

Enterprise Data Warehousing

SQL Server 2005

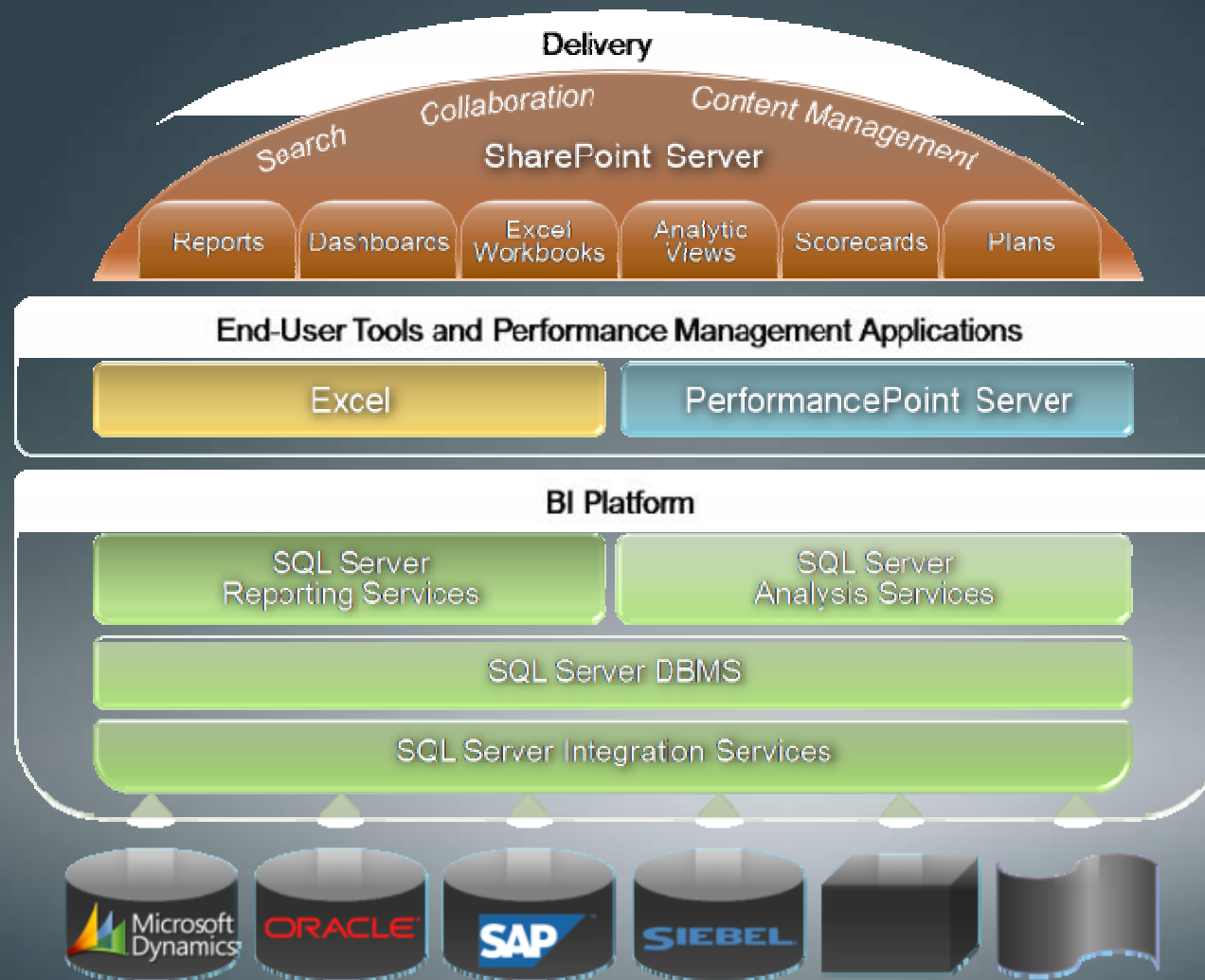
Ron Dunn
Data Platform
Technology Specialist



Business Intelligence Summit 2007

Integrated BI Platform

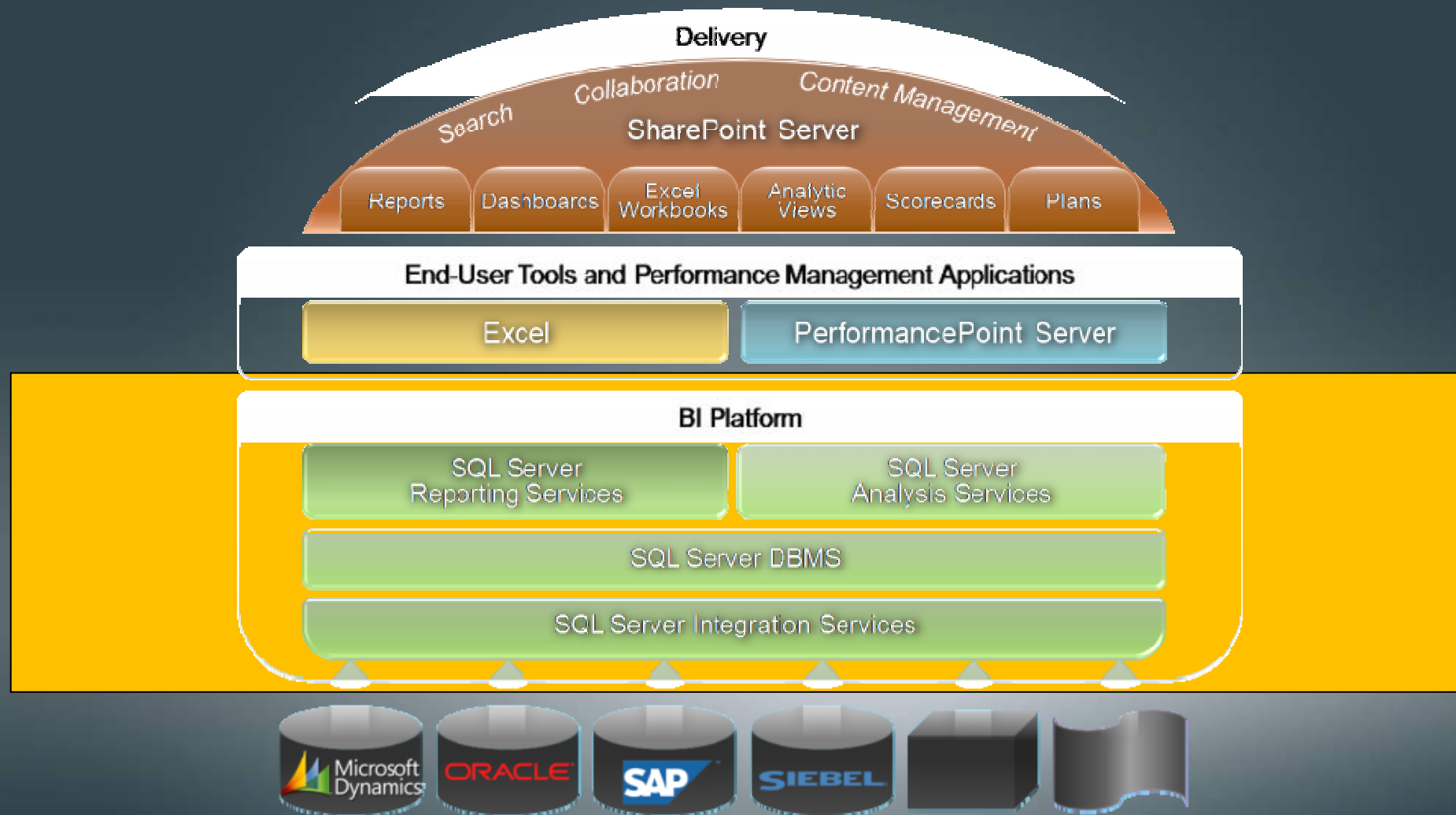
Microsoft



Business Intelligence Summit 2007

Integrated BI Platform

Microsoft



Business Intelligence Summit 2007

Agenda

Microsoft

- **Can SQL Server cope?**
- **Do I need Enterprise Edition?**
- **Will I avoid the maintenance nightmare?**
- **Does it work with my other databases?**
- **Is support available?**

Agenda

Microsoft

- **Can SQL Server cope?**
- Do I need Enterprise Edition?
- Will I avoid the maintenance nightmare?
- Does it work with my other databases?
- Is support available?

Question:

Who has the biggest
data warehouse?



SQL Server 2005 Users

Microsoft



- Problem: POS data loads from 300 million to 4 billion rows,
- Solution: 4 terabytes of data in a single instance with 25 TB of raw storage accessed by over 350 power users.



- Problem: Aging, costly DB2 mainframe data warehouse
- Solution: 2 Terabyte SQL Server data warehouse



- Problem: Covered member base expected to triple in 2006
- Solution: 6 Terabyte warehouse helps improve patient care



- Problem: Identify which members would most benefit from proactive intervention to prevent health deterioration.
- Solution: 5 Terabyte warehouse with over 3000 OLAP users.

SQL Server 2005 Users

Microsoft



- Problem: Complex analyses of credit card customer base.
- Solution: 10 Tb data, growing 300Gb/month



- Problem: Efficient and economical support for analyst growth
- Solution: 6Tb data warehouse, 30% growth pa, 2000 users



- Problem: Centralise and analyse data from 1000 stores
- Solution: 8Tb data warehouse, 500Gb SSAS cubes



- Problem: Improve decision making around profitability, sales and inventory
- Solution: 5Tb data warehouse

SQL Server 2005 Availability

Microsoft



99.999% availability,
12,000 concurrent users



99.999% availability,
500 concurrent users



99.999% availability
433 tps



100% availability
900 tps



99.990% availability
2,000 concurrent users



99.998% availability
3,200 tps



99.997% availability
500 tps

Data Warehouse Needs

- Relational Database
- Multi-dimensional Database
- Extract, Transform, Load (ETL)
- Rich Reporting
- Advanced Analytics

SQL Server 2005 Solution

- **Relational Database**
 - SQL Server 2005 Database Engine
- **Multi-dimensional Database**
 - SQL Server 2005 Analysis Services
- **Extract, Transform, Load (ETL)**
 - SQL Server 2005 Integration Services
- **Rich Reporting**
 - SQL Server 2005 Reporting Services
- **Advanced Analytics**
 - Microsoft Office 2007

Can SQL Server Cope?

- ✓ Scalability
- ✓ Availability
- ✓ Complete, integrated BI product set

Agenda

Microsoft

- Can SQL Server cope?
- **Do I need Enterprise Edition?**
- Will I avoid the maintenance nightmare?
- Does it work with my other databases?
- Is support available?

Standard or Enterprise?

Microsoft

Category	Scenario	Standard	Enterprise
Performance & Scalability	Small to Medium Scale OLTP Systems	✓	✓
	High Transactional/Large Scale OLTP Systems		✓
	Server Consolidation		✓
	Data Warehousing		✓
High Availability	Reduce Downtime due to HW/SW Failure	✓	✓
	Zero downtime during maintenance		✓
	Reduce downtime due to human error		✓
	Database Mirroring for High Transactional Systems		✓
Data Integration & Cleansing	Integration from Multiple Data Sources	✓	✓
	Data Cleansing		✓
	Replication with Oracle		✓
Business Intelligence	High Scale & Real Time Analysis and Reporting		✓
	Complex Analytic & Data Mining		✓
	Report Server Scale-out Deployment		✓

Scalability and Performance

Microsoft

Feature	Standard	Enterprise
Number of CPUs	4	64
Memory	OS Maximum	OS Maximum
Multi-Instance Support	16	50
Dynamic AWE Memory Management		✓
Table and Index Partitioning		✓
Database Snapshot		✓
Scalable Shared Databases		✓
Indexed Views		✓
Parallel Index Operations		✓
Parallel Consistency Checks (DBCC)		✓
Enhanced Read-ahead and Scan (Super Scan)		✓

High Availability

Microsoft

Feature	Standard	Enterprise
Failover Clustering	2-nodes Only	Up to 8 nodes
Database Mirroring	Safety FULL Only	All Modes
Database Snapshot		✓
Fast Recovery		✓
Online Index Operations		✓
Online Page and File Restore		✓
Parallel Index Operations		✓
Updateable Distributed Partitioned Views		✓
Table and Index Partitioning		✓

Business Intelligence

Microsoft

Category	Scenario	Standard	Enterprise
Data Mining	Parallelism for Model Processing		✓
	Parallelism for Model Prediction		✓
	Algorithm Plug-in API		✓
	Advanced Configuration & Tuning Options for Data Mining Algorithms		✓
Advanced Analytics	Account Intelligence		✓
	Linked Measures and Dimensions		✓
	Perspectives		✓
	Semi-additive Measures		✓
	Writeback Dimensions		✓
Scalability & Performance	Proactive Caching		✓
	Auto Parallel Partition Processing		✓
	Partitioned Cubes		✓
	Distributed Partitioned Cubes		✓

Business Intelligence

Microsoft

Category	Feature	Standard	Enterprise
Integration Services	Data Mining Query Transformation		✓
	Data Mining Model Training Destination Adapter		✓
	Fuzzy Grouping Transformation		✓
	Fuzzy Lookup Transformation		✓
	Term Extraction Transformation		✓
	Term Lookup Transformation		✓
	Dimension Processing Destination Adapter		✓
	Partition Processing Destination Adapter		✓
Reporting Services	Data-driven Subscriptions		✓
	Report Scale-out Deployment		✓
	Infinite Clickthrough in Ad-hoc Reports		✓

Enterprise Edition?

- ✓ **Performance and Scalability**
- ✓ **High Availability**
- ✓ **Data Integration and Cleansing**
- ✓ **Business Intelligence Capability**

Agenda

- Can SQL Server cope?
- Do I need Enterprise Edition?
- **Will I avoid the maintenance nightmare?**
- Does it work with my other databases?
- Is support available?

How many DBAs?

Typical industry metrics:

- DB2 1+ DBAs per database
- Oracle 10 databases per DBA
- SQL Server 100 databases per DBA

Manage

Microsoft

Scale and Manage your large data warehouse

- Partition your large data sets into manageable chunks
- Improved Query Optimizer to provide scalable query performance
- Simplify Aggregations using Indexed Views and compact Cubes
- Streamlined Index Management to reduce maintenance windows
- Management tools to identify and resolve potential issues
- Bring your Data Warehouse online faster by backing up and restoring active partitions only



Manageability Tools

Microsoft



- Management Studio to manage your entire BI infrastructure
- Dynamics Management Views to get real time insight into the database
- SQL Query Analyzer provides prescriptive guidance on tuning query performance
- Management Objects allows integration into existing management infrastructure

How much downtime?

- Online Backups
- Online re-orgs
- Online loads

Maintainable?

- ✓ **Fewer support staff**
- ✓ **Less configuration and tuning**
- ✓ **Less downtime**

Agenda

Microsoft

- Can SQL Server cope?
- Do I need Enterprise Edition?
- Will I avoid the maintenance nightmare?
- **Does it work with my other databases?**
- Is support available?

Integration Options

- **Bulk Import**
- **Linked Servers**
- **Replication**
- **ETL**

Traditional ETL

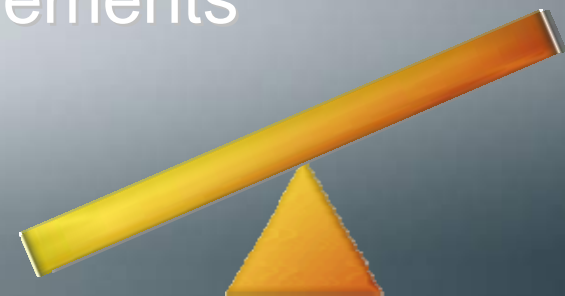
- Traditional Solutions
 - Shrink-wrapped Tool
 - Or Build a Custom Solution
 - Or Hybrid (Staged Data + Glue)



Limitations of Traditional ETL

Microsoft

- Integration Today
 - Increasing Data Volumes
 - Increasingly Diverse Sources
 - More Users & Use Cases
- Requirements reached the Tipping Point
 - Traditional Tools unable to Deliver
 - More Stringent Business Requirements
 - Difficult to Maintain Agility
 - Eroded Ability to Compete



SQL Server Integration Services

Microsoft



Enterprise data integration

- Connectivity to any source
- Scalable pipeline architecture
- Comprehensive set of transformations
- DW smarts handling Slowly changing dimensions and direct loading of cubes
- Clean and Validate using data mining and data quality



 **.NET Framework**

Extensible platform

- Pluggable custom tasks
- Componentized development using .NET

Data Integration?



Application Systems

- SAP, Siebel, Peoplesoft



Relational DBMS

- SQL Server, Oracle, DB2, Sybase



Semi-structured Data

- Excel, Text, XML, EDI



Queues and Protocols

- MSMQ, (s)FTP, HTTP(S)

Agenda

- Can SQL Server cope?
- Do I need Enterprise Edition?
- Will I avoid the maintenance nightmare?
- Does it work with my other databases?
- **Is support available?**

{database} “business intelligence”

- SQL Server 74
- Oracle 68
- DB2 20
- Teradata 6

{database} data warehouse

- SQL Server 51
- Oracle 36
- DB2 5
- Teradata -

Some Local Partners

- Angry Koala
- Avanade (Accenture)
- Bearing Point
- BiQ
- Biztools
- Capiotech
- Centura
- ComOps
- Data#3
- Dimension Data / DataCraft
- Evolution
- F1
- HP
- Interline
- James and Munro
- Oakton
- Pollak Partners
- Power Business Solutions
- Praxa
- Prima
- Professional Advantage
- Red Rock
- SDM
- Spirit Lake
- Unique World

3rd. Party Software

Microsoft

- SQL Server is the biggest installed base for our *competitors'* products!

Support?



Books



Staff



Training courses



Contractors and consultants



3rd party solutions

Agenda Summary

- ✓ **Can SQL Server cope?**
- ✓ **Do I need Enterprise Edition?**
- ✓ **Will I avoid the maintenance nightmare?**
- ✓ **Does it work with my other databases?**
- ✓ **Is support available?**

Industry Accolades

Microsoft®

"The 2005 results reveal significant strengths for Microsoft SQL Server in the very large database (VLDB) market for both **data warehouse** (DW) and transaction processing (OLTP) systems."



"They can scale and handle data warehouses of any size; we know that now,"... "I routinely talk to SQL Server customers that have built 7TB data warehouses or Windows Datacenter [systems] with 3,000 users."

Gartner

Donald Feinberg
VP & Distinguished Analyst
Information Week

"The recent release of the 2006 Gartner DW DBMS Magic Quadrant demonstrates the **significant progress Microsoft** has made in the Data Warehouse (DW) Database Management Systems (DBMS) market, moving from deep placement in the challenger's quadrant to riding the line between leader and challenger."

Gartner

"Microsoft's BI Tools revenue growth in 2005 was more than 25 percent, growing at more than twice the rate of the overall market. Microsoft has seen strong growth over the past several years as it has expanded and enhanced its database-embedded BI features **Microsoft's impact on the BI tools market cannot be overemphasized.** " *Dan Vesset*



Microsoft

Your Questions



Business Intelligence Summit 2007