



# Deploying SAS Enterprise Business Intelligence in an AIX virtual environment

*Guide for installing in a medium-scale  
POWER6 environment*

*Alfredo Mendoza  
Joe Pu*

*ISV Business Strategy and Enablement  
January 2009*



## Table of contents

<b>Abstract</b> .....	<b>1</b>
<b>Introduction</b> .....	<b>1</b>
<b>Describing the test suite</b> .....	<b>1</b>
<b>Deploying the SAS 9 eBI midsize example</b> .....	<b>1</b>
Deploying the hardware .....	2
Deploying hardware by using PowerVM.....	5
Deploying the software .....	7
<b>Benchmark Results</b> .....	<b>8</b>
<b>Summary</b> .....	<b>10</b>
<b>Resources</b> .....	<b>11</b>
<b>About the authors</b> .....	<b>12</b>
Acknowledgements.....	12
<b>Trademarks and special notices</b> .....	<b>13</b>



## Abstract

---

*The SAS Enterprise Business Intelligence (BI) Server solution, running on the IBM System p 570 server with the IBM AIX Version 6.1 operating system, is a perfect match for midsized enterprises that are looking for performance, scalability and reliability. Using IBM PowerVM technology, this white paper shows a reference architecture that accommodates the components of the SAS Enterprise BI Server solution in a single-server platform.*

## Introduction

---

Midsized enterprises, just as with large enterprises, need speed in decision-making abilities to meet today's challenges of a global economy. SAS Enterprise Business Intelligence (BI) Server provides a suite of business intelligence (BI) tools, as well as an integrated BI platform. A suite of user interfaces that provide dashboard, scorecard and other visualization tools offers users many expanded capabilities beyond the traditional query, reporting and OLAP functions.

Unmatched IBM expertise in hardware and software technology enables the SAS Enterprise BI solution for IBM® AIX® 6.1 on IBM POWER6™ to deliver significant benefits. It can be deployed on an infrastructure that is designed to improve reliability, performance and scalability. Combined with IBM PowerVM and System Storage Proven™ solutions, the reference architecture described in this paper provides unparalleled function and performance for implementing BI solutions in a midsized environment.

## Describing the test suite

---

The reference architecture was tested against the SAS Enterprise BI Server reference architecture test suite. The test suite is based on scenarios that show how SAS users address their business- and data-reporting needs. Although the data that is used in this scenario is retail-based, the workload is representative of reporting functions that are required across a wide range of industries.

- 10 SAS Enterprise Guide sessions developing reports
- 20 advanced analytics sessions performing SAS statistical functions
- 10 SAS Enterprise Guide sessions developing reports and SAS Data Integration Server code
- 80 SAS Add-In for Microsoft® Office sessions
- 80 SAS Web Report Studio ad hoc, single-user sessions
- 70 SAS Web Report Studio static report sessions
- 80 SAS Stored Process user sessions
- 80 SAS OLAP cube-viewer single-user sessions
- 70 SAS Enterprise BI Server portal and dashboard viewer sessions

Each type of user is patterned after users that are commonly encountered in SAS Enterprise BI Server deployments.

## Deploying the SAS 9 eBI midsize example

---

The SAS environment can be loosely defined as a three-tier architecture, comprising a Web and application tier, a services middle tier and a computational back-end tier. Typically, this type of architecture is deployed in a distributed environment. However, with IBM PowerVM™ virtualization technologies, the alternative approach of logical server partitioning is recommended.

## Deploying the hardware

To support 500 concurrent users at 77 percent processor utilization, the SAS Enterprise BI solutions was deployed as four physical partitions of varying sizes across a single IBM Power™ 570 POWER6 processor-based system. Figure 1 shows the basic deployment of the hardware environment.

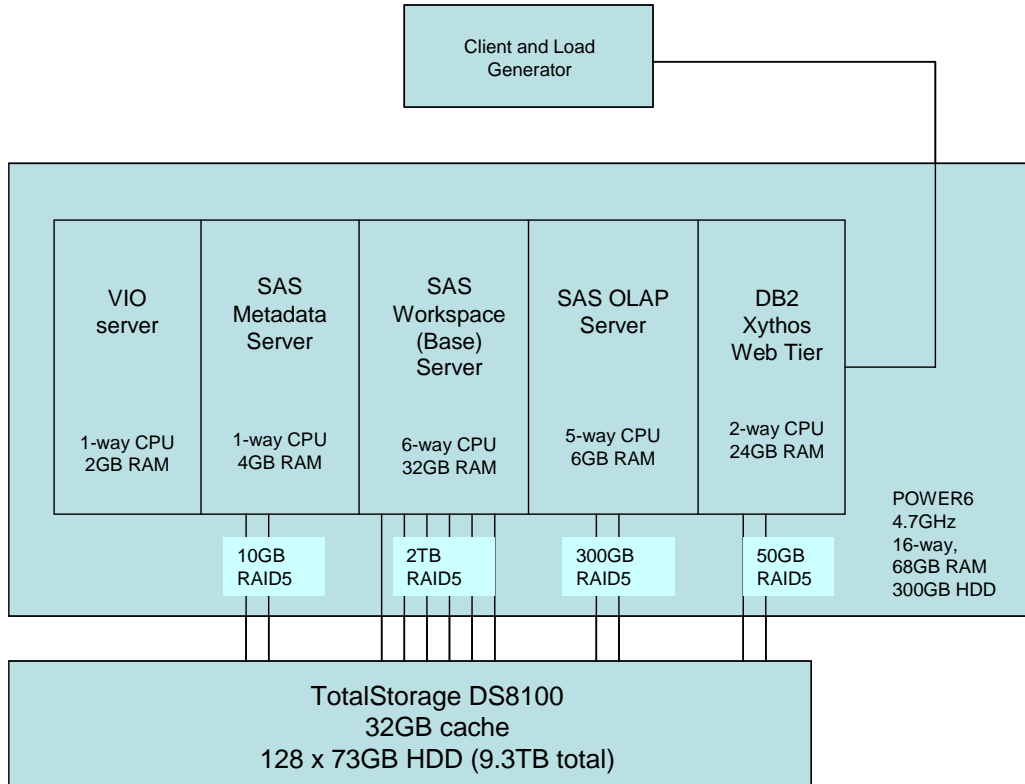


Figure 1. Dedicated-processor logical partition (LPAR) server configuration

The configuration shows the use of a virtual I/O server (VIOS). The VIOS shared serially attached SCSI cards between LPARs. Internally attached disk drives accommodated each partition's system file systems (rootvg). This was done to show general VIOS use and the small amount of processor resource that was needed by the VIOS during the benchmark. Although the use of VIOS for flexibility is beyond the scope of this white paper, users are encouraged to investigate the use of VIOS if they plan to use more advanced features of PowerVM, such as live partition mobility (LPM) in the future (at the IBM Redbooks® Web site, see *PowerVM Live Partition Mobility on IBM System p*, [www.redbooks.ibm.com/abstracts/sg247460.html](http://www.redbooks.ibm.com/abstracts/sg247460.html)).

The Web tier was comprised of a partition that was configured with two POWER6 processors, each with 24 GB RAM. The Web tier LPAR was provided with a 1 Gbps Ethernet connection to an Ethernet switch. The Ethernet switch delivered network connectivity to a LAN that was comprised of an IBM xSeries® Intel® system to provide the load generation of the workload. The Web tier LPAR also accommodated both the Xythos WebFile server and the IBM DB2® instance.

The SAS Metadata Server was deployed as a separate partition that was dedicated to metadata functions. This server had two 1 Gbps fibre-channel connections to the IBM TotalStorage® DS8100 device. The Metadata log files were stored to VIOS-attached disk.



Third-tier computational processing was deployed on two partitions. The SAS 9.1.3 base functions (Workspace engine, and other tasks) are deployed in a dedicated seven-way partition with 32 GB RAM. The SAS OLAP cube is deployed in a separate five-way partition with 6 GB RAM. Drives that were VIOS-attached to both partitions contained the SAS binary files and various log files. The /saswork temporary file system and /data file system were allocated to logical volume groups deployed on the shared DS8100 storage device. The SAS Workspace engine and OLAP cube LPAR were deployed with five and two 1 Gbps fiber-channel paths, respectively, to the storage array.

Network connectivity between SAS partitions (Web tier, SAS Workspace Server, OLAP cube and metadata server) was provided through inter-LPAR communications by using virtual Ethernet adapters that were connected to an IEEE 802.1q (VLAN) style virtual Ethernet switch. The underlying architecture that supported the VLAN technology, known as the IBM POWER Hypervisor™, transmits packets by copying the packet directly from the memory of the sender partition to the receive buffers of the receiver partition without any intermediate buffering of the packet. This technology allows processor cycles that would have otherwise been used for network communications, to be used for application operations.

Using an LPAR configuration allows system administrators to accommodate varying workloads by easily moving processors between partitions. This can be useful if, at certain parts of the workday, the OLAP server needs more processor capacity than the SAS Workspace Server. In this case, the system administrator can manually move one or more processors from the SAS Workspace Server to the OLAP server.

Table 1 shows the storage allocation for each LPAR. The storage was provided by the DS8100 storage solution.

Server type	Server name	Allocated disk space
Metadata server	ibmaix3	10 GB
OLAP server	ibmaix4	300 GB
Workspace Server	ibmaix5	2 TB
Midtier server	ibmaix6	50 GB

Table 1. Storage allocation for each LPAR

The IBM DS8100 storage system consisted of 128 physical disks. Best practice dictates that SAS runtime data be spread out to as many physical disks as possible. In this test deployment, the Workspace Server partition needed the most disk space<sup>1</sup>. The storage configuration for the Workspace Server was used to illustrate some main points in managing disk space that is provided by the DS8100 storage system.

To strike a balance between disk I/O performance and the economics, the storage system was configured as RAID5. The minimal unit on the system that a piece of storage, a logical unit number (LUN), can be created, is called an *array*. Each array consists of eight physical disks. Figure 2 shows a representation of a LUN consisting on an array with 8 disks.

---

<sup>1</sup> Previous SAS experience also supports the observation that the SAS Workspace Server generates a lot of I/O activity.

### 8-disk Array (RAID5)

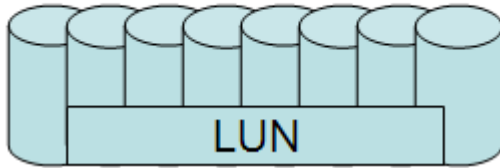


Figure 2. Representation of a LUN that consists of an array with eight disks

With 128 physical disks, the disks were grouped into arrays of eight disks, each resulting in 16 arrays. To simplify AIX disk management, a LUN was created from each array and then presented to AIX (see Figure 3).

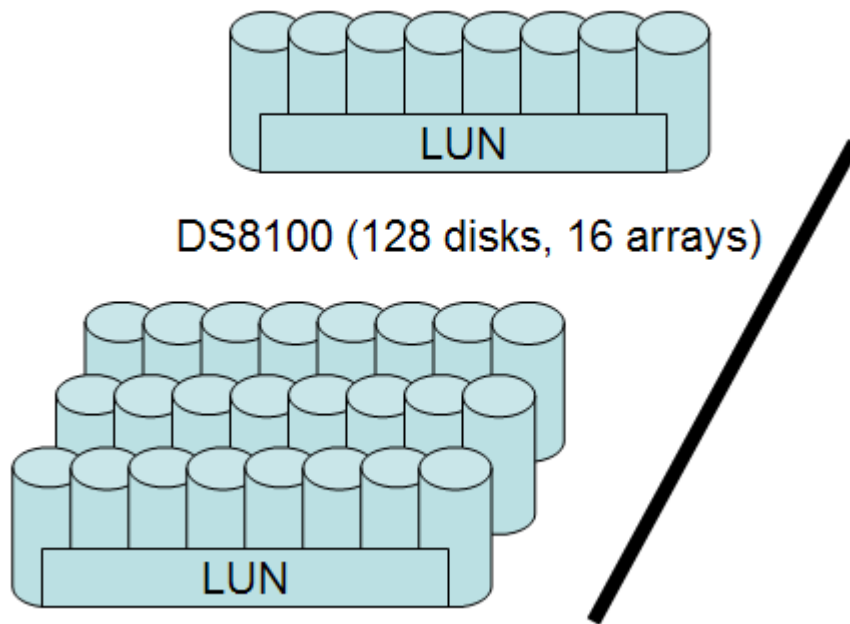


Figure 3. A 128-disk subsystem configured with 16 arrays

From the receiving AIX partition point of view the LUNs were seen as 16 logical hdisks from which a volume group, called *saswrk2TBvg*, was created (Figure 4). Striped logical volumes were added, resulting in a file-system layout that was evenly distributed among all the disks.

```

root@ibmaix5 /=>lspv | grep saswrk
hdisk1      00c08b426fa55922      saswrk2TBvg      active
hdisk2      00c08b426fa55a0d      saswrk2TBvg      active
hdisk3      00c08b426fa55af0      saswrk2TBvg      active
hdisk4      00c08b426fa55bbb      saswrk2TBvg      active
hdisk5      00c08b426fa55c9a      saswrk2TBvg      active
hdisk6      00c08b426fa55ddb      saswrk2TBvg      active
hdisk7      00c08b426fa55f43      saswrk2TBvg      active
hdisk8      00c08b426fa560af      saswrk2TBvg      active
hdisk9      00c08b426fa561b3      saswrk2TBvg      active
hdisk10     00c08b426fa56286      saswrk2TBvg      active
hdisk11     00c08b426fa5633a      saswrk2TBvg      active
hdisk12     00c08b426fa56408      saswrk2TBvg      active
hdisk13     00c08b426fa564d7      saswrk2TBvg      active
hdisk14     00c08b426fa56592      saswrk2TBvg      active
hdisk15     00c08b426fa5664b      saswrk2TBvg      active
hdisk16     00c08b426fa567c2      saswrk2TBvg      active

```

Figure 4. lspv command shows volume group with 16 disks

During testing, system monitoring showed no disk I/O bottlenecks and also showed that disk I/O was equally distributed among the 16 logical disks.

## Deploying hardware by using PowerVM

Another set of tests was conducted to perform a benchmark on the IBM's PowerVM features that use the same test scripts. The IBM's PowerVM features lower the total cost of ownership while providing world-class performance.

Figure 5 shows the shared LPAR configuration with a pool of 15 processors that the LPARs share. The following features were used during testing:

- **IBM PowerVM Micro-partitioning™:** Allows processor resources to be assigned at a granularity of 1/100th of a core.
- **Shared processor pooling:** Allows for automatic non-disruptive balancing of processing power between partitions that were assigned to shared pools which resulted in increased throughput.
- **Virtual I/O server:** Allows resource sharing between LPARs thus reducing costs by eliminating the need for dedicated network and disk adapters,
- **IBM PowerVM features:** Increase resource utilization through dynamic load balancing and resource sharing. Traditional reference architectures look at each resource as a unique and separate entity. PowerVM changes that to have a more holistic view of the environment.

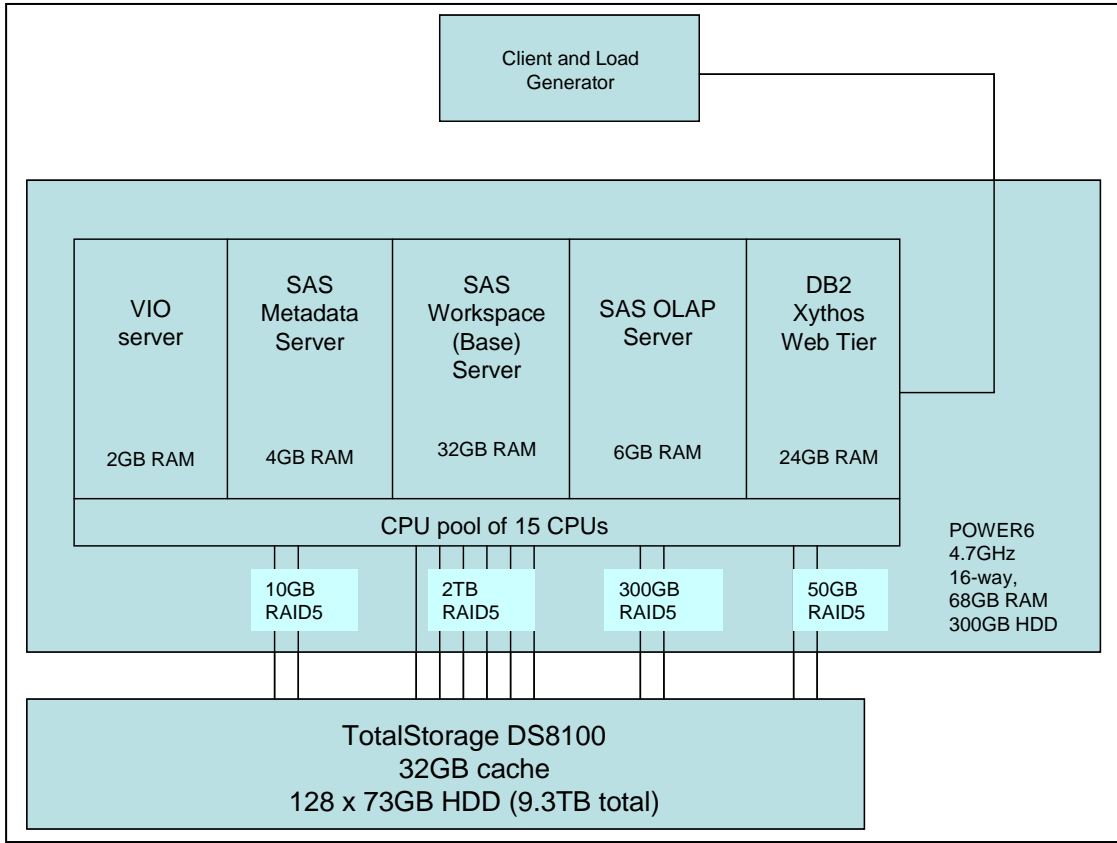


Figure 5 POWER 570 Shared-processor configuration.

The pool of 15 processors were configured to share between the different SAS LPARs. This was done to show the benefits of shared-processor configuration on throughput.

In the shared-processor configuration, processor utilization for each SAS LPAR was measured as follows:

- **Workspace Server:** 5.13 processor\* with 32 GB RAM.
- **SAS OLAP Server:** 4.72 processor\* with 6 GB RAM.
- **SAS Metadata Server:** 0.74 processor\* with 4 GB RAM.
- **Web application server:** 2.24 processor\* with 24 GB RAM.
- **Virtual I/O server** – 0.03 processor\* with 2GB RAM.
- **Storage** – IBM System Storage DS8000 (model DS8100) with 73GB 15K drive sets and 4 GB Fibre-Channel connections.
- **Software** -- SAS 9.1.3 with Service Pack 4; SAS Enterprise BI Server; SAS Scalable Performance Data Engine; SAS OLAP Server; one WebSphere ND 6.1.0.17 cluster with four application server members; and Xythos WebDAV with DB2 9.5.0.1 FP 1 WebDAV data server, all running on AIX 6.1.
- **Network** – 1-gigabit-per-second Ethernet connections. Internal VLAN across virtual partitions.

\* - A shared processor pool (consisting of 15 processors) were virtualized for this configuration. The processor configuration shown in Figure 5 is the average processor usage by each server from this pool. One processor (which was not in the pool) was configured to be used as a staging server.



Table 2 shows individual LPAR configuration for shared resources. The options to set desired, minimum and maximum values are needed when LPARs use the Micro-Partitioning feature of PowerVM.

System	Memory (in GB)			Processor-entitlement			Virtual processor			Weighting
	Online	Minimum	Maximum	Online	Minimum	Maximum	Online	Minimum	Maximum	Uncapped
ibmaix3	4	1	16	1	0.10	4.0	1	1	4	128
ibmaix4	6	1	16	4	0.10	7.0	6	1	7	128
ibmaix5	32	1	36	3	0.10	7.0	6	1	7	128
ibmaix6	24	1	32	2.6	0.10	5.0	4	1	5	128
VIOS	2	2	4	0.40	0.20	2.0	2	1	8	128

Table 2 Shared LPAR configuration table.

## Deploying the software

Each LPAR was installed with identical AIX version 6.1 operating systems. The SAS Enterprise BI Server software (version 9.1.3 SP4) was used with recommended hot fixes from the installation documentation. The software deployment is described by Table 3.

Server	Software	Version	Provided by
Metadata Server	AIX	AIX 6.1	IBM
	SAS Metadata Server	9.1.3	SAS
SAS Workspace (Foundation, IT and others) Server	AIX	AIX 6.1	IBM
	SAS Base and support for Workspace service	9.1.3 SP4	SAS
SAS OLAP Server	AIX	AIX 6.1	IBM
	SAS OLAP and support	9.1.3 SP4	SAS
Web tier, Xyθος and DB2	AIX	AIX 6.1	IBM
	WebSphere ND (32-bit)	6.1.0.17	IBM
	DB2	DB2 version 9.5.0.1 FP1	IBM
	Xyθος WebFile Server	2.2	SAS

Table 3. List of deployed software

IBM WebSphere® Network Deployment version 6.1.0.17 was installed and configured with a four-member cluster configuration. The topology represents a vertical-scaling topology where additional application servers are necessary to accommodate a large number of concurrent SAS Web sessions. This type of configuration provided for high availability in the event that one instance of the application server went down.

## Benchmark Results

For the benchmark results to be certified, several criteria needed to be met. They are as follows:

- *Failed to Passed* transaction ratio must be less than 0.1 percent.
- *Errors to Passed* transaction ratio must be less than 0.1 percent.
- Average transaction response times must be less than five seconds.
- Average transaction response times under load should not be larger than 50 percent of the average transaction-response times for the same transaction with a single-user or a small number of users.

The 100- through 500-user benchmark results that met the benchmark criteria are shown in Figure 6. The results shows linear scalability as numbers of concurrent users increased. Both the dedicated and shared-processor configuration showed the same scalability and average processor utilization. Varying resource utilization for the 500-user test was recorded as follows:

- The total overall average processor utilization was 77 percent.
- The I/O subsystem delivered approximately 6 MB per second per core of sustained I/O to the SAS Enterprise BI Server workspace server. Writes to reads were a ratio of 2-to-1.
- The network traffic approached 0.3 gigabits per second.

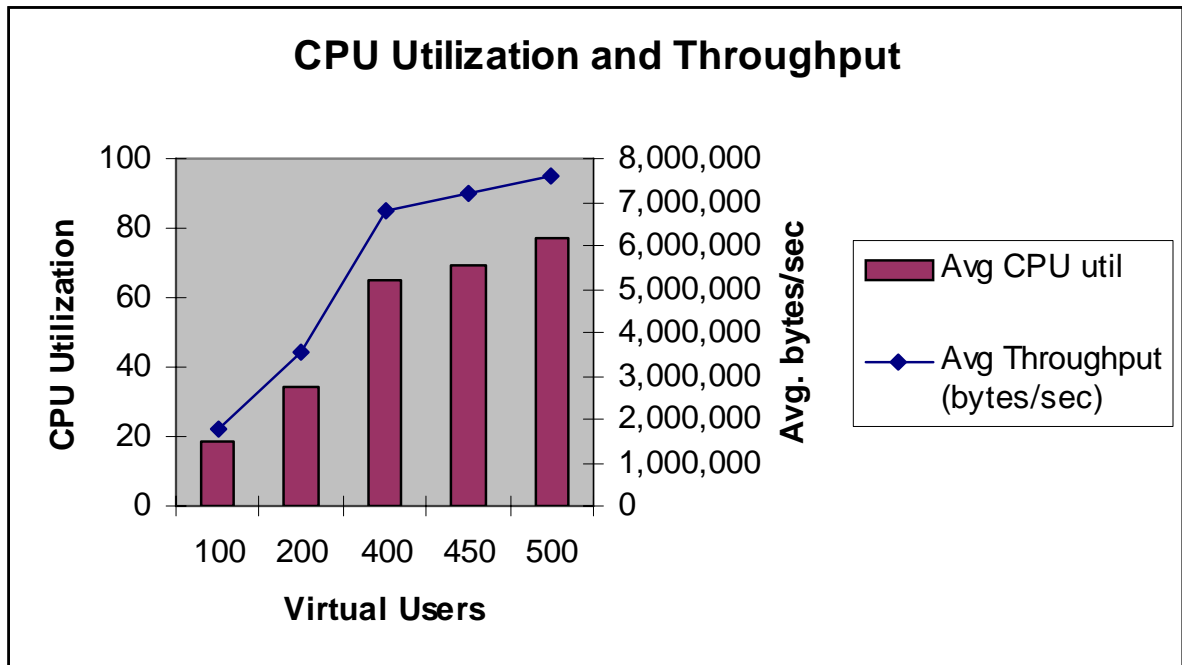


Figure 6 15-way processor utilization for varying number of concurrent users.

Figure 7 shows a comparison of the average throughput while using PowerVM, specifically shared-processor configuration. Figure 7 also shows a compelling reason to use shared-processor resources. Although the average throughput rate was similar in both the dedicated and shared-processor configuration when the numbers of users were below 400, it is possible to see the advantage of using a shared-processor configuration as the number of users increased beyond 400. At 450 and 500 users, there were throughputs of 7,176,251 and 7,605,463 bytes per second, respectively, when using a dedicated-processor configuration. For 450 and 500 users, there were throughputs of 7,485,043 and 8,248,841 bytes per second, respectively, for shared processor configuration. For 500 users, there was a 7.8 percent improvement in throughput.

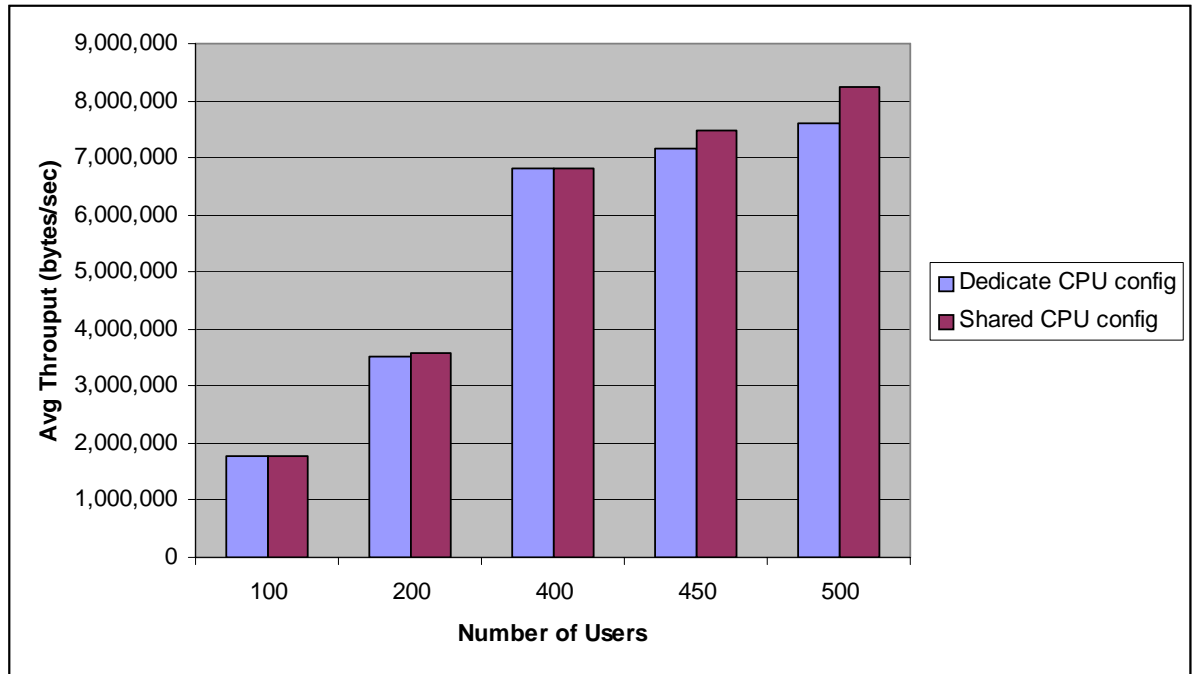


Figure 7. Average throughput comparison between the dedicated- and shared-processor configurations.

Although average transaction-response times remained the same for most of the workload, improvements to the average transaction- response times for some user workloads, such as the OLAP single user view, contributed to the overall improvement in throughput when using a shared-processor configuration.

The results of the performance tests provide an understanding of how the SAS Enterprise BI server solution efficiently uses IBM POWER6 hardware resources. It can also help if you correlate these performance numbers with the anticipated usage profiles of your possible deployments.



## Summary

---

The successful completion of performance testing of the SAS Enterprise Business Intelligence reference architecture showed that an IBM Power 570 system can easily sustain a concurrent, steady-state load of 500 SAS users with varying workloads. The environment was deployed in a consolidated fashion in four partitions on a Power 570, using a pool of 15 physical processors and 68 GB RAM. The environment included the use of a Total Storage DS8100 storage array.



## Resources

---

These Web sites provide useful references to supplement the information contained in this document:

- IBM System p Information Center  
<http://publib.boulder.ibm.com/infocenter/pseries/index.jsp>
- System p on IBM PartnerWorld®  
[ibm.com/partnerworld/systems/p](http://ibm.com/partnerworld/systems/p)
- AIX on IBM PartnerWorld®  
[ibm.com/partnerworld/aix](http://ibm.com/partnerworld/aix)
- IBM Systems on IBM PartnerWorld  
[ibm.com/partnerworld/systems/](http://ibm.com/partnerworld/systems/)
- IBM Publications Center  
[www.elink.ibm.link.ibm.com/public/applications/publications/cgibin/pbi.cgi?CTY=US](http://www.elink.ibm.link.ibm.com/public/applications/publications/cgibin/pbi.cgi?CTY=US)
- IBM Redbooks  
[ibm.com/redbooks](http://ibm.com/redbooks)
- PowerVM Live Partition Mobility on IBM System p, Harding, M., et al. (2007)  
[www.redbooks.ibm.com/abstracts/sg247460.html](http://www.redbooks.ibm.com/abstracts/sg247460.html)
- IBM developerWorks®  
[ibm.com/developerworks](http://ibm.com/developerworks)
- Hayes-Hall, E, et al., The Deployment of a SAS Enterprise Business Intelligence Solution in a large IBM POWER6 Environment, IBM, 2006  
[ibm.com/support/techdocs/atsmastr.nsf/WebIndex/TD103051](http://ibm.com/support/techdocs/atsmastr.nsf/WebIndex/TD103051)
- Jain, V., Virtual Networking on IBM AIX 5L, IBM, 2004  
[ibm.com/systems/resources/systems\\_p\\_os\\_aix\\_whitepapers\\_pdf\\_aix\\_vn.pdf](http://ibm.com/systems/resources/systems_p_os_aix_whitepapers_pdf_aix_vn.pdf)



## About the authors

---

**Afredo (Freddie) Mendoza** is an IT specialist in the ISV Business Strategies and Enablement Group at IBM. He has more than 19 years of experience in application architecture and systems management. His current work involves helping solution providers enable their applications on IBM cross-platform systems. Mr. Mendoza has authored or coauthored several patent-invention publications on thread management and software-metering algorithms that are used in advanced Web services environments. He is the author of *Guide to Utility Computing Strategies and Technologies* and co-author of *The Linux to UNIX Porting Guide*.

**Joseph Pu**'s focus is in the area of AIX performance, tuning and sizing. He has extensive experience in software development, from graphics to software simulation. He started his AIX development experience more than 20 years ago. Joe graduated from the University of Texas at Austin with a degree in Computer Science. He can be reached at [joepu@us.ibm.com](mailto:joepu@us.ibm.com).

## Acknowledgements

Thanks go to the following IBM people who reviewed this paper for technical accuracy:  
Frank Battaglia and Harry Seifert.



## Trademarks and special notices

---

© Copyright IBM Corporation 2009. All rights Reserved.

References in this document to IBM products or services do not imply that IBM intends to make them available in every country.

IBM, the IBM logo, and [ibm.com](http://ibm.com) are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. If these and other IBM trademarked terms are marked on their first occurrence in this information with a trademark symbol (® or ™), these symbols indicate U.S. registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" at [www.ibm.com/legal/copytrade.shtml](http://www.ibm.com/legal/copytrade.shtml).

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

Intel, Intel Inside (logos), MMX, and Pentium are trademarks of Intel Corporation in the United States, other countries, or both.

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

Information is provided "AS IS" without warranty of any kind.

All customer examples described are presented as illustrations of how those customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics may vary by customer.

Information concerning non-IBM products was obtained from a supplier of these products, published announcement material, or other publicly available sources and does not constitute an endorsement of such products by IBM. Sources for non-IBM list prices and performance numbers are taken from publicly available information, including vendor announcements and vendor worldwide homepages. IBM has not tested these products and cannot confirm the accuracy of performance, capability, or any other claims related to non-IBM products. Questions on the capability of non-IBM products should be addressed to the supplier of those products.

Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput or performance 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 or performance improvements equivalent to the ratios stated here.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.