

# IBM BI Solution with Microsoft® SQL Server 2008®

---

## Business Intelligence Solution on IBM System x™ 3950M2 Server and IBM System Storage™ DS5000 with Microsoft® SQL Server 2008®

### Highlights:

- ❑ **IBM System x™ provides the most scalable Intel-based system**
- ❑ **IBM System Storage™ provides high performance and robust disk solutions for mission-critical applications**
- ❑ **Microsoft SQL Server 2008® provides all the components needed to build a comprehensive low-cost BI solution**

**IBM System x3950 M2 server** is based on X4 – the 4<sup>th</sup> generation of the proven IBM Enterprise X-Architecture® chipset. The X4 chipset including the Hurricane4 memory controller provides the x3950 M2 server with the ability to scale-up based on a modular “pay-as-you-grow” design, high performance as demonstrated by leadership TPC benchmarks, and mainframe-like reliability in an x86 environment.

### Key features

- True 2-to-16-socket scalability up to 64 cores
- Revolutionary Intel Xeon dual-core and quad-core MP 7300 Series processors
- Up to 1TB of registered DIMM memory for better workload density and up to 20–30% less power consumption than competitors' fully buffered DIMM technology\*
- IBM Memory ProteXion™ with redundant bit-steering offers twice the memory resilience of the competition
- 4th generation snoop filter 4 times larger than the competition's best
- IBM Predictive Failure Analysis®, not just on hard drives and memory but, unlike competitors, also on processors, power supplies, fans, and voltage regulator modules
- 40% lower memory latency than the nearest competition
- More flexible memory configurations than competitors, at significantly lower costs
- IBM System x3950 M2 servers provide an uncomplicated, cost-effective and highly flexible solution. With the ability to scale while maintaining balanced performance between processors, memory and I/O, these servers can easily accommodate business expansion and the resulting need for additional application space

**The NEW IBM System Storage™ DS5000** sets new standards for performance, scalability, reliability, availability, and flexibility for midrange storage systems. As IBM's most powerful midrange storage system, the DS5000 is the ideal platform for a database environment that can keep pace with an organization's business growth. Organizations can buy only the capacity needed initially, and can then dynamically upgrade and reconfigure additional capacity & features later to meet changing business requirements, all without any system downtime. The DS5000 delivers class-leading performance and is equally adept at supporting transactional-

# IBM BI Solution with Microsoft® SQL Server 2008®

---

applications, such as databases and OLTP, throughput-intensive applications, such as HPC and rich media, and concurrent workloads, well-suited for consolidation and database. With its relentless performance and architected to provide the highest reliability and availability, the DS5000 storage system comfortably supports the most demanding service level agreements (SLAs). And when requirements change, the DS5000 can easily be reconfigured “on-the-fly” to add or replace host interfaces, increase performance, grow capacity, or add cache – ensuring it keeps pace with a growing company.

## DS5000 key features

- Flexible host interface options are 8 Gb/s Fibre Channel and 10 Gb/s iSCSI ready
- Sixteen 4 Gb/s Fibre Channel drive interfaces for support up to 256 drives in initial release, with future support for 448 FC/SATA drives, using EXP5000/EXP810 drive expansion units.
- Up to 16 GB of dedicated data cache (8 GB per controller) in initial release, with future support for 32 GB. Dedicated cache mirroring channels and persistent cache backup in the event of a power outage
- Support for RAID 6, 5, 3, 10, 1, 0
- Two performance levels (base and high) with ability to field-upgrade
  - Base model is DS5100 – High end model is DS5300
- Remote Volume Mirroring and FlashCopy premium features for Volume Shadow Copy (VSS) supported backups and flexible DR scenarios
- Break-through performance levels over 5X greater than the DS4800

**Microsoft SQL Server 2008®** running on Microsoft Windows Server 2008® enables organizations to build comprehensive, enterprise-scale analytic solutions that deliver intelligence where customers want it.

## Key features

- 64-bit native DBMS
- Dynamic large memory support
- High Availability via Failover Clustering, Database Mirroring, Log Shipping and Replication
- Database snapshots, a point in time, instantaneous read-only copy of a database which can be used for reporting and reverting back
- Data compression for reduced on disk space utilization, faster data retrieval and faster backups and restores
- Management Data Warehouse for monitoring SQL Server performance across the enterprise
- Policy Based Management for managing SQL Server instances across the enterprise
- Includes all the components needed to provide a comprehensive low cost Business Intelligence solution

## SQL Server 2008® BI Components

SQL Server 2008 includes the SQL relational database engine, Analysis Services, Integration Services and Reporting Services. Together these components form a complete data warehousing and business intelligence solution. The Database Engine provides the services for the relational databases which includes the operational databases and the data warehouses. Integration Services provides the means for extracting, transforming and loading (ETL) the data

from the operational databases into the data warehouses. Analysis Services provides the means for creating, processing and serving OLAP cubes. A cube is a highly structured multidimensional database that contains fact data and aggregations associated with dimensions which can be organized into hierarchies. An example is product sales information aggregated by Time (Year, Month, Day) and Product (Bikes, Vendor, Model). A cube can be queried efficiently by a decision maker to explore and plan new business actions. Cubes can be queried using the industry standard MDX (Multidimensional Expressions) query language. Several products support OLAP query, including Reporting Services. Reporting Services provides a comprehensive reporting solution for SQL Server data stored in relational databases and cubes.

## SQL Server® Analysis Services

Microsoft SQL Server 2008® Analysis Services builds on the value delivered with the significant investments in Analysis Services 2005 around scalability, advanced analytics, and Microsoft Office interoperability. Through substantially improved performance, scalability, and developer productivity, customers can build enterprise-scale Online Analytical Processing (OLAP) solutions. The Unified Dimensional Model consolidates data access and provides a wide range of analytical capabilities, while deep integration with the Microsoft Office system and an open, embeddable architecture allows customers to reach every user with familiar tools and drives actionable insight to users across the enterprise.

## Business Intelligence enhancements in SQL Server 2008®

- Improved Data Warehouse performance – star join optimization (bitmap filter), improvements in partitioning – more parallelism, partition-aligned indexed views
- Change tracking – support for automatically tracking database changes for efficiently updating the Data Warehouse.
- New MERGE statement – process inserts, updates, deletes in one pass for efficiently bringing the Data Warehouse up to date.
- Improved tools for designing cubes, dimensions and aggregations.
- Improved backup and restore performance.

**Emulex 42D0494 (LPe12002) Fibre Channel host bus adapters** (HBAs) provide streamlined installation and management, outstanding scalability and industry-leading database support well-suited for small-to-large enterprises and Microsoft Windows Server 2008 and SQL Server 2008 storage area network (SAN) environments. With powerful management tools and broad System x and support, the LightPulse family of IBM-branded 4Gb/s and 8Gb/s HBAs (IBM Server Proven Validation) delivers high performance for a broad range of applications and environments.

## Key features

- Exceptional performance and full-duplex data throughput
- Comprehensive database capabilities with support for N-Port ID Database (NPID)
- Simplified installation and configuration using AutoPilot Installer®
- Administration via HBAnyware® integrated with IBM Director

# IBM BI Solution with Microsoft® SQL Server 2008®

---

## Goal

The goal of this paper is to report on a series of tests conducted to test the performance characteristics of building and processing cubes.

## System Configuration

<b>Server</b>	<b>IBM System x3950 M2 4 Quad core Intel Xeon X7350, 2.93GHz, 2x4MB L2 cache, 16MB L3 cache</b>
<b>Memory</b>	<b>64 GB</b>
<b>HBAs</b>	<b>3 Emulex LPE12002 8Gb dual-port FC adapters</b>
<b>Storage</b>	<b>IBM DS5000 controller, 13 EXP 810 enclosures, 208 300GB 15K rpm drives</b>
<b>Software</b>	<b>Windows server 2008 x64 Enterprise edition, version 6.0.6001, Service Pack 1 SQL Server 2008 Enterprise Edition, version 10.0.1600.20</b>

## Server configuration

The single node x3950 M2 server used for these tests was connected to the IBM DS5000 storage controller with 6 direct connections from the 3 Emulex LPe 12002 M8 HBAs to 6 ports on the controller. The Emulex LPE 12002 HBAs are 8GBps adapters. The first generation of DS5000 controller used in this configuration is only capable of running at 4GBps making the connections run at 4GBps. The next release of IBM DS5000 will support 8GBps. The Emulex LPE 12002 HBAs and the accompanying drivers have built-in support for MSI-X and NUMA IO and take advantage of NUMA IO improvements in Windows server 2008. There are 7 PCIe slots on the IBM x3950 M2 server. Slots 1, 3 and 5 were populated with the Emulex dual-port adapters. The Microsoft Multi-Path Bus Driver was used to manage the multiple paths. The DS3000/DS4000 DSM was used with the 'least queue depth' policy. The least queue depth policy compensates for uneven loads by distributing proportionately more I/O requests to lightly loaded processing paths. Other policies available are 'fail over only', 'round robin' and 'weighted paths'.

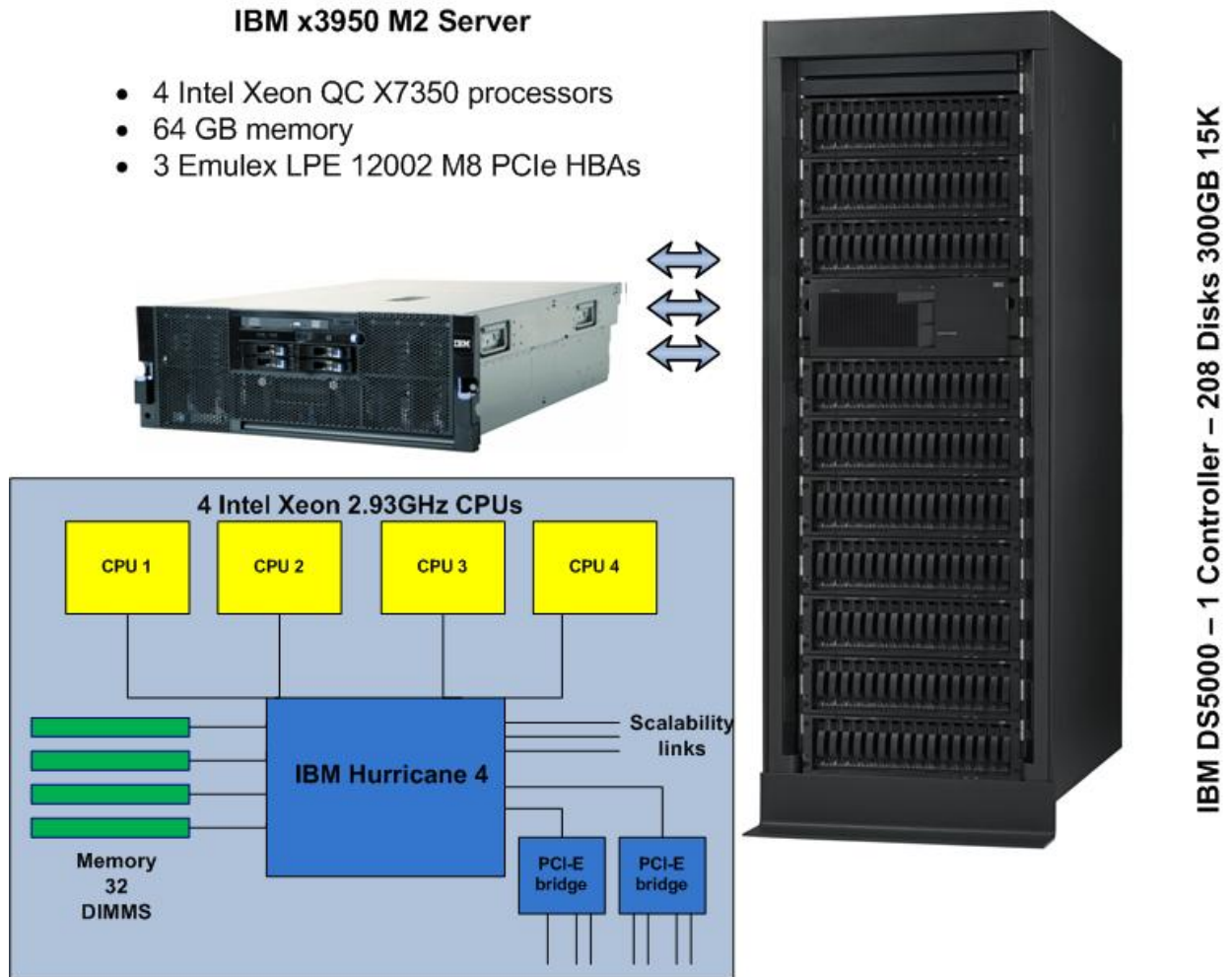
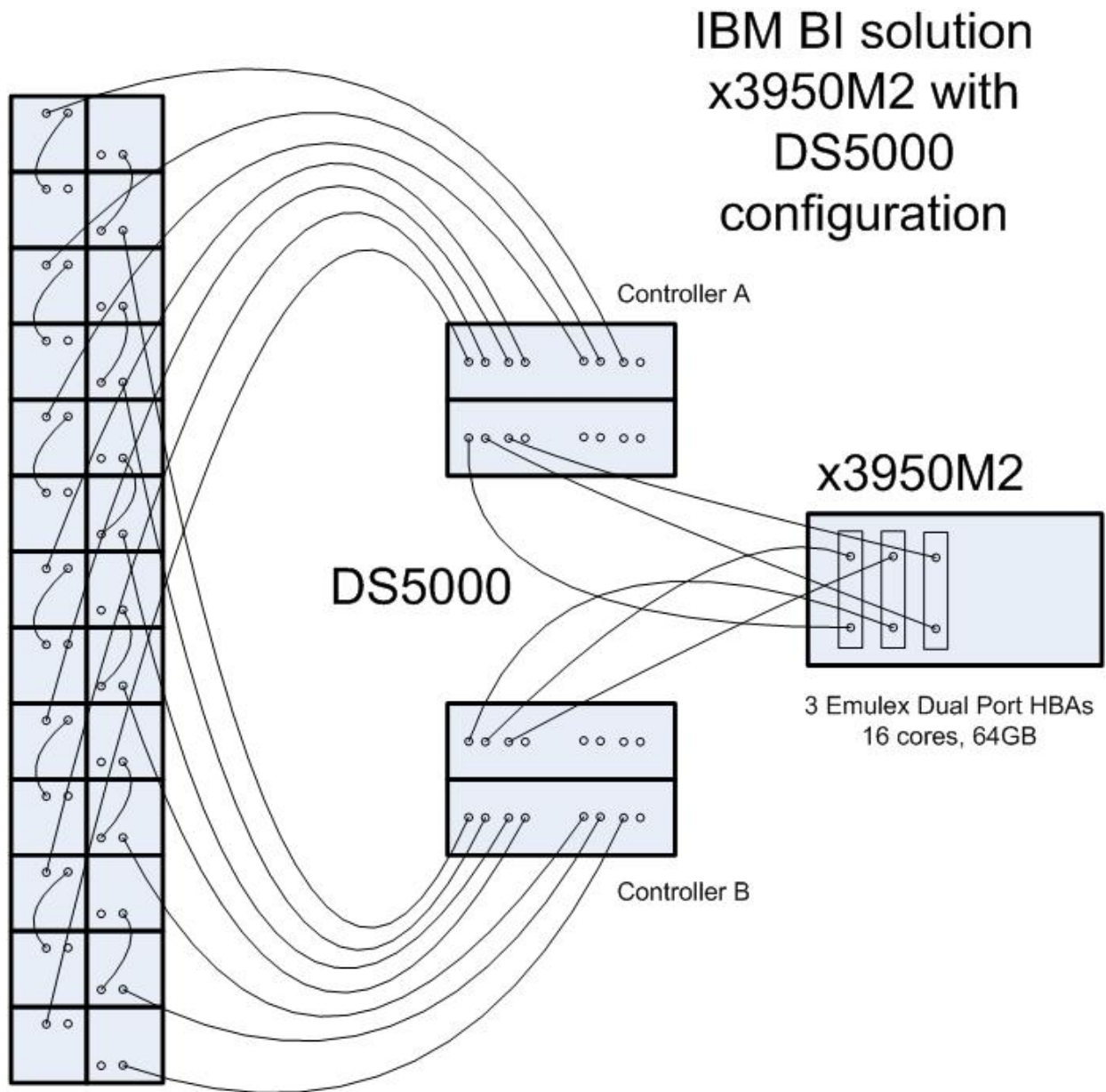


Figure 1. Server and storage configuration

## Storage configuration

The IBM DS5300 dual-controller had 13 IBM EXP810 enclosures attached to it. Each of these EXP 810 enclosures was populated with 16 300GB 15K rpm drives for a total of 208 drives. One of these enclosures with 16 drives was set up as a RAID-5 LUN for archiving the backup of the database. The data warehouse consisted of 14 months of data. 14 RAID-5 LUNs of 7 disks each were carved out to hold these 14 months of data. 4 log LUNs were set up with 4 disks each. 2 miscellaneous LUNs were carved out for the remaining data in the DW. The rest of the disks, 6 disks each on 14 enclosures for a total of 84 disks, were set up as one big RAID-10 LUN for the OLAP cube. Figure 2 shows the cabling between the x3950 M2, DS5300 and the EXP 810 enclosures. Figure 3 shows the OLAP Cube LUN configuration.



13 EXP810 Enclosures  
208 Disks, 300 GB, 15K rpm

Figure 2. x3950M2 server and DS5300 storage cabling Diagram

## DS5300 LUN Configuration

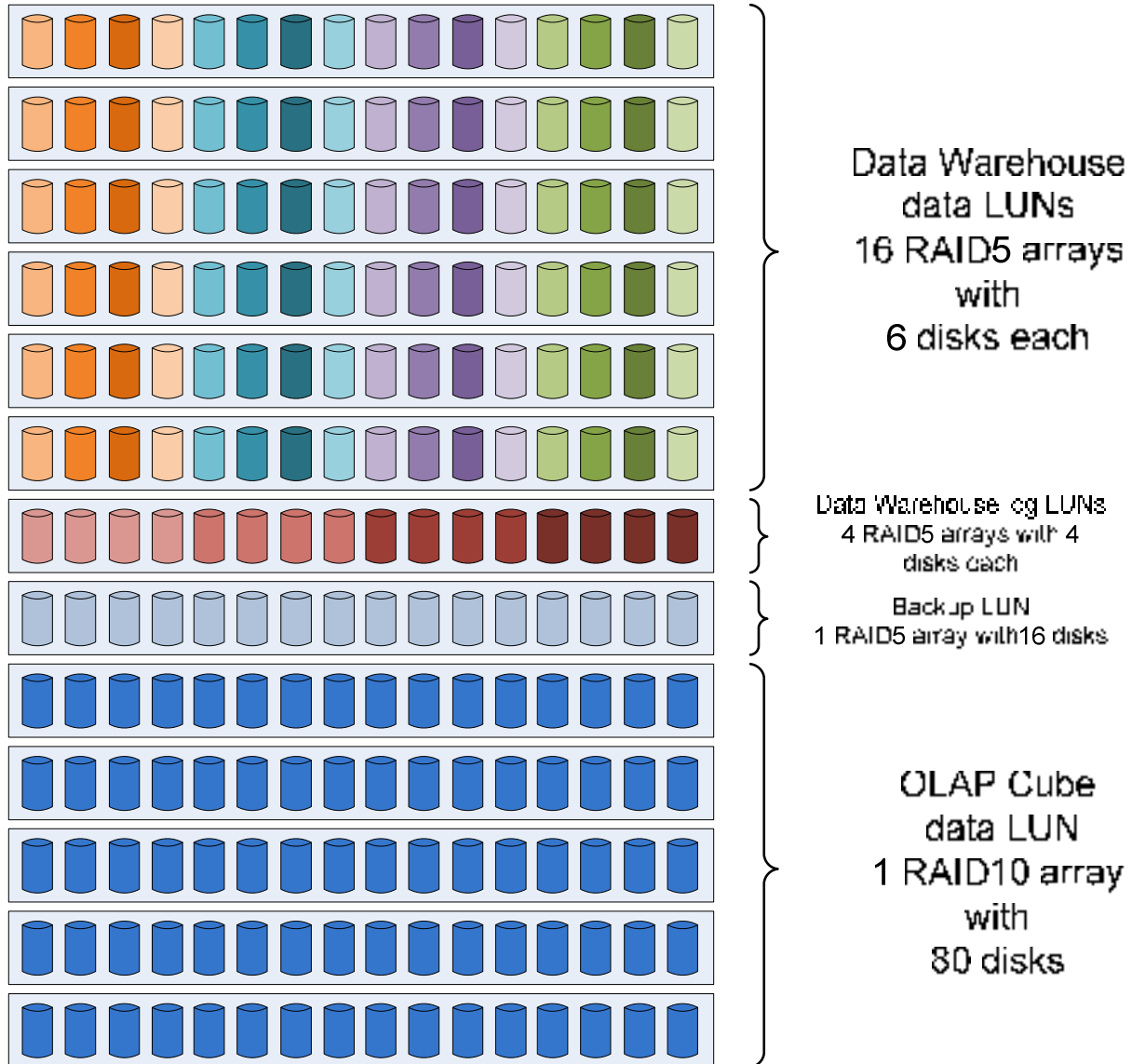


Figure 3. DS5300 LUN configuration

### BI workload size

The DW database size used in the test is around 1.77 TB and the Cube size ended up to be around 412GB.

## SQL Server 2008® configuration

SQL server 2008® database engine and Analysis Services were installed on one SMP server and run in a shared configuration. CPU and memory for the database engine were limited by using the sp\_configure settings for 'affinity mask' and 'max server memory'. Of the 16 CPUs and 64GB memory on the server, the database engine was limited to 8 processors and 28000MB memory. The following sp\_configure settings were changed:

```
sp_configure 'affinity mask', 0xFF
go
sp_configure 'max server memory', 28000
go
```

## Cube Processing

An Analysis Services Cube is a multidimensional structure that enables customers to access high volumes of pre-aggregated data and extract useful data very quickly. A cube is a set of related measures and dimensions that is used to analyze data. When a cube is processed, the aggregations designed for the cube are calculated and the cube is loaded with these calculated aggregations and data. The tests in this configuration used a data warehouse of web tracking information as a data source to create a multidimensional cube. A simple Cube object is composed of: basic information, dimensions, and measure groups. Basic information includes the name of the cube, the default measure of the cube, the data source, the storage mode, and others.

Dimensions are the actual set of dimensions used in the cube. All dimensions have to be defined in the dimensions collection of the database before being referenced in the cube. Measure groups are sets of measures in the cube. A measure group is a collection of measures that have a common data source view and a common set of dimensions. A measure group is the unit of process for measures; measure groups can be processed individually and then browsed. The generated OLAP cube is then processed with default measure groups and partition sizes.

## Results

### Cube Processing times

#### SQL Server 2005, 16 cores, 64 GB memory, DW capped at 8 cores, 28000MB memory

Phase	Time to process hh:mm:ss	Time to process in minutes	size GB
Measure group one	5:19:03	319	123
Measure group two	8:23:54	504	121
Measure group three	3:45:43	226	146
Total	17:28:40	1049	412



# IBM BI Solution with Microsoft® SQL Server 2008®

SQL Server 2008, 16 cores, 64 GB memory, DW capped at 8 cores, 28000MB memory

Phase	Time to process hh:mm:ss	Time to process in minutes	size GB
Measure group one	3:22:27	202	123
Measure group two	4:09:27	249	121
Measure group three	2:42:09	162	146
Total	10:14:03	614	412

## SQL server 2005 to SQL server 2008 improvements

Phase	SQL 2005 minutes	SQL 2008 minutes	Time Reduction	Percentage Improvement
Measure group one	319	202	117	37
Measure group two	504	249	254	50
Measure group three	226	162	64	28
Total	1049	614	435	41

There is a significant 41% reduction in cube processing times by upgrading from SQL server 2005 to SQL server 2008.

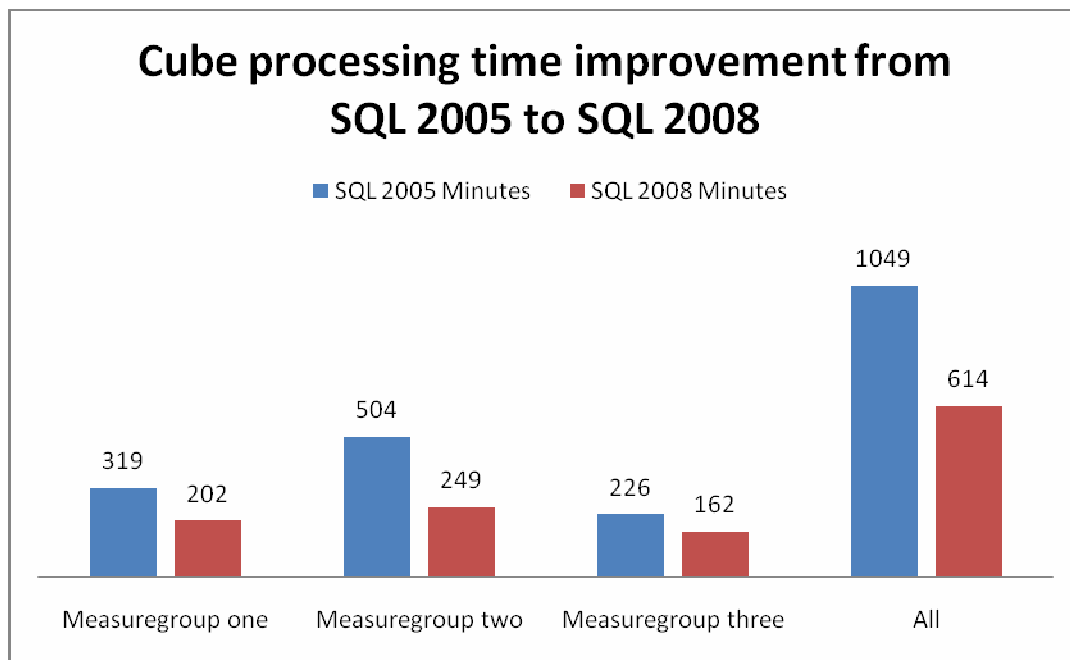


Figure 4. Cube processing improvement

## Conclusion:

Microsoft SQL Server 2008® ushers in a new era of application database affordability to the masses, offering new opportunities for increased resource utilization, ease of management, and improved ROI. IBM has worked closely with Microsoft to ensure our products are optimized for SQL Server 2008® deployments.

IBM System x™ 3950 M2 server with IBM System Storage™ DS5000 storage running Microsoft Windows Server 2008® and Microsoft SQL server 2008® provide a great platform for customers to design a cost effective scalable BI solution in a traditional x86 environment.

Copyright © 2008 by International Business Machines Corporation. This document could include technical inaccuracies or typographical errors. IBM may make changes, improvements or alterations to the products, programs and services described in this document, including termination of such products, programs and services, at any time and without notice. Any statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only. The information contained in this document is current as of the initial date of publication only, and IBM shall have no responsibility to update such information. Performance data for IBM and non-IBM products and services contained in this document was derived under specific operating and environmental conditions. The actual results obtained by any party implementing and such products or services will depend on a large number of factors specific to such party's operating environment and may vary significantly. IBM makes no representation that these results can be expected or obtained in any implementation of any such products or services. THE INFORMATION IN THIS DOCUMENT IS PROVIDED "AS-IS" WITHOUT ANY WARRANTY, EITHER EXPRESS OR IMPLIED. IBM EXPRESSLY DISCLAIMS ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR INFRINGEMENT. References in this document to IBM products, programs, or services does not imply that IBM intends to make such products, programs or services available in all countries in which IBM operates or does business. Any reference to an IBM program or product in this document is not intended to state or imply that only that program or product may be used. Any functionally equivalent program or product, that does not infringe IBM's intellectually property rights, may be used instead. It is the user's responsibility to evaluate and verify the operation of any non-IBM product, program or service. The provision of the information contained herein is not intended to, and does not grant any right or license under any IBM patents or copyrights. Inquiries regarding patent or copyright licenses should be made, in writing, to: IBM Director of Licensing IBM Corporation North Castle Drive Armonk, NY 10504-1785 U.S.A. IBM, the IBM logo, System x, and System Storage are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries or both. Microsoft, Windows and SQL Server are trademarks of Microsoft Corporation in the United States, other countries or both.