



SAS deployment on IBM Power servers with IBM PowerVM dedicated-donating LPARs

Narayana Pattipati

*IBM Systems and Technology Group ISV Enablement
January 2013*

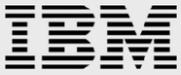
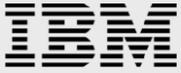


Table of contents

Abstract	1
IBM PowerVM overview	1
Processor virtualization.....	1
Memory virtualization	3
I/O virtualization	3
Dedicated-donating LPARs	3
Dedicated-donating processing mode and its benefits.....	3
How to configure an LPAR in dedicated-donating mode.....	3
SAS software deployment on dedicated-donating LPARs	4
Deployment example – SAS grid on Power servers with dedicated-donating LPARs	5
Use case – SAS grid with dedicated-donating LPARs	5
Use case demonstration.....	6
Performance of the grid workloads in dedicated-donating mode	8
Platform LSF considerations in dedicated-donating mode	8
Summary	9
Resources	9
Acknowledgements	10
About the author	10
Trademarks and special notices	11



Abstract

The paper demonstrates the use cases and benefits of deploying the SAS software on IBM Power Systems using dedicated-donating logical partitions (LPARs) on IBM PowerVM hypervisor. It shows an example of how SAS Grid can be deployed on dedicated-donating LPARs.

IBM PowerVM overview

IBM® PowerVM® provides enterprise virtualization solution for IBM Power Systems™ servers and blades. Based on more than a decade of innovation and evolution, PowerVM represents state of the art in enterprise virtualization and is broadly deployed in production environments worldwide by most Power Systems clients.

PowerVM virtualization enables continuous, dynamic resource adjustments across all partitions and operating environments, independent of physical placement, to optimize performance while minimizing energy usage. Supported environments include IBM AIX®, IBM i, Linux for Power, all on the same system. Please note SAS is only available on AIX on IBM Power systems. PowerVM also offers a secure and resilient virtualization environment, built on the advanced reliability, availability, and serviceability (RAS) features, extreme scalability and leadership performance.

The IBM POWER Hypervisor™ is integrated as part of the system firmware and supports multiple operating environments on a single system. The PowerVM virtualization technology offers the flexibility of combining dedicated and shared resources in the same partition. PowerVM helps in the following ways to improve efficiency, utilization, and flexibility:

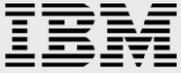
- Consolidation of multiple workloads
- Rapid deployment and scaling of workloads to meet changing business needs
- Aggregation of system resources, such as processors, memory, and storage into shared pools for dynamic reallocation between workloads
- Live transfer of running workloads between servers to support migrations and system balancing or to avoid planned downtime
- Control server sprawl and thereby reduce system management costs

Processor virtualization

In PowerVM technology, LPAR is the terminology used to refer to a virtual machine that runs AIX, IBM i, or Linux operating system. Some of the other terminologies and technologies associated with PowerVM processor virtualization are described in this section.

Dynamic logical partitioning (DLPAR)

Dynamic logical partitioning (DLPAR) increases the flexibility, allowing selected system resources, such as processors, memory, and I/O components to be added and deleted from logical partitions while they are running. This feature provides the ability to reconfigure LPARs dynamically through Hardware Management Console (HMC), Integrated Management Console (IVM), or other Power Systems virtualization management consoles.



Micro-Partitioning

The IBM Micro-Partitioning® technology allows you to allocate fractions of processors to a logical partition. A logical partition using fractions of processors is also known as a shared processor partition or micro-partition. Micro-partitions run over a set of processors called a shared processor pool. It provides the ability to distribute the processing capacity of one or more physical processors among one or more logical partitions. Thus, processors are shared among logical partitions. PowerVM Micro-Partitioning supports up to 10 logical partitions per processor core.

PowerVM processing modes

PowerVM allows the following processing modes for LPARs.

- Dedicated
- Dedicated-donating
- Shared

Processing mode – dedicated:

- In the dedicated processing mode, physical processors are assigned as a whole to LPARs.
- A dedicated LPAR cannot borrow any cores when needed.
- During low workload, the unused processing resources remain idle, causing lower system utilization.

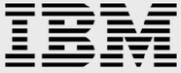
Processing mode – dedicated-donating:

Dedicated donating refers to the ability to donate idle or spare processors owned by a dedicated LPAR to a shared processor pool. The LPAR maintains absolute priority over the processors that it owns. The donated processors are returned instantaneously (within micro-seconds) to the dedicated-donating LPAR, when it needs them back. Enabling this feature can help to increase system utilization, without compromising the computing power for critical workloads in a dedicated processor.

Processing mode – shared:

In the shared mode, logical partitions use virtual processors to access fractions of physical processors. Shared LPARs can define any number of virtual processors (the maximum number is 10 times the number of processing units assigned to the partition). Shared processors use the Micro-Partitioning technology to access fractions of a processor.

- Entitled capacity (referred to as EC in this paper): Represents the guaranteed capacity that an LPAR gets under all circumstances.
- Virtual processors (referred to as VP in this paper): Virtual processors provide the flexibility to a shared LPAR to access fractions of a processor. While creating a shared LPAR, this number can be greater than or equal to entitled capacity. The maximum value of virtual processors should be less than or equal to 10 times that of entitled capacity.
- Capped: Shared LPAR cannot get more processors than EC (even through the value of VP is greater than EC)
- Uncapped: Shared LPAR can get processing units beyond EC up to VP value.



- A shared LPAR can borrow processing units from a shared pool in the uncapped mode.

Memory virtualization

IBM PowerVM technology features Active Memory Sharing, a technology that intelligently and dynamically reallocates memory from one partition to another for increased utilization, flexibility, and performance. Active Memory Sharing enables the sharing of a pool of physical memory among logical partitions on a single server.

I/O virtualization

The Virtual I/O Server (VIOS) is an integral part of the PowerVM technology. The VIOS, which is a special-purpose partition, eliminates the need for dedicated network adapters, disk adapters and disk drives, and tape adapters and tape drives in the guest partitions running as virtual machines (VMs). It can reduce costs by virtualizing I/O resources to those partitions. VIOS owns the resources that are shared with clients; a physical adapter assigned to the VIOS partition can be shared by one or more other partitions.

Dedicated-donating LPARs

The section describes the dedicated-donating processing mode of PowerVM along with its benefits. The dedicated donating mode is also called the *shared dedicated* mode. The term *dedicated-donating* is used across this paper.

Dedicated-donating processing mode and its benefits

IBM POWER6® processor-based and later servers offer the capability of harvesting unused processor cycles from dedicated-processor partitions. The dedicated-donating capability provides flexibility to donate idle or spare processors owned by a dedicated LPAR to a shared processor pool. The LPAR maintains absolute priority over the processors it owns. The donated processors are returned instantaneously (within micro-seconds) to the dedicated-donating LPAR, when it needs them back. The system administrator can control which dedicated-processor partitions can donate unused cycles. The dedicated-processor partition must be identified as a donating partition. When the processor utilization of the core goes below a threshold, and all the simultaneous multithreading (SMT) threads of the processor are idle from a hypervisor perspective, the processor will be donated to the shared processor pool.

The PowerVM hypervisor ensures that only spare processor cycles are donated. Hence, enabling this feature helps in increased system utilization, without compromising the processing power needs for workloads in a dedicated processor.

How to configure an LPAR in dedicated-donating mode

A dedicated LPAR can be made to donate idle processor cycles, by selecting the **Allow when partition is active** processor sharing option, as shown in the following figure. The option, which is available on POWER6 processor-based systems or later, indicates whether the dedicated processors are made available to shared processor partitions when the logical partition that is associated with this partition profile is active. An LPAR can be switched from the dedicated mode to the dedicated-donating mode or

from the dedicated-donating mode to the dedicated mode, using the dynamic LPAR operation. This can be done by editing the profile of an existing LPAR.

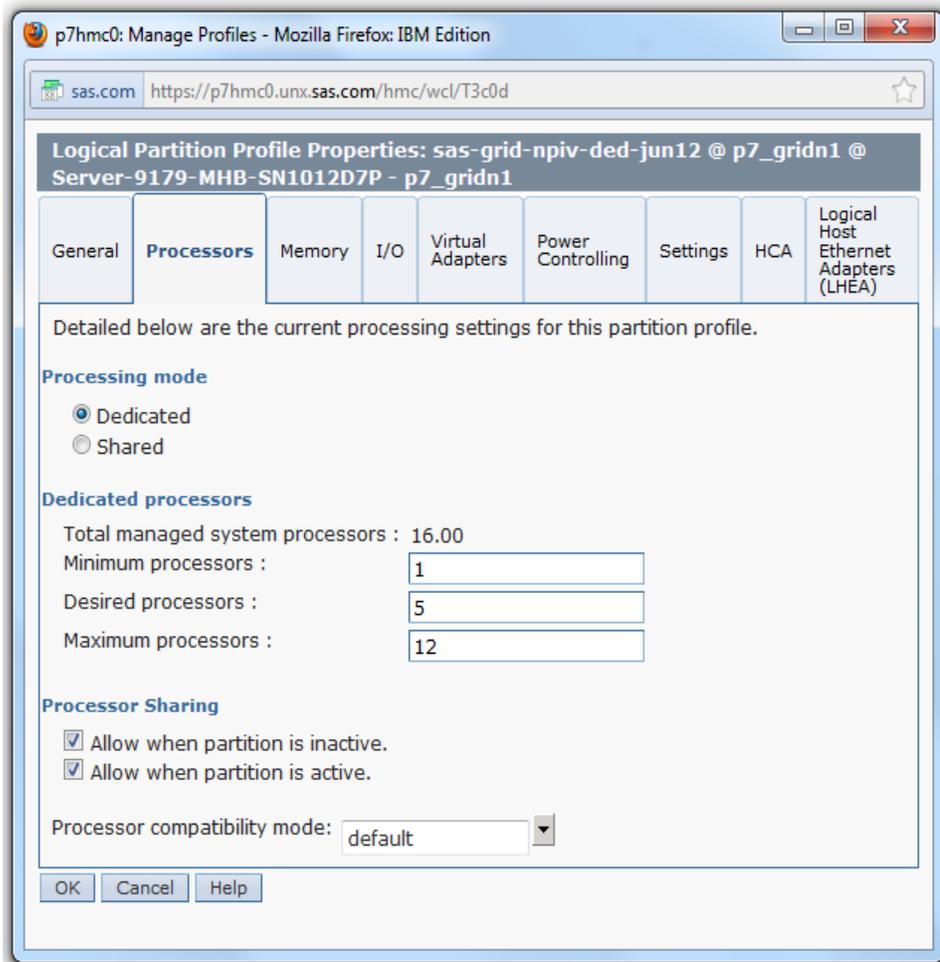
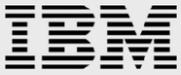


Figure 1: Configure an LPAR in the dedicated-donating mode

SAS software deployment on dedicated-donating LPARs

PowerVM dedicated-donating LPARs can be used for any SAS software deployment, regardless of the application and release version. The feature helps in optimizing server utilization and performance of SAS workloads on Power servers.

In the benchmark setup in the lab, the benchmark team had SAS Grid Computing environment on an IBM Power® 780 server with four LPARs in dedicated mode. So the team reconfigured the LPARs to dedicated-donating mode for SAS Grid quickly to demonstrate the feature and benefits. The following section demonstrates the SAS Grid deployment on dedicated-donating LPARs. The SAS grid deployment on dedicated-donating LPARs is given as an example only. As noted earlier, any SAS application can be deployed on dedicated-donating LPARs.



Deployment example – SAS grid on Power servers with dedicated-donating LPARs

The section demonstrates SAS grid deployment on Power servers with dedicated-donating LPARs. This deployment is very similar to the SAS grid deployment with dedicated LPARs, described in the SAS 9.3 *grid deployment on IBM Power servers with IBM XIV Storage System and IBM GPFS* paper published at ibm.com/support/techdocs/atmsastr.nsf/WebIndex/WP102192. No additional changes are needed for IBM Platform LSF® configuration when using dedicated-donating LPARs, instead of dedicated LPARs.

Use case – SAS grid with dedicated-donating LPARs

The following figure demonstrates SAS Grid deployment with dedicated-donating LPARs compared to dedicated LPARs. In a SAS Grid deployment with dedicated-donating LPARs, grid nodes (LPARs) donate idle processors to the shared processor pool. Non-grid LPARs (configured as shared) can use the idle processors from the shared pool, with no impact on the grid workload.

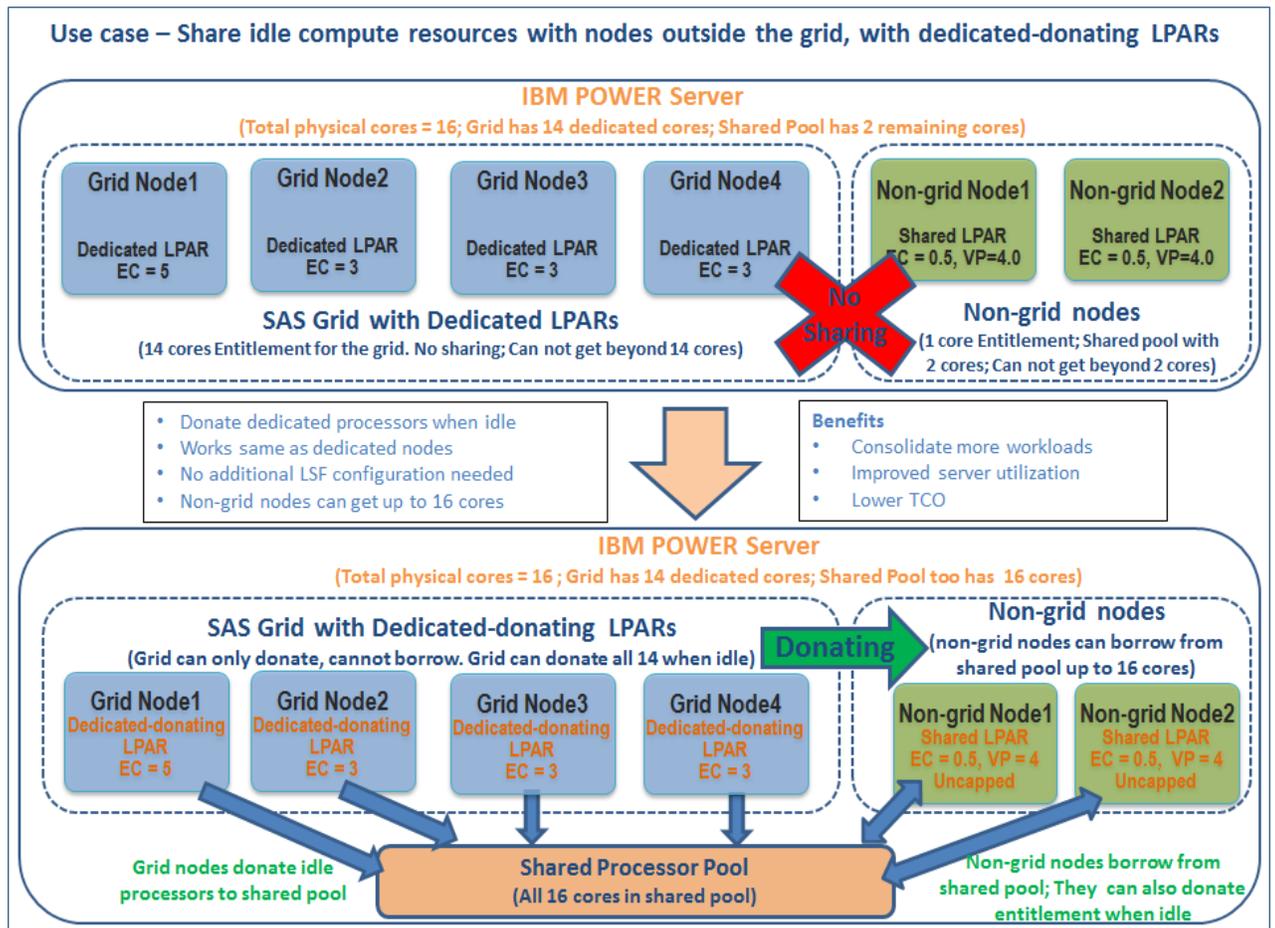


Figure 2: Sharing idle compute resources from the grid with nodes outside the grid using dedicated-donating LPARs

Use case demonstration

Test scenario:

- The Power 780 server was configured with four SAS Grid nodes (LPARs), and one non-grid LPAR.
- The four grid nodes are configured in the dedicated-donating mode with 5, 3, 3, and 3 processors entitled capacity (EC) respectively.
- The non-grid LPAR is configured in shared mode, with EC = 0.5 and VP = 4 and uncapped, which allows it to borrow up to four cores from a shared processor pool.
- A 40-session SAS Grid workload was run on the grid, and at the same time, a compute-intensive test workload was started on a non-grid LPAR.

Test results:

- When the grid is completely idle, it donated all its 14 processors to the shared pool.
- When the grid is partially idle, it donated the idle processors to the shared pool.
- The non-grid nodes could borrow from the shared pool depending on the configuration (VP)
- The grid nodes got back their donated cores instantaneously when needed.
- The following figures (Figure 3 to Figure 7) describe the test results and benefits of dedicated-donating LPARs to the SAS Grid.

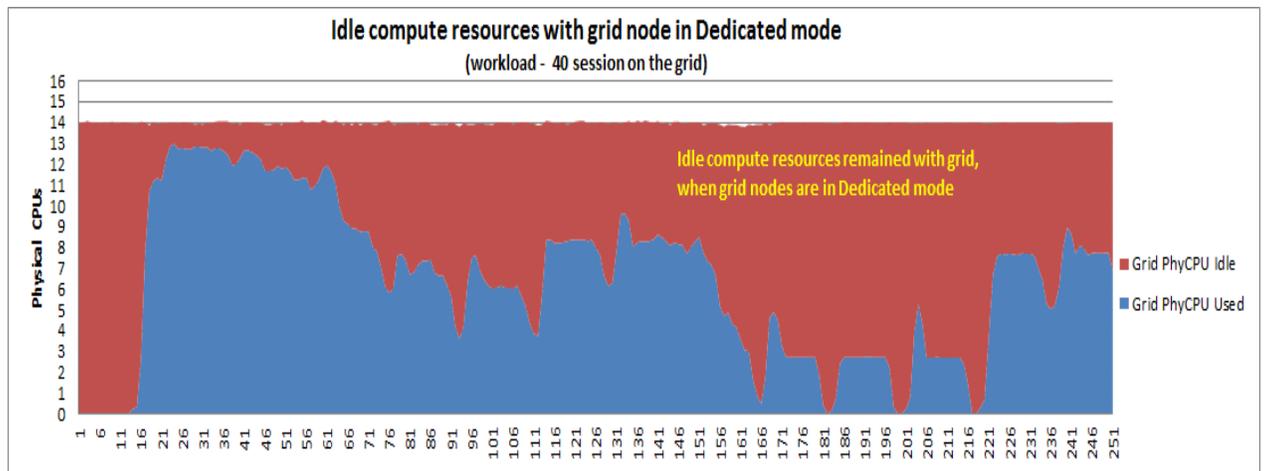


Figure 3: Idle processors remained with the grid in dedicated mode

As shown in Figure 3, in the dedicated mode, the idle processors remained with the grid nodes (LPARs) unused. Figure 4 shows how individual grid nodes donate idle processor cycles to a shared pool. Figure 5 shows the cumulative idle processors donated to the shared pool by the grid. The donated processor cycles can be borrowed by the non-grid node and use them to run test workloads effectively.

The AIX performance monitoring tool, *nmon*, shows the processors donated by a dedicated-donating LPAR. The data from *nmon* charts is used to draw the graphs in Figure 3, Figure 4, and Figure 5.

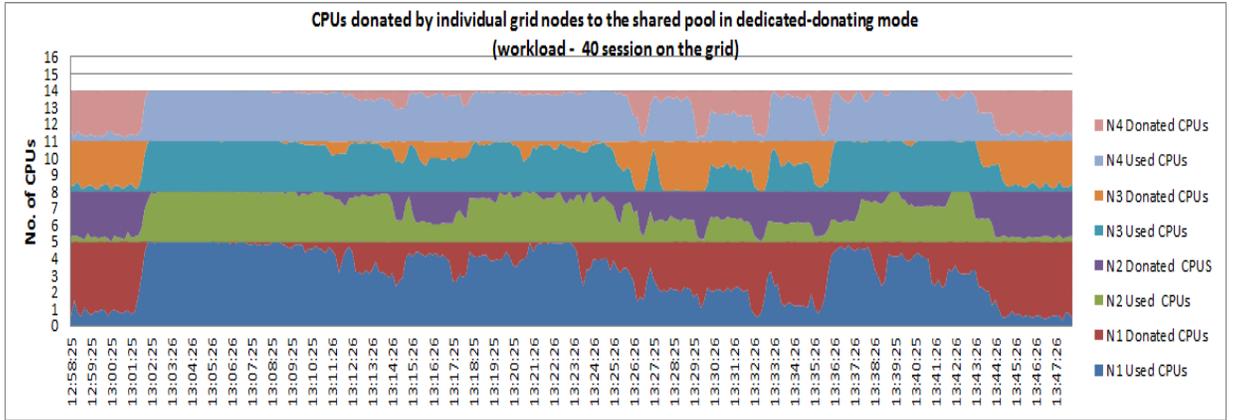


Figure 4: Idle processors donated by individual grid nodes in the dedicated-donating mode

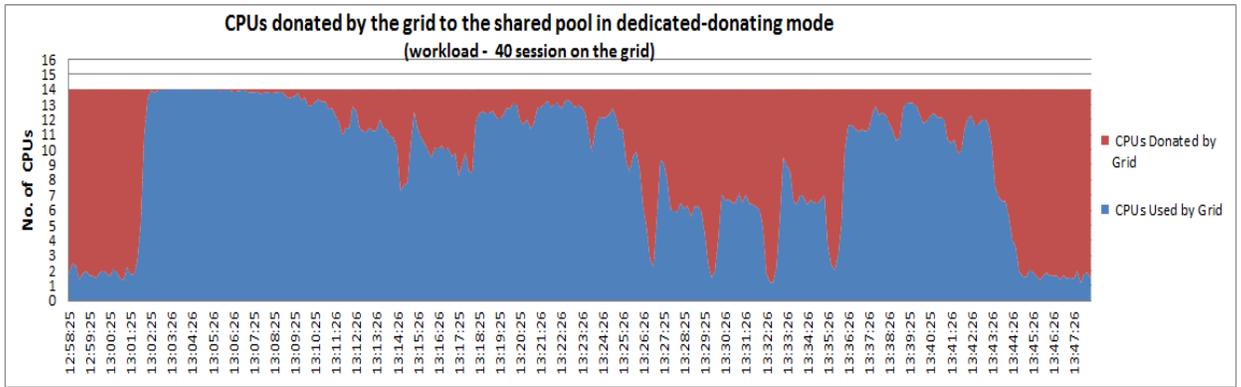


Figure 5: Idle processors (cumulative) donated by the entire grid in the dedicated-donating mode

The idle processors, donated by the dedicated-donating LPARs of the grid, can be borrowed by the non-grid LPAR to run test workloads faster. Figure 6 shows the processor usage for a compute-intensive test workload on the non-grid LPAR, when the grid is in the dedicated-mode. Even though the non-grid node is shared LPAR with EC=0.5, VP=4.0, it can only get up to two processors from the shared pool. As the test requires more than two processors, it suffers performance impact.

Figure 7 shows how the processors donated by the grid in the dedicated-donating mode are used effectively by the non-grid LPAR. The non-grid LPAR could get up to 3.5 processors to run the test workload faster.

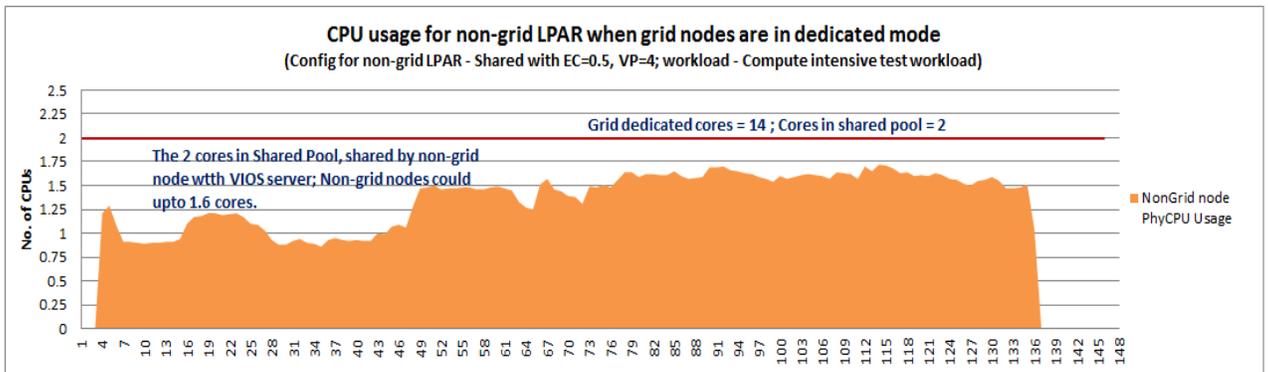


Figure 6: Processor usage for non-grid test workload when grid is in the dedicated mode

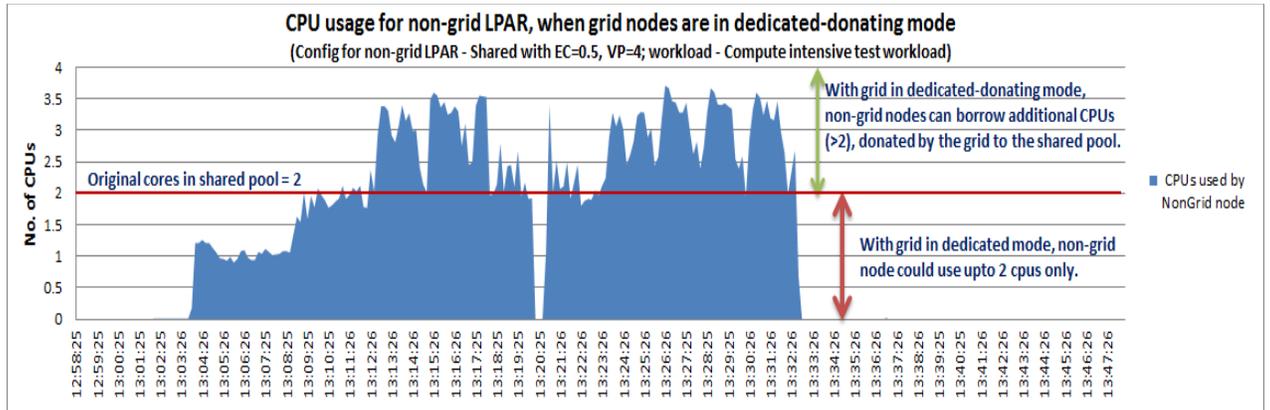


Figure 7: In dedicated-donating mode, non-grid LPAR borrowing idle cycles donated by the grid to shared pool

Performance of the grid workloads in dedicated-donating mode

During the demonstration, the benchmark team ran a 40-session SAS grid workload, with the nodes in the dedicated compared to the dedicated-donating mode. The workload showed similar performance in both the modes. In the dedicated mode, the workload took 40466 seconds of total real time (response time). In dedicated-donating mode, the workload took 39219 seconds of time real time (response time). This indicates that the dedicated-donating mode did not have any additional performance impact on the grid workload.

Platform LSF considerations in dedicated-donating mode

Platform LSF does not need any additional configuration changes when deploying SAS grid on dedicated-donating LPARs. The LSF configuration is same for either dedicated or dedicated-donating mode deployments. Figure 8 and Figure 9 show the Platform LSF utilization and run queue trends for a 40-session workload, for dedicated and dedicated-donating modes. The utilization and run queue trends are similar for both the modes. This indicates that any scheduling thresholds set based on utilization and run queue work same for dedicated and dedicated-donating mode deployments.

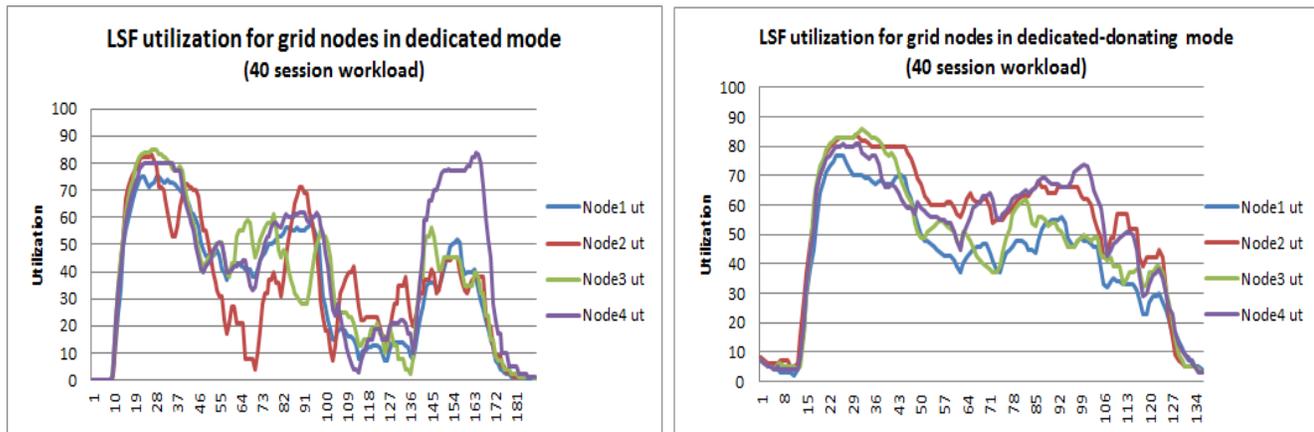


Figure 8: Platform LSF utilization trend for 40 session workload for dedicated and dedicated-donating modes

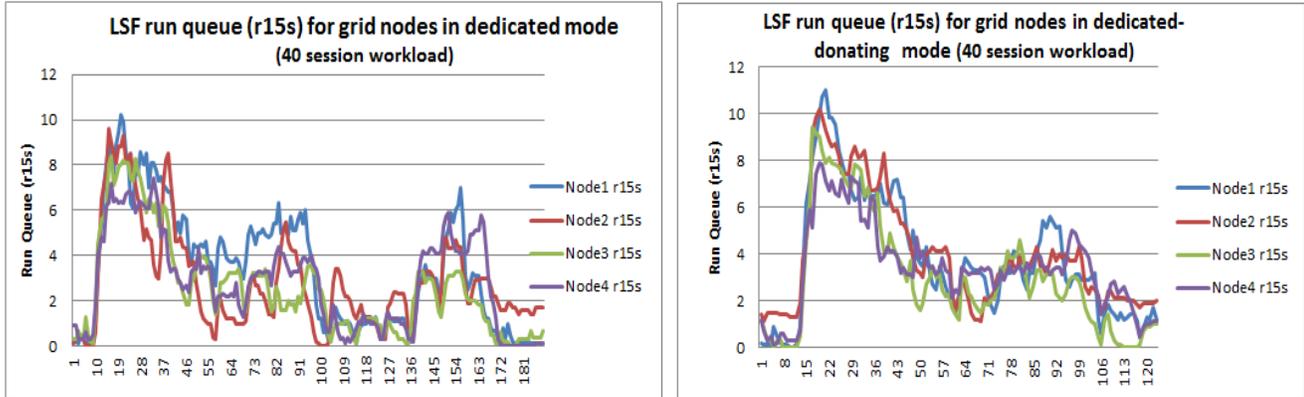


Figure 9: Platform LSF run queue (r15s) trend for 40 session workload for dedicated and dedicated-donating modes

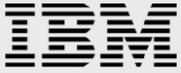
Summary

IBM PowerVM hypervisor's dedicated-donating processing mode provides flexibility to the SAS workloads to donate idle processing cycles and help in improved system utilization. The white paper demonstrates the use cases and benefits of deploying SAS applications on PowerVM dedicated-donating LPARs.

Resources

These web resources provide useful references to supplement the information contained in the paper.

- IBM Power Systems Information Center
<http://publib.boulder.ibm.com/infocenter/pseries/index.jsp>
- Power Systems on IBM PartnerWorld®
ibm.com/partnerworld/systems/p
- AIX on IBM PartnerWorld
ibm.com/partnerworld/aix
- Running SAS applications on Virtual I/O based storage configuration
ibm.com/support/techdocs/atmastr.nsf/WebIndex/WP101664
- IBM PowerVM Virtualization Managing and Monitoring
ibm.com/redbooks/abstracts/sg247590.html
- SAS 9.3 grid deployment on IBM Power servers with IBM XIV Storage System and IBM GPFS
ibm.com/support/techdocs/atmastr.nsf/WebIndex/WP102192
- Grid Computing in SAS 9.3
<http://support.sas.com/documentation/cdl/en/gridref/64808/PDF/default/gridref.pdf>
- SAS Grid Computing
 - <http://www.sas.com/resources/factsheet/sas-grid-computing-factsheet.pdf>
 - <http://www.sas.com/technologies/architecture/grid/index.html>



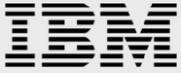
- Running Jobs with Platform LSF
http://support.sas.com/rnd/scalability/platform/PSS5.1/lsf7.05_users_guide.pdf

Acknowledgements

The author would like to thank Frank Battaglia (IBM), Harry Seifert (IBM), and Margaret Crevar (SAS) for reviewing the paper.

About the author

Narayana Pattipati works as a technical consultant in the IBM Systems and Technology Group ISV Technical Enablement team. He has more than 12 years of experience in the IT industry. He currently works with software vendors to enable their software solutions on IBM platforms. You can reach Narayana at npattipa@in.ibm.com.



Trademarks and special notices

© Copyright IBM Corporation 2013.

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 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.

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.

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 websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.