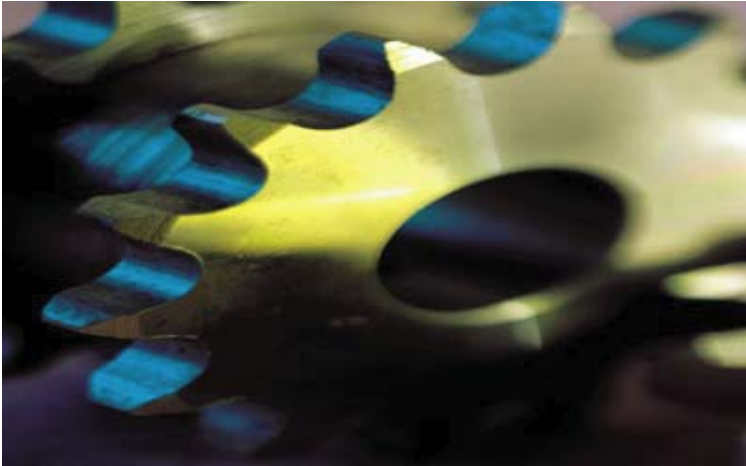




SAS and IBM - Achieving Systems Excellence Together



Competing in the on demand world requires complex yet responsive systems that can grow and change with a business. This requires the right balance between simplicity and flexibility that results from cooperative alliances between industry leaders like IBM – the world’s largest IT hardware and services company – and SAS, a leader in Business Intelligence (BI) solutions.

Recently SAS and IBM proved the power of their alliance by performing a SAS ETL test that leveraged the power of IBM’s technology. Extract, transform and load, or ETL, is a key database function for helping to transform data into information.

SAS: The Power to Know®
Clients trust SAS to help them drive profitable growth, unlock the value of their existing business data, and help ensure the security and integrity of their global data warehouse.

SAS provides complete BI solutions. IBM provides industry leading technologies and products that help maximize the

use of SAS® software systems in your business today.

IBM and SAS Leverage IBM ~® p5 - UNIX® Servers
Flexibility, performance, reliability and security top the list of what’s required from today’s computing environments.

To gain competitive advantage, an organization must be able to respond to changes in business demands - anticipated or not - while reducing IT costs and ensuring that data remains protected. IBM ~ pSeries systems deliver on every count.

IBM is changing the way the world buys UNIX servers. The performance of the 64-bit IBM POWER5™ platform, enabled by simultaneous multithreading and AIX 5L™, the industrial-strength UNIX operating system from IBM, has is setting new industry performance standards. IBM’s continued investment in the AIX 5L operating system and unique advanced virtualization technologies like the IBM Micro-Partitioning™ technology enable users to accomplish more with a single system.

Mainframe-inspired reliability features designed into every system, along with embedded autonomic security features in the AIX 5L platform and the POWER5 processor, help keep the system up and running around the clock. pSeries servers are designed to provide unparalleled UNIX performance and an attractive price/performance ratio with POWER5 technology.

The IBM ~ pSeries is designed for agility to enable quick response to changes in UNIX processing requirements, so anticipated and unanticipated business demands can be met. pSeries servers provide the ability to dramatically increase system utilization in a consolidated server environment.

IBM ~ pSeries servers are designed to provide a highly secure UNIX processing platform that helps safeguard data from unauthorized access. pSeries systems also make migrations and upgrades easy, so anyone, including existing IBM clients and users of competing systems, can take advantage of these game-changing capabilities designed to provide excellent overall value today—and tomorrow.

The IBM TotalStorage® Difference

IBM provides organizations what they need to make the right storage decisions now and in the future.

IBM provides a focus on open standards, support for heterogeneous environments and continuous leverage of innovative systems technologies.

This storage system approach provides the following key benefits:

- TotalStorage solutions that are designed to provide a lower total cost of ownership.
- IBM TotalStorage SAN Volume Controller and TotalStorage SAN File System (SFS) software enable the virtualized management of heterogeneous storage resources.

IBM TotalStorage solutions can help provide industry-leading business continuity capabilities, to help support critical business decision making.

IBM and SAS – Partnering for SAS ETL Performance

The IBM and SAS engineering teams have worked together to help ensure SAS software works seamlessly on IBM infrastructure and to improve SAS software performance on IBM pSeries servers and storage products.

Recently IBM and SAS achieved outstanding performance running the ETL test scenario based upon a business scenario utilizing SAS® ETL Studio and SAS@9 Platform to help demonstrate the SAS ETL Server Solution. This Reference Architecture was developed by SAS as a means to compare performance between various hardware/operating system platforms.

The test scenario builds a warehouse and features two workflows that were predominately built using SAS ETL Studio. The first half of the test builds out a data warehouse composed of two star schemas from 50 raw text files (census data files). These files are converted into two fact tables with associated dimensions. One

fact table is for census personal information and the other is for household information. All dimension tables are related back to the fact table by a generated surrogate key. The surrogate keys are built from multiple business values found in each dimension table. The final output is nine dimension tables and two fact tables.

In a common data warehouse model, it is typical to update or add 10-20% of the total warehouse records during the nightly ETL process. The second half of the test demonstrates Slowly Changing Dimensions during the update phase of the ETL process. As updates arrive, older records are marked with the date when they became invalid. Indexes are created on fact tables upon completion of the initial load and/or update process.

Test Suite Execution Highlights

The test scenario used was the SAS Extra Large Reference Architecture suite. It involved the loading and validation of three jobs (49.5GB each). Three Large scenarios run simultaneously from independent data sources. This involved an update/upload of three jobs (10.8 GB each).

System Configuration:

- 8-way 1.65 GHz IBM ~ p5-590 server partition
- AIX 5L Version 5.3
- 128GB RAM
- TotalStorage DS4400 (56 15K rpm disks, two controllers)
- SAS ETL Server (Base SAS® 9.1.3

These outstanding test results confirm the superior performance advantages of IBM pSeries systems when used to run critical business intelligence applications in real business situations. These Reference Architecture tests offer clients a look into the performance potential of IBM systems running SAS ETL applications under the most demanding of circumstances.

First Half of Test: Initial Warehouse Build Out	
Data Validation and Load	1,097 Million rows/hour
	258.6 GB / Hour
Data Transformation:	293 Million Rows / Hour
	72.3 GB / Hour
Second Half of Test: Update Load – 21% of Initial Data Warehouse	
Data Validation and Load	1,096 Million Rows / Hour
	265 GB / Hour
Data Transformation:	176 Million Rows / Hour
	45.7 GB / Hour

For more information

To learn more about IBM ~
pSeries systems and offerings for
SAS, please contact your IBM
sales representative, or visit:
<http://www.ibm.com/servers/eserver/pseries>

Please note: SAS ETL widely varies
between implementations and will be
highly customized due to business
requirements. This particular test may
be more or less complex than a
particular SAS ETL requirement. It is
not meant to be a benchmark of
absolute performance. These results are
shared as a guideline of what a typical
user might build with SAS ETL Server.

**The following terms are registered
trademarks of International Business
Machines Corporation in the United
States and/or other countries:** AIX,
AIX/L, AIX/L(logo), e business(logo),
e(logo)business, e(logo)server, IBM
TotalStorage Proven, pSeries,
TotalStorage,

**The following terms are trademarks of
International Business Machines
Corporation in the United States
and/or other countries:** Advanced
Micro-Partitioning, AIX/L(logo), AIX
5L, e-business(logo), e-business on
demand, eServer, Virtualization Engine,
Micro-Partitioning, POWER, **A full list
of U.S. trademarks owned by IBM may
be found at:**
<http://www.ibm.com/legal/copytrade.shtml>.

UNIX is a registered trademark in the
United States, other countries or both.

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

SAS and all other SAS Institute Inc.
product or service names are registered
trademarks or trademarks of SAS Institute
Inc. in the USA and other countries.

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