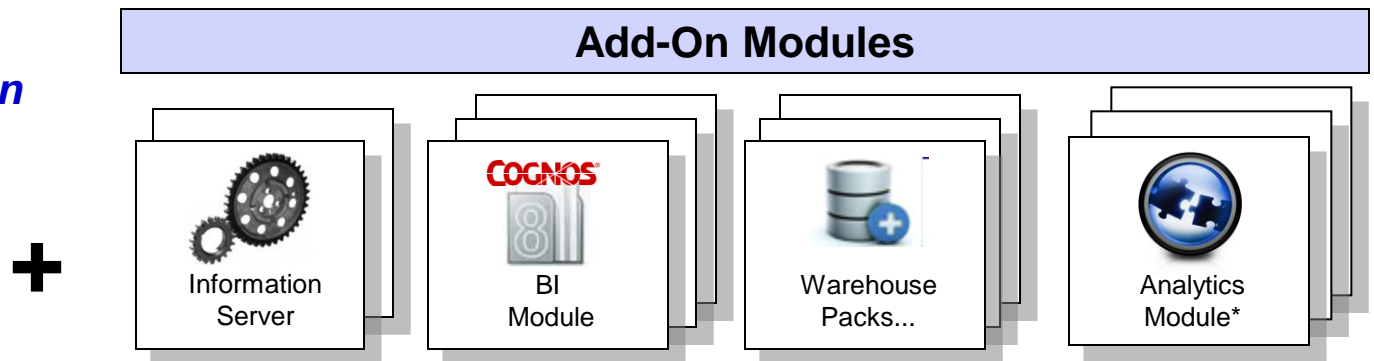
# IBM Smart Analytics System 7700

## *Transparent Modular Architecture*



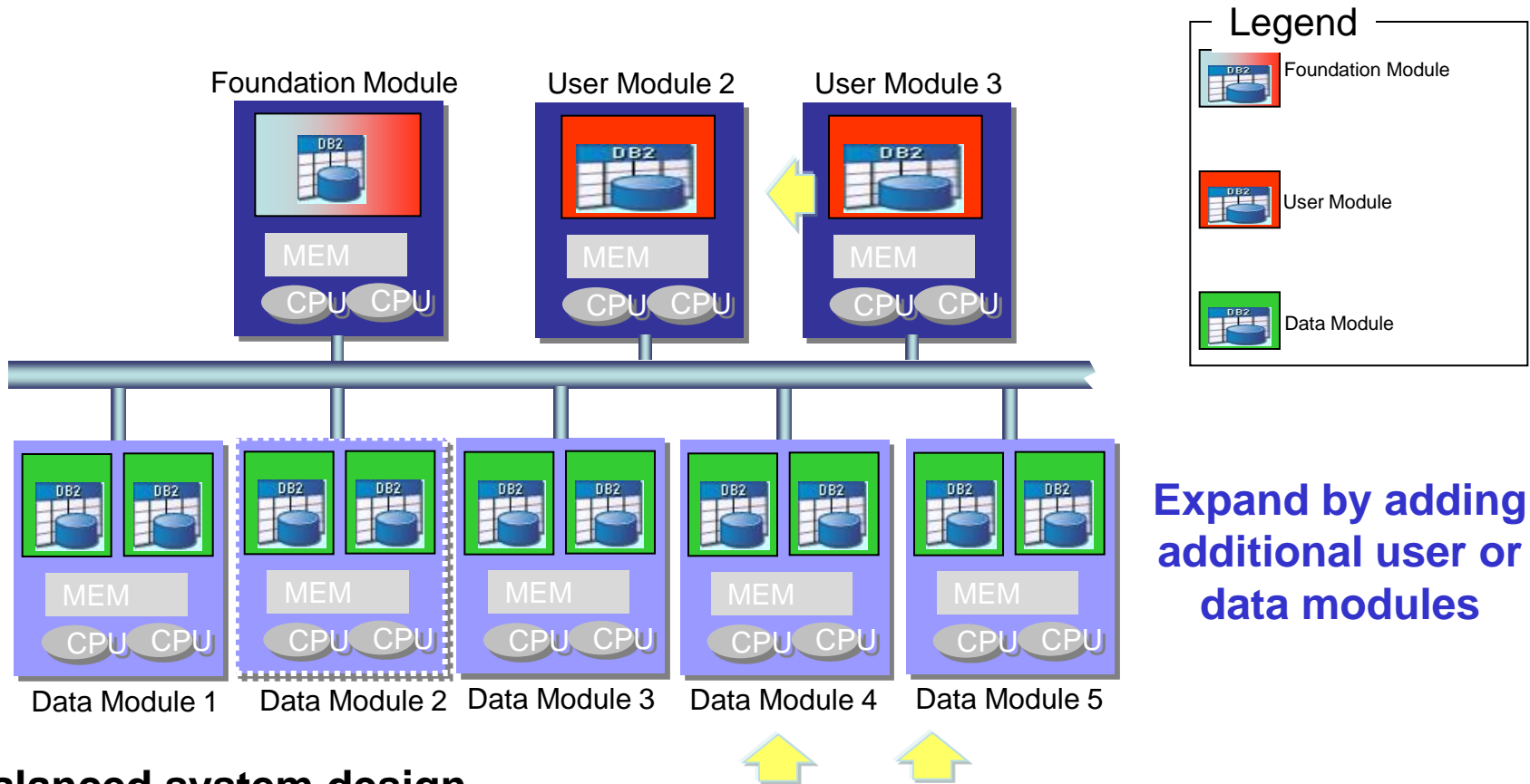
**Modular design**

**+ SSD**



# Scaling out to support more data or more users

## Shared Nothing, Massively Parallel Design



- **Balanced system design**
  - System modules with optimal processor, memory, and I/O specifications
- **Scale-out by adding additional system modules**
  - Which always include balanced I/O
- **Proven “best practice“ for large scale data warehousing**

## 7700 Solid State Disk Exploitation

- Every 7700 Data Module includes 800 GB Solid State Disk
  - Can grow to 4.8 TB per data module
- Configured out of the box to be used for transient data
  - Temporary tables
  - Spilled sort operations
- Significant offload of write IOPS and associated reads from the data disks
- Concurrent workloads with 25% of queries driving spilled sorts have seen 2x improvement in execution time

# InfoSphere Warehouse is powered by DB2

- DB2 offers many unique, industry leading capabilities that are advantageous to data warehousing environments
  - “Shared Nothing” Architecture
  - Advanced Cost Based Parallel Query Optimizer
  - Flexible Partitioning
  - Multi-dimensional Clustering (MDC)
  - Materialized Query Tables (MQT)
  - Industry leading compression
  - Workload management

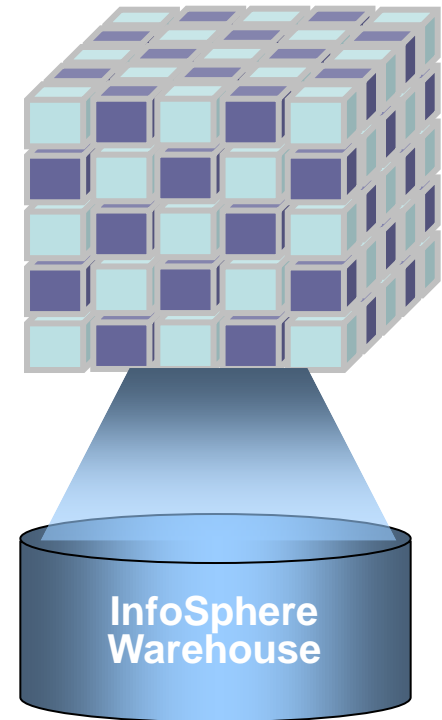


# Online Analytical Processing (OLAP) Cubing Services

Cubing Services is a multidimensional analysis server that enables OLAP applications to access to large data volumes stored inside a DB2 database

## Benefits

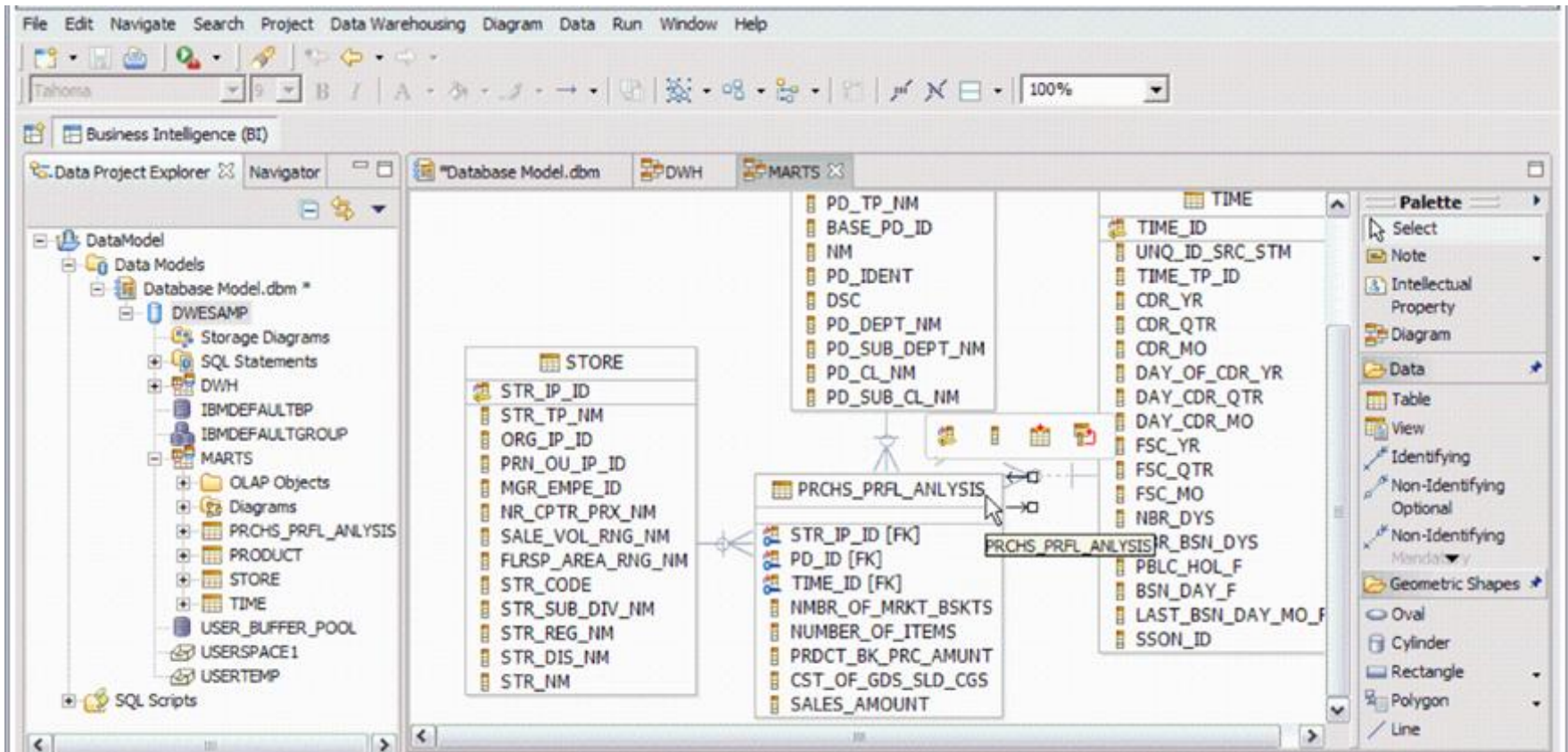
- Empowers users with ad hoc access to business information.
  - What is the profitability for Product A across the Branches X,Y,Z?
- Speed of thought access to OLAP data managed by DB2
- OLAP and SQL shared access to the same information
- Single point of management, maintenance, and performance tuning
- Accessible via Cognos 8 BI, Microsoft Excel, Alphablox, IBM DataQuant, and Cubeware Cockpit





# Modeling & Design

## InfoSphere Warehouse Design Studio



Jump start implementations with industry specific models

# Power Smart Analytics System 7710

**7710 provides an even lower price, size and footprint. Single server implementation of 7700 offering based on LPAR'd architecture. Estimated at \$1M list and 10TB.**

**Targets non-production / development environments along with departmental and small enterprise production environments.**

Comprised of

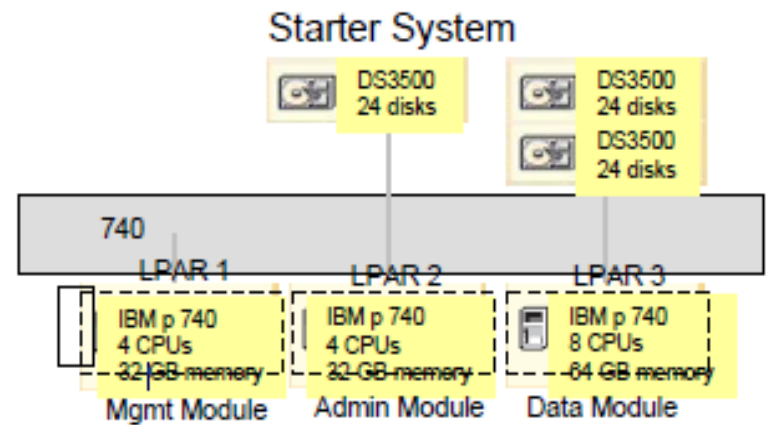
- Smart Analytics System 7710 Enterprise Software
- Smart Analytics System 7700 OS Software
- One POWER7 P740 server 16 core (3 LPARs for each 7700 standard Module support)
- Three DS3524 disk modules
- CSC deployment, Lab Service engagement, Single point of entry support and AVP
- Planned Console, HA and ATK support

Does not support

Add on Modules (Business Intelligence, IIS, Data/User)

Licensing Model

Per Install with Production and Non-Production options



# Questions & Answers





[ibm.com/software/data/infosphere/smart-analytics-system](http://ibm.com/software/data/infosphere/smart-analytics-system)

# Trademarks

**The following are trademarks of the International Business Machines Corporation in the United States, other countries, or both.**

Not all common law marks used by IBM are listed on this page. Failure of a mark to appear does not mean that IBM does not use the mark nor does it mean that the product is not actively marketed or is not significant within its relevant market.

Those trademarks followed by © are registered trademarks of IBM in the United States; all others are trademarks or common law marks of IBM in the United States.

For a complete list of IBM Trademarks, see [www.ibm.com/legal/copytrade.shtml](http://www.ibm.com/legal/copytrade.shtml):

\*, AS/400®, e business(logo)®, DBE, ESCO, eServer, FICON, IBM®, IBM (logo)®, iSeries®, MVS, OS/390®, pSeries®, RS/6000®, S/30, VM/ESA®, VSE/ESA, WebSphere®, xSeries®, z/OS®, zSeries®, z/VM®, System i, System i5, System p, System p5, System x, System z, System z9®, BladeCenter®

**The following are trademarks or registered trademarks of other companies.**

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.

Cell Broadband Engine is a trademark of Sony Computer Entertainment, Inc. in the United States, other countries, or both and is used under license therefrom.

Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

ITIL is a registered trademark, and a registered community trademark of the Office of Government Commerce, and is registered in the U.S. Patent and Trademark Office.

IT Infrastructure Library is a registered trademark of the Central Computer and Telecommunications Agency, which is now part of the Office of Government Commerce.

\* All other products may be trademarks or registered trademarks of their respective companies.

## Notes:

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.

All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.