

The IBM® DB2® Integrated Cluster Environment leads the way with 3-tier, 32-node SAP Business Information Warehouse benchmark results

An SAP BW 3.5 and DB2 Universal Database™ (DB2 UDB) V8.2 benchmark overview



*Benita Hsueh
IBM Toronto Lab*

1. Introduction

Historically seen, the concept of data warehousing typically brings to mind complexity and expensive hardware. The introduction of Linux® clustering changed that view over the last few years by using concepts from the “Linux Beowulf” High Performance Computing cluster¹ in the enterprise arena. The IBM DB2 Universal Database (DB2 UDB) product family is the known scalability leader for databases and, to be more specific, data warehouses. This combination of Linux, DB2 UDB, and hardware, provided through the DB2 Integrated Cluster Environment (ICE) solution stack, drives a completely new level of simplification to the marketplace, allowing companies of every size to deploy scalable data warehouses that require very limited resource investments.

The SAP Business Information Warehouse (BW) is one of the most complex pre-packaged data warehouse application frameworks because of its feature richness and deployment. Moreover, its tight integration into the SAP enterprise resource planning (ERP) landscape makes it easy to manage. For many years, the default answer to deploying data warehouses meant large scale-up SMP technology because a scale-out solution was considered too difficult – even though DB2 UDB has been able to do so for a long time. This paper describes the ultimate proof, with the SAP BW benchmark, that the days of scale-up as the ultimate answer to data warehousing are over and that, in reality, scale-out with Linux is the “must-do” choice for any serious implementation.

Furthermore, the DB2 ICE provides an added advantage with its simple uniform pre-packaged building blocks that can make planning a cluster easier than planning a large SMP system. The hardware and software environments used in this SAP BW benchmark, along with its published results, will show that DB2 ICE can provide the infrastructure to meet all of your business information warehousing needs and keep implementation time short.

2. SAP BW Benchmark Overview

Business information warehousing is a means for companies to take advantage of the data they gather to better understand their business and how to increase efficiency, retain customers, and grow revenues. SAP Business Intelligence Warehouse (BW) grew within the SAP product offering from the need for more flexible and powerful reporting capabilities to process the data stored in the SAP ERP R/3 systems. SAP BW is fully integrated within the SAP NetWeaver 2004 portfolio and can be used as a standalone component; it is the central reporting tool for almost all SAP business intelligence solutions.

The SAP BW information model uses InfoObjects as its main building block. InfoObjects are used to describe business processes and information requirements and can be reused in other key elements of the SAP information model such as InfoSources, Operational Data Store (ODS) objects, and InfoCubes. InfoSources are groups of InfoObjects that logically belong together from a business point of view. They can contain transactional data from Online Transactional Processes (OLTP) and master data, which contains data such as customer addresses that do not change over time. ODS objects are data sets resulting from the merging of one or more InfoSources and are used for day-to-day decision making, preparing reports, and quality assurance purposes. InfoCubes are defined and created using an extended star schema database design to organize data for reporting and analysis. A basic star schema contains independent dimension tables that are linked together through a fact table. The extended star schema adds to the star schema capability by also storing master data about attributes, hierarchies, and text, in separate tables, which are shared between InfoCubes. This means that data redundancy is reduced because master data only needs to

¹ For more information on the Beowulf cluster project, visit <http://www.beowulf.org>

be stored once but can be used by various InfoCubes. Each InfoCube is made up of a fact table and a set of dimension tables. Fact tables contain data that describes specific events within a business and are the largest tables in a star schema; they can contain billions of entries. Dimension tables contain further information about attributes in a fact table but are independent of each other and are linked only through fact tables. **Figure 1** shows what an extended star schema looks like.

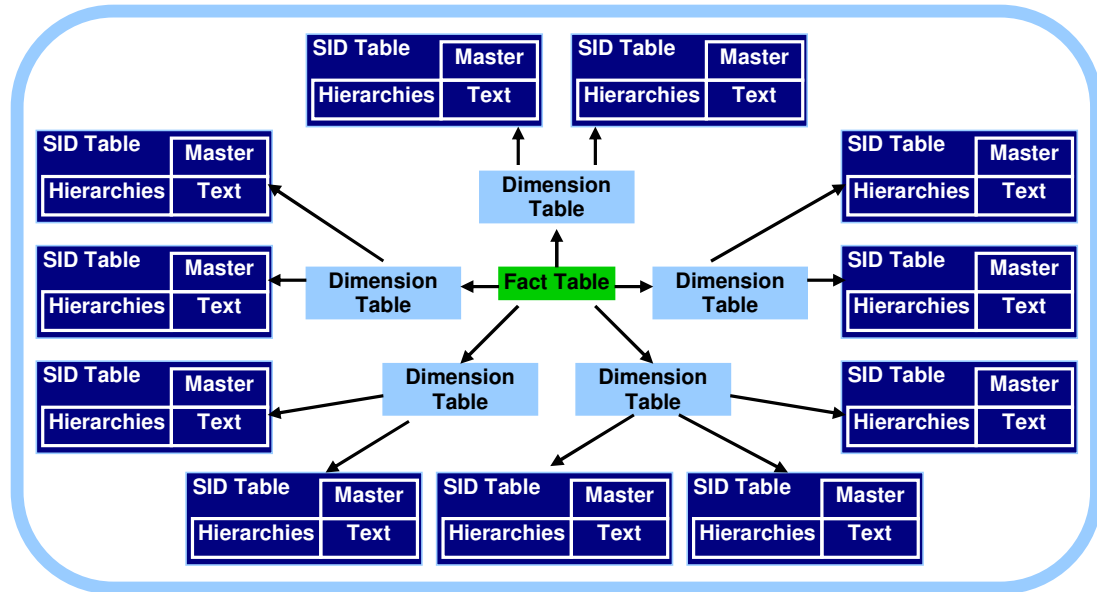


Figure 1 : Extended Star Schema

SAP benchmarks are continuously being improved to better fit real-world scenarios and to accommodate innovative new hardware and software technology. With the latest release of the SAP Business Information Warehouse (BW) benchmark 3.5, the query phase is more oriented towards actual data warehouse queries, but still remains comparable to past SAP BW benchmark implementations².

The layout of the SAP BW benchmark uses a 3-tier benchmark configuration on the SAP NetWeaver 2004 platform. SAP uses one driver for all their different benchmark scenarios and versions. The SAP application server benchmark driver represents the presentation layer and is often called the front-end, or first, tier. This driver has the job of simulating end users by driving business transactions in dialog steps. The next layer is the application layer, or so-called second tier, and holds the SAP logic. The last layer is the database layer (or third tier). From the logical point of view, all layers reside on their own hardware but these layers are often placed together on the same physical system. In this benchmark configuration, an SAP dialog instance and a DB2 partition reside on each of the 32 physical nodes. Connectivity between each node is facilitated by onboard Gigabit Ethernet and Voltaire InfiniBand. This configuration can be seen in **Figure 2**. (Physical details of the system are further outlined in **Section 4** of this document.)

² For correctness reasons, it should be pointed out that SAP is only allowing actual comparisons of equal benchmark releases in very tight comparable factors.

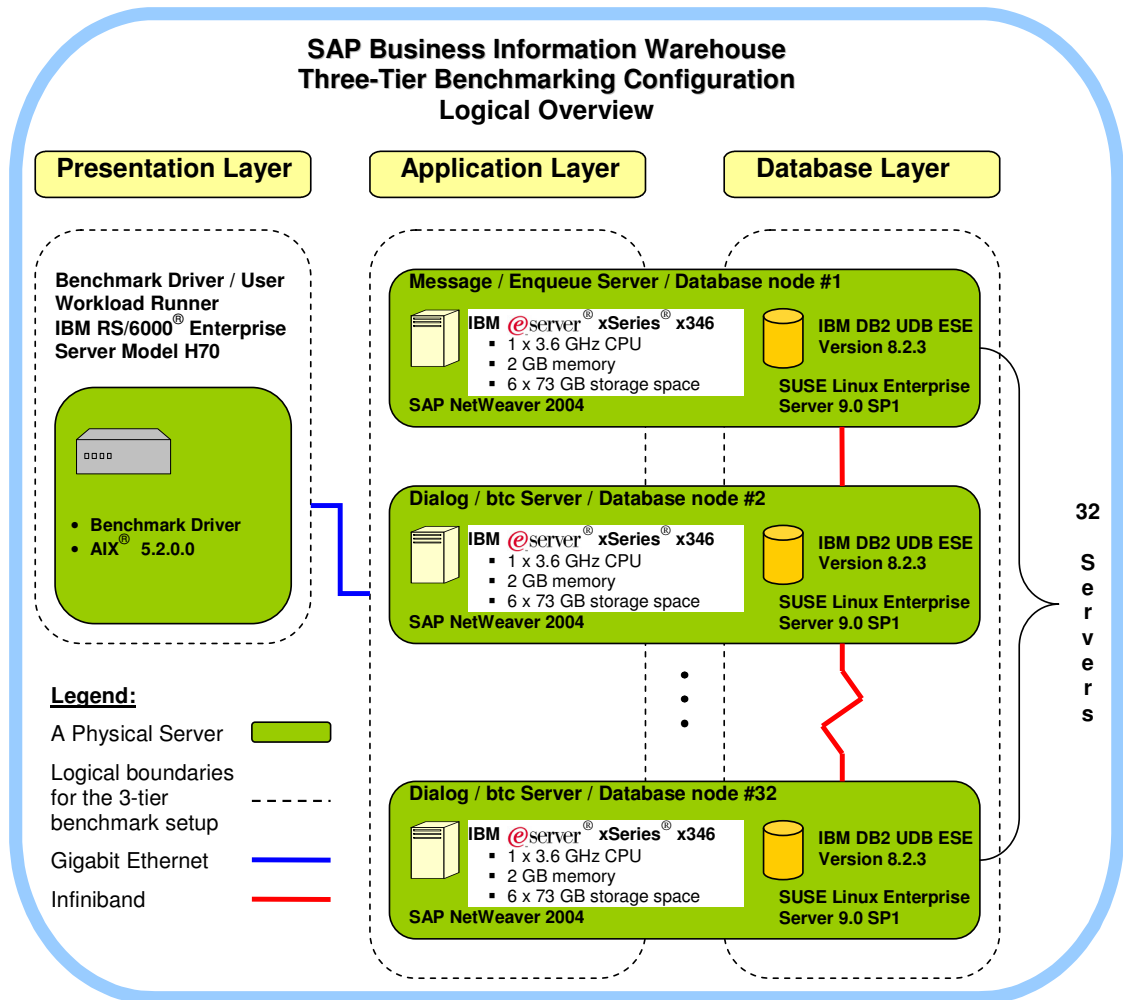


Figure 2 : SAP BW 3-tier Benchmark Configuration

The execution of the SAP BW Benchmark³ follows a set of steps outlined in the SAP BW benchmark kit⁴ and tests two phases. The first phase simulates the steps needed to load data into the InfoCube OSD_C01 and the Operational Data Store (ODS), ODSBENCH. The second scenario simulates multi-user queries.

During the load phase, the amount of data that should be loaded into the benchmark InfoCube is dependent on the amount of memory and disk space available. The implemented benchmark configuration used 32 servers, which provided a total of 64 GB of available main memory correlating to almost 1 billion records in the fact table. The master data sizes, which are constant for all benchmarks, are listed in **Figure 3**.

³ For more detailed information on the SAP BW benchmark, visit <http://www.sap.com/benchmark>

⁴ Contact your SAP sales office or www.sap.com/benchmark, to learn more about the SAP BW benchmark.

- ▶ 100000 products
- ▶ 100000 customers
- ▶ 9 divisions
- ▶ 9 distribution channels
- ▶ x incoming orders / day in total 10 *365 days = 3650
- ▶ 100 version
- ▶ 1000 sales orgs
- ▶ 22 countries
- ▶ 100 types

Figure 3 : Master data

For SAP BW benchmarks, the data load is run without having any secondary indices created on the fact tables. As such, they are recreated after the data load, and statistics are gathered using runstats for DB2 to use when creating query access plans. At this point, there are 10 aggregates that need to be defined on the OSD_C01 InfoCube. Aggregates are subsets of InfoCube data that is pre-summarized so that data can be read directly from the database during query execution instead of having to perform this summarization during runtime. This technique allows us to improve query performance by reducing the volume of data to be read from the database, speed up query execution time, and reduce the overall load on the database. For real-world implementations, the SAP BW system will generate suggestions for creating optimal aggregates, and then it is up to the system administrator to decide if these aggregates are created or not. The aggregate definitions are shown in **Figure 4**.

Division: F, B1 Material number: * Country of the sold-to party: * Valuetype for reporting: *	Distr. Channel: * Division: * Material number: H/BW_SDK_BMK///0MATERIAL 2 Country of the sold-to party: * Valuetype for reporting: *
Division: F, B2 Material number: * Country of the sold-to party: * Valuetype for reporting: *	Distr. Channel: * Division: * Material number: H/BW_SDK_BMK///0MATERIAL 3 Country of the sold-to party: * Valuetype for reporting: *
Division: F, B3 Material number: * Country of the sold-to party: * Valuetype for reporting: *	Distr. Channel: * Division: * Material number: H/BW_SDK_BMK///0MATERIAL 4 Country of the sold-to party: * Valuetype for reporting: *
Division: F, B4 Material number: * Country of the sold-to party: * Valuetype for reporting: *	Division: * Country of the sold_to_party: * Valuetype for reporting: *
Division: F, B5 Material number: * Country of the sold-to party: * Valuetype for reporting: *	Division: * Valuetype for reporting: *

Figure 4: Aggregates

The ODS load is similar to the InfoCube load except that no master data needs to be maintained for this step so less processing and fewer selects are required. The final stage for

the load phase is the ODS activation of data. This step verifies that the data loaded is consistent and is made available for query reporting within the Business Explorer (BEX). This step is the most time-consuming step of the load and requires a significant amount of CPU processing per row.

The query phase is the most important part of this benchmark; it generally consumes the most resources in a customer environment. This multi-user phase simulates the actions in a Sales and Distribution (SD) InfoCube, against either of the data stores that were created during the load phase. The benchmark runs through 20 navigational steps, and for each navigational loop there are 11 query steps performed, and the results are measured by the number of dialog steps per hour.

The SAP Benchmark steps have been summarized in **Figure 5**

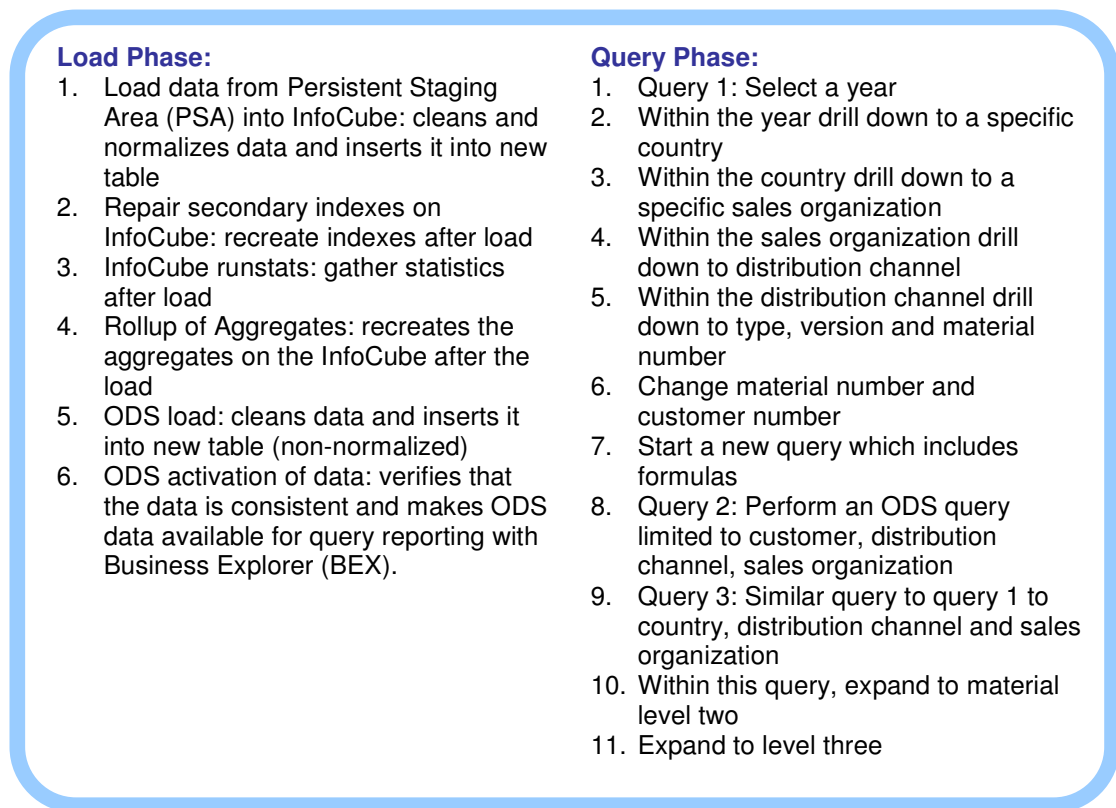


Figure 5 : Benchmark steps

3. The DB2 Integrated Cluster Environment

The introduction of this paper outlined the complexity of data warehouses. The secret behind simplifying the overall solution lies now in the underlying methodology from DB2 UDB for cluster deployments. This methodology is called IBM DB2 Integrated Cluster Environment, or DB2 ICE, for short. DB2 ICE and its components are architected to deliver optimum price/performance and reliability on commodity components. The DB2 ICE solution comes completely integrated: the hardware is assembled, cabled, and labeled; the software is installed; and the operating system is configured with customer-specified TCP/IP addresses. This integration helps to reduce the customer deployment time of SAP BW solutions and improves the time-to-value and return on invested capital for the overall deployment because,

in many cases, up to two months of setup time can be shaved off the project plan. The IBM DB2 balanced configuration unit⁵ (BCU) methodology is the engineering brain behind DB2 ICE and is designed with ease of understanding and deployment in mind.

So what is a BCU? A BCU is a single building block containing the minimum replicable hardware and software requirements necessary to start or expand the infrastructure of a Business Intelligence system. This balanced configuration methodology provides precise guidance as to what ratios of disks, memory, and CPU power are needed for a certain workload scenario. Based on these data points, the IBM hardware and software team jointly engineers a solution stack, not only to fulfill those ratios, but also to allow for a very simplified deployment that contains the knowledge of correct resource allocation for a project – a project such as deploying SAP BW.

Parameter	Specification
SUSE Linux Enterprise Server Edition	Version 9.0 SP 1 64-bit
Kernel	Kernel 2.6.5-7.139-smp
Linux Swap Space	50 GB

Figure 6 : Operating System

The hardware base for DB2 ICE and this benchmark configuration is the IBM eServer™ Cluster 1350. The software components include IBM DB2 UDB Enterprise Server Edition (ESE) and the Data Partitioning Feature (DPF). SUSE Linux Enterprise Server (SLES) and Red Hat Enterprise Linux (RHEL) are the official SAP supported operating systems⁶ and this benchmark uses SLES9 SP1, as seen in **Figure 6**. DB2 ICE lets you choose between Gigabit Ethernet and InfiniBand to connect each physical node, this choice should be based on the size of your configuration which will be explained in Section 4 of this document. InfiniBand is a high-speed interconnect used to connect servers with remote storage, networking devices, and other servers. **Figure 7** outlines the components of this product offering.

⁵ For more information on BCUs, visit <http://www.ibm.com/software/data/DB2bi/bcu>.

⁶ Visit <http://www.ibm.com/DB2/linux/validate> to view supported operating systems for DB2. Visit <http://www.sap.com/linux/platforms> for supported operating system releases for SAP NetWeaver 2004.

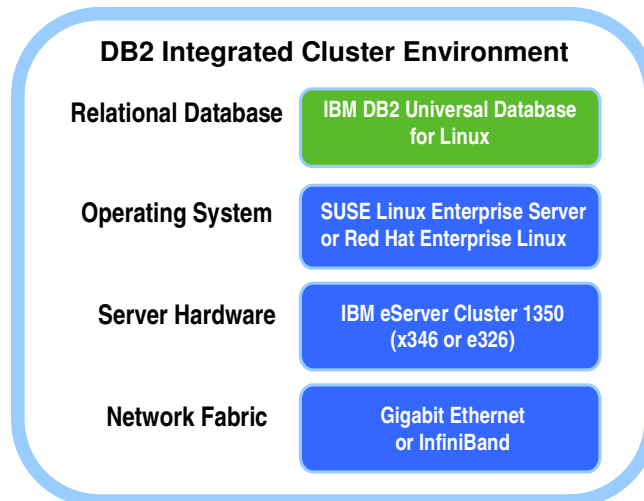


Figure 7 : DB2 Integrated Cluster Environment

4. Physical Setup of the Benchmark Environment

The SAP BW benchmark system was configured to use entry-level hardware to show the ability of DB2 ICE to provide a robust Business Warehouse infrastructure, even with the most minimal of hardware configurations. The components for this benchmark were taken from a previous existing 64-node cluster as seen in **Figure 8**. To demonstrate the capabilities of DB2 ICE, only half the number of nodes (32 nodes) and internal hard drives were used.



Figure 8 : Benchmark rack configuration

The balanced configuration unit (BCU) that this benchmark was based on comprises IBM xSeries® x346⁷ servers, which are designed to deliver mission-critical performance and reliability for data-dense environments. Each x346 server, as seen in **Figure 9**, comes with one 3.6 GHz Intel® Xeon® processor with Extended Memory 64 Technology (EM64T), 2 GB of RAM, and six 73 GB hard drives. The new Intel EM64T Xeon processor enables 64-bit computing and improves performance by allowing the system to address more than 4 GB of both virtual and physical memory. In addition, the whole Intel Xeon processor family comes with Hyper-Threading (HT) technology, which increases the servers' computing power and throughput by providing thread-level parallelism on each processor, resulting in more efficient use of processor resources, higher processing throughput, and improved performance for today's multi-threaded software.

⁷ For more information, visit <http://www-03.ibm.com/servers/eserver/xseries/x346/specs.html>


IBM  x346	
Processors	1 x 3.6GHz Intel Xeon with EM64T technology
Memory	4 x 512MB PC3200 CL2.5 ECC DDR SDRAM RDIMM
Disk Controllers	1 x IBM ServeRaid™ 7M
Networks Controllers	1x 1Gbps onboard Ethernet, Voltaire HCA
Disk Drives	6 x 73GB 15Krpm Ultra320 SCSI HS HDD

Figure 9 : IBM eServer x346

As mentioned earlier, the size of the SAP BW benchmark was based on the amount of available main memory and disk space. To demonstrate the minimum disk capacity that DB2 ICE can provide with maximum redundancy configured, 73 GB hard drives were chosen. Each server used two RAID 5 arrays so as to ensure high data reliability and transfer capacity, resulting in four usable disks and two disks for data parity. The total amount of available disk space per server was approximately 55 GB. Each server rack can accommodate 16 servers, so two racks, as seen in **Figure 8**, were used for a total of 32 machines. With this configuration, the benchmark can provide almost 2 TB of raw data.

For this benchmark configuration, on-board Gigabit Ethernet was used for connectivity between application servers and the end user. The Voltaire 10 Gbps InfiniBand⁸ was used for connectivity between each physical node in the Linux cluster. InfiniBand components include Voltaire Host Channel Adapters (HCA) for each server and two Voltaire ISR 9024 switches creating the InfiniBand fabric (see **Figure 10**). The Voltaire ISR 9024 switch comes with twenty-four 10 Gbps ports in a 1U chassis and the capacity to build InfiniBand fabrics that scale from several to hundreds of nodes.

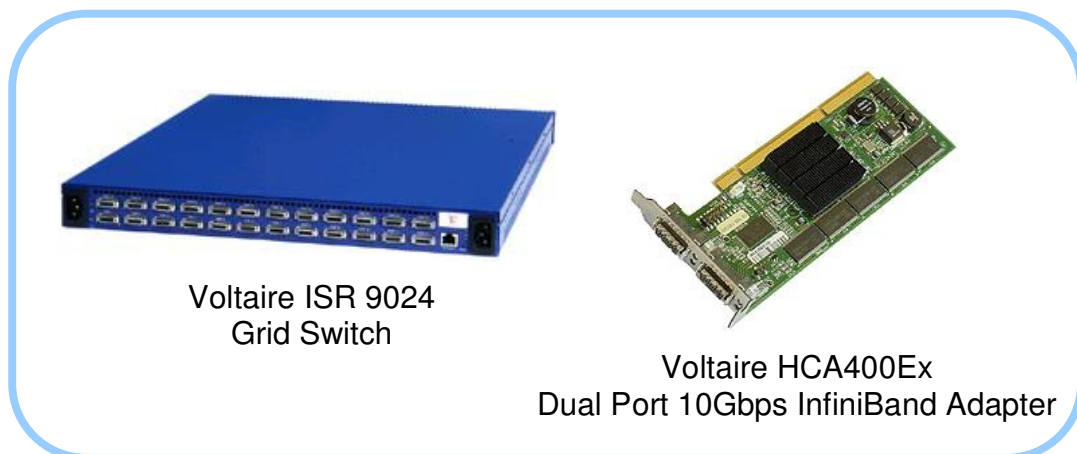


Figure 10 : Voltaire grid backbone products used in the benchmark

For very small clusters, Gigabit Ethernet may be a good-enough solution as a cluster interconnect; however, bandwidth requirements increase when the configuration grows larger.

⁸ For more information on Voltaire InfiniBand products, visit <http://www.voltaire.com>

So for larger clusters, such as 32 nodes or more, Gigabit Ethernet becomes a bottleneck on the performance of the cluster.

In contrast to the Ethernet environment, an InfiniBand fabric allows 10 Gbps per each adapter port (and the common adapter comes with two of those for redundancy). The additional cost of the InfiniBand adapter is quickly offset by the extra performance and scalability that is achieved when using InfiniBand. Voltaire InfiniBand Grid Interconnect switches are also simple to manage and provide efficient management of the entire cluster. The Voltaire switches do not need additional configuration or special certified staff.

From a system administration perspective, the InfiniBand fabric at the network level represents a standard cluster of TCP/IP devices interacting with standard network management tools. However, when it comes to database communication, the performance provided by the InfiniBand interconnects low latency and fast response times make it clear that this is no ordinary network.

Figure 11 shows a 64-node cluster utilizing the Voltaire InfiniBand Grid Interconnect infrastructure. The five ISR 9024 switches connect the servers into a single 10 Gbps fabric through an ISR 6000 Switch Router that provides the gateway to the corporate Ethernet environment and potential application servers. With this benchmark, only three of the ISR 9024 switches were utilized since only half of the servers (32) of the original configuration participated in the benchmark.

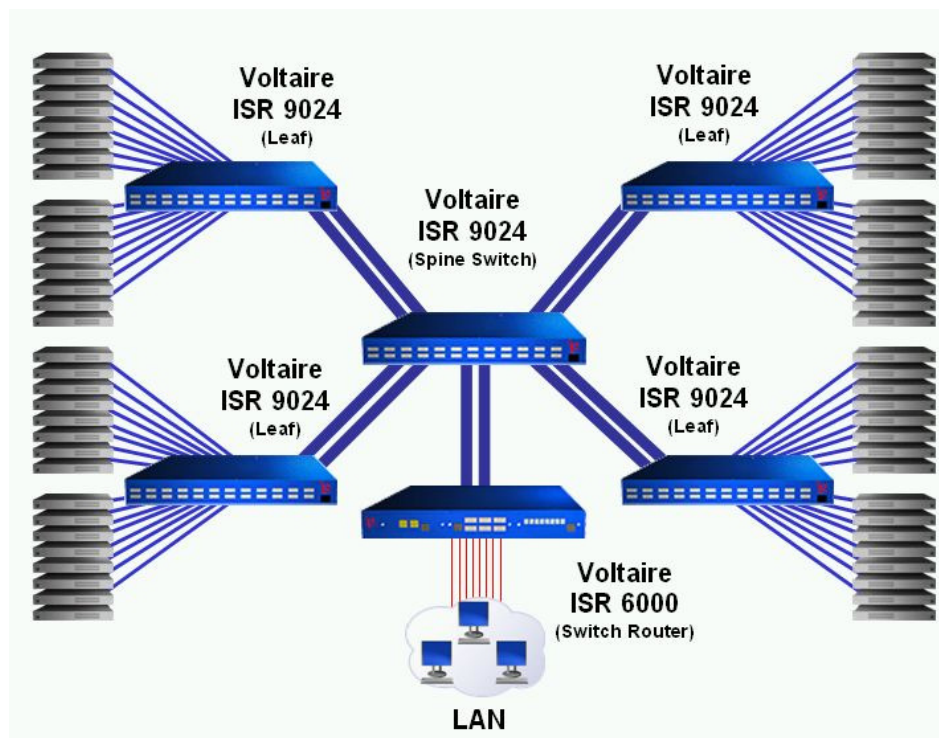


Figure 11 : Voltaire InfiniBand Grid Interconnect Infrastructure

In addition to the interconnect functionality of the cluster, the 10 Gbps fabric allows convenient fast backups that can be directed to an additional backup server. It reduces the costs for the backup infrastructure and can shorten the backup windows. While a usual backup using fibre connection to a backup system is limited again to 2 Gbps (assuming single path fibre connections), the InfiniBand-based solution features a five times higher bandwidth. In this case, the file-based backup can be made to an additional backup storage system that can be a simple server with S-ATA disks on the database InfiniBand fabric, and the final tape backup

is then executed asynchronously from there, saving the expensive fibre channel ports at each server and simplifying the backup process.

5. The DB2® database advantage

When an SAP BW system grows in size, number of users, and query execution time, it may result in performance bottlenecks. SAP by design is scalable at the presentation layer and application layer. With DB2 UDB V8.2, SAP BW is also scalable at the database layer

DB2 UDB V8.2 and its Data Partitioning Feature (DPF) enable us to scale a single database over more than one server. With DPF, a large database can be distributed over multiple partitions (logical and physical) using a shared-nothing architecture. Shared-nothing means that every database node in a cluster has its own dedicated CPU and hard disk resources, and each node is responsible for its own subset of the data, determined when you deploy the cluster.

As data volumes grow, new logical or physical partitions can be added into an existing environment and the response time for queries remains virtually constant. DB2 UDB with DPF is an excellent choice for implementing business intelligence solutions because it allows you to 'pay-as-you-grow', enabling a business to control its cost in a predictable fashion. DB2 UDB for Linux can scale up to 1000 nodes.

DB2 UDB also has many features⁹ to help improve SAP BW systems performance. To help improve I/O, a 32 KB page size was used for the FACTD, FACTI, ODSO, and ODSI table spaces. Business intelligence workloads can benefit from a large page size because more data can be stored on a page that is frequently scanned or sorted.

Multidimensional clustering (MDC) and materialized query tables (MQT) are used to improve query performance. MDC provides a method for clustering data in tables along multiple dimensions in a flexible, continuous, and automatic way. Maintenance overhead is greatly reduced because there is no need for data maintenance operations such as reorganization and index maintenance during inserts, updates, and delete operations. An MQT is a table containing data in the form of pre-computed results derived from one or more tables that define the MQT.

For this benchmark configuration, MDC tables were created on the Persistent Staging Area¹⁰ (PSA), and fact tables and eleven MQTs were created for each benchmark query.

6. Benchmark Results

Previously published SAP BW benchmarks have only implemented the 2-tier Internet configuration with anywhere from 116,800,000 rows (proportional to having 8 GB of main memory) to 233,600,000 rows (proportional to having 16 GB of main memory) in their fact table¹¹. This particular SAP BW benchmark publication is currently the only 3-tier Internet configuration with 64 GB of main memory providing the capacity to have up to four times more data with almost 1 billion rows in the fact table.

⁹ For more DB2 features and further details on features mentioned in this paper, visit <ftp://ftp.software.ibm.com/software/data/DB2/linux/DB2stingerlinux.pdf>

¹⁰ In the SAP BW information model, data is physically stored in PSA objects. These are collections of flat tables holding extracted data that has not yet been cleaned or transformed.

¹¹ Visit the SAP benchmark Web site: http://www.sap.com/solutions/benchmark/bw2_results.htm to view the details of the published results of 2-tier configurations.

Though we cannot directly compare a 2-tier configuration with that of a 3-tier configuration, we can still evaluate the results of this benchmark in terms of the amount of memory used and the number of rows that were loaded into the fact table. The results of 2-tier benchmark configurations on similar Intel hardware and distributed platforms, with memory ranging from 8 GB to 16 GB, are anywhere from 24,108,018¹² to 66,918,119¹³ rows per hour for load phase 1, and 66,852¹⁴ to 154,656¹⁵ dialog steps per hour for the query phase. **Figure 12** shows the results of this 3-tier benchmark. For a bare-bones DB2 ICE configuration, load phase part 1 results were 168.360.360 rows per hour, and the query phase resulted in 311,004 dialog steps per hour.

Phase	Measure
Load 1	168,360,360 rows/hour
Load 2	6,511,120 rows/hour
Query	311,004 dialog steps/hour

Figure 12 : SAP BW Benchmark results

With this benchmark, DB2 ICE remains at the head of the class by leading the way on 3-tier benchmark configurations and proving that it provides superior performance and data capabilities in comparison to other infrastructures.

DB2 ICE, backed by the power and scalability of Linux, brings together the joint effort of IBM DB2 software, IBM xSeries servers, Intel processor technology, and Voltaire's industry-leading interconnect solutions to provide a robust infrastructure for SAP Business Information Warehouse.

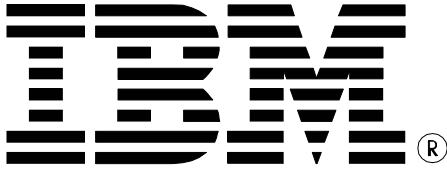
¹² Refer to SAP BW benchmark certification number [2003043](#)

¹³ Refer to SAP BW benchmark certification number [2004026](#)

¹⁴ Refer to SAP BW benchmark certification number [2003043](#)

¹⁵ Refer to SAP BW benchmark certification number [2004026](#)

IBM DB2 ICE leads the way with 3-tier, 32-node SAP BW benchmark results



© Copyright IBM Corporation 2005
All Rights Reserved.

IBM Canada
8200 Warden Avenue
Markham, ON
L6G 1C7
Canada

Printed in United States of America
10/05

IBM, IBM (logo), AIX, DB2, DB2 Universal Database, eServer, e(logo)server, RS/6000, ServeRAID, and xSeries are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both.

Intel and Xeon are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

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

Other company, product, or service names may be trademarks or service marks of others.

References in this publication to IBM products or services do not imply that IBM intends to make them available in all countries in which IBM operates. The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law:

INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you. This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurement may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of 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.

The information in this white paper is provided AS IS without warranty. Such information was obtained from publicly available sources, is current as of 12/05/2005, and is subject to change. Any performance data included in the paper was obtained in the specific operating environment and is provided as an illustration. Performance in other operating environments may vary. More specific information about the capabilities of products described should be obtained from the suppliers of those products.