

# DB2 and the Live Partition Mobility feature of PowerVM on IBM System p using storage area network (SAN) storage

How to do many maintenance tasks without disruption to service

Skill Level: Intermediate

[Punit Shah \(punit@us.ibm.com\)](mailto:punit@us.ibm.com)  
IBM Advisory Software Engineer  
IBM

[Basker Shanmugam \(basker@us.ibm.com\)](mailto:basker@us.ibm.com)  
Technical Consultant, ISV Enablement  
IBM

05 Aug 2008

Learn about Live Partition Mobility, a feature of the System p™ virtualization PowerVM™ Enterprise edition. See how Live Partition Mobility can be applied to DB2® deployments, and how it helps you migrate AIX® and Linux® partitions and hosted applications from one physical server to another compatible physical server. Live Partition Mobility allows hardware maintenance, firmware upgrades, system maintenance, and on-the-fly server consolidation without application outage. Setup, configuration, best-practices, and performance characterization for Storage Area Network (SAN) and DB2 are covered.

## Introduction

System p virtualization technologies offer a rich set of platform deployment features. Features range from simple resource isolation to an array of the most advanced and powerful functions, including: server-resource partitioning, autonomic dynamic resource reallocation, and workload management.

Live Partition Mobility, a new feature on POWER6™ processor-based systems, is designed to enable migration of a logical partition (LPAR, or partition) from one system to another compatible system. Live Partition Mobility uses a simple and automated procedure that transfers the migrating LPAR's running state (primarily memory pages belonging to the migrating LPAR) and configuration information (such as virtual network interface, virtual Small Computer System Interface (SCSI) interface, or LPAR profiles) from the source server to the destination server without disrupting the hosted operating system, network connectivity, and hosted applications.

Live Partition Mobility gives the administrator greater control over the system resources in a data center. It increases flexibility in workload deployment and workload management, which previously wasn't possible due to narrow maintenance windows, stringent application availability requirements, or strict service level agreements that don't allow an application to be stopped.

The migration process can be performed either with a partition in powered off state or a live partition. There are two available migration types:

#### **Inactive migration**

The logical partition is powered off and moved to the destination system.

#### **Active migration**

The migration of the partition is performed while service is provided, without disrupting user activities.

Maintenance activities that previously required down time can now be performed with no disruption to services. Activities include, but are not limited to:

- Preventive hardware maintenance
- Hardware upgrades
- Server consolidation that involves reallocation of partitions among servers

DB2 V9.5 is the industry's leading information management platform. DB2 V9.5 is designed to leverage the System p virtualization technology. Innovative features such as host dynamic-resource awareness, online tuning of configuration parameters, and self-tuning memory management (STMM) make DB2 well suited for the virtualization environment.

## Prerequisites for Live Partition Mobility

It is assumed that you are knowledgeable about System p virtualization concepts, storage management, network administration, and basic system administration skills.

Live Partition Mobility has specific requirements in terms of the operating system level, firmware level, DB2 data storage layout, and network interfaces. Setting up for the migration, regardless of whether it's active or inactive, requires careful deployment planning beforehand. Sometimes you can qualify a partition for mobility by taking additional steps, such as removing physical adapters (non-virtual adapters) using a dynamic logical partitioning (DLPAR) operation.

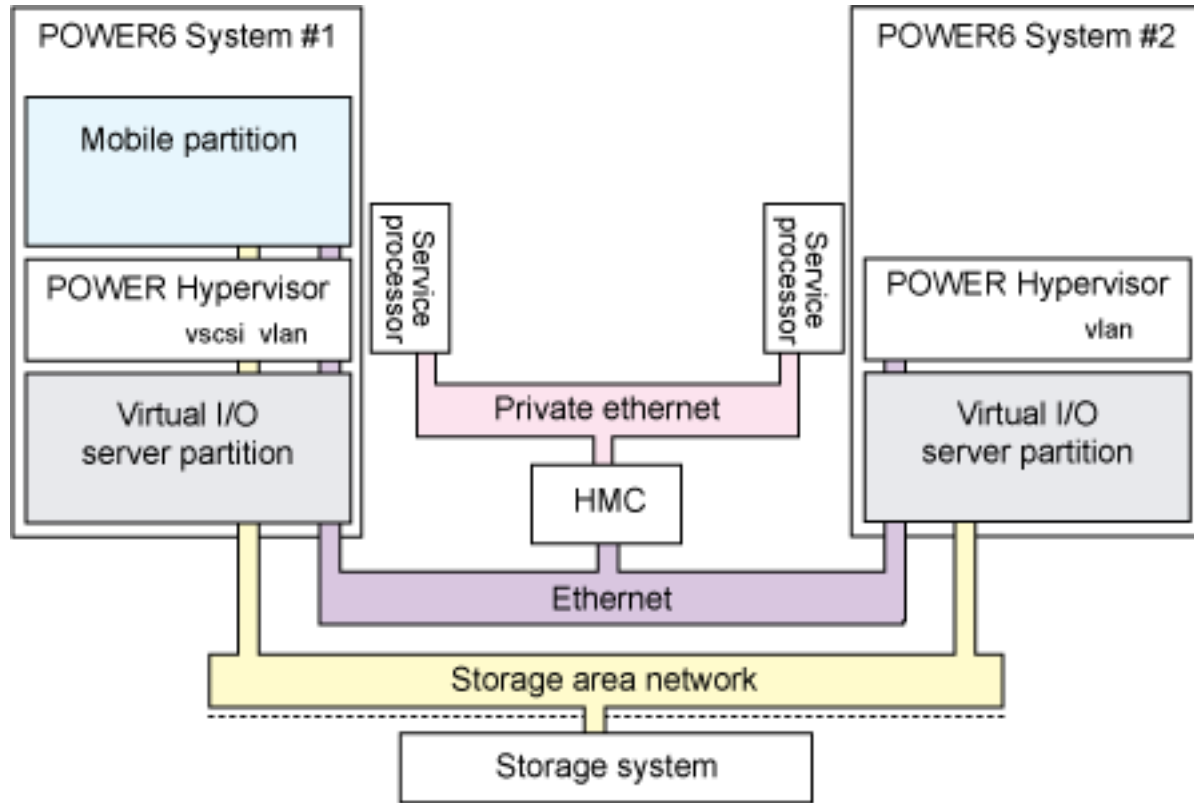
Only the partitions meeting the following requirements are eligible for migration. The typical requirements for the migration of a logical partition are:

- Two POWER6 processor-based systems controlled by the same Hardware Management Console (HMC). An optional redundant HMC configuration is supported.
- The destination system must have at least an equal amount of available memory and processors as the migrating partition's current usage. DLPAR operation might have altered partition resources (including memory, processor, and adapters) at run time after it is booted with a partition profile. Hence, the current amount of partition resources may be different than the amount specified as the "desired" in the partition profile. A partition can have multiple profiles. At operating system boot time, the system administrator selects which profile a partition is booted with. In general, a partition profile specifies operating system boot time resource requirements, such as the desired amount of memory, processor entitlement, adapters, and so on. The partition profile also specifies the minimum and maximum amount of memory, and number of virtual processors and processor entitlement.
- The operating system, applications (DB2) installation, DB2 instance data, and DB2 table data must reside on Virtual I/O storage on an external storage subsystem.
- The mobile partition's all network and disk access must be virtualized using one or more Virtual I/O Servers (VIOS).
- No other physical adapters, such as serial I/O adapter, SCSI adapters, and so on, may be used by the mobile partition at the time it is migrated. The pre-migration validation phase will fail if any physical adapter belongs to a migrating partition.
- The VIOS on both systems must have a shared Ethernet adapter configured to bridge to the same Ethernet network used by the mobile partition.

- The VIOS on both systems must be capable of providing virtual access to all storage resources that the mobile partition is using.

Figure 1 shows a general schematic to enable mobility. Each POWER6-based system is configured with a VIOS partition. The mobile partition has access to network and disk resources only through virtual adapters. The VIOS on the destination system is connected to the same network, and is configured to access the same storage used by the mobile partition.

**Figure 1. Partition mobility infrastructure**



## Configuration to enable Live Partition Mobility

This section provides a reference checklist of configuration settings for Live Partition Mobility. In general, one method to verify items in the checklist is to navigate to the respective Property screen. For example, for items in the Systems checklist navigate to **System -> Property screen**. For a VIOS partition, navigate to its Property screen. The Redbook *PowerVM Live Partition Mobility on IBM System p* (see [Resources](#)) has detailed instructions for verifying and changing these settings.

### Systems

- The source and destination servers must be at least POWER6 processor

architecture-based systems. Both systems must be controlled by the same HMC. Live Partition Mobility requires HMC Version 7.3.2 or higher.

- The memory region size (also called logical memory block (LMB) size) must be the same on the source and destination systems. It can be verified on HMC by navigating to the Memory tab on the managed system Properties screen. See [Figure 2](#) below.
- The destination system cannot be running on battery power for migration to take place.
- The destination system must have at least an equal amount of processors and memory available (as dictated by the active partition profile) as the host mobile partition.

**Figure 2. Verify system memory region (LMB) size on HMC**

General	Processors	Memory	I/O	Power-On Parameters	Capabilities	Advanced
Details of the managed system's memory are listed below.						
Installed memory:	131072MB					
Deconfigured memory:	0MB					
Available memory:	4608MB					
Configurable memory:	131072MB					
Memory region size:	256MB					
Current memory available for partition usage :	121344MB					
System firmware current memory:	9728MB					

## VIOS partitions

- At least one VIOS at release level 1.5 or higher has to be installed both on the source and destination systems.
- The partition attribute “Mover service partition” (MSP) must be enabled on both the source system and target system VIOS. You can enable this by selecting a check box when creating the VIOS partition profile, or later by navigating to the partition property and selecting a check box, as shown in [Figure 3](#).
- On the VIOS where MSP is enabled, a Virtual Asynchronous Service Interface (VASI) device must be defined and configured in order to be considered a valid MSP candidate.
- Virtual disks for use by the mobile partition have to be backed by the devices; they cannot be part of any logical volume group on VIOS.

**Figure 3. Enable or check status of mover service partition**



## Mobile partition

- AIX 5.3 technology level 7 or later.
- Ensure that Resource Monitoring and Control (RMC) connections are established.
- Redundant error path reporting has to be disabled on the migrating partition.
- The mobile partition cannot have any required virtual serial adapters, except for the two reserved for the HMC.
- The mobile partition cannot be part of a partition workload group. A partition workload group is a group of logical partitions whose resources are managed collectively by a workload management application. Workload management applications can balance memory and processor resources within groups of logical partitions without intervention from the HMC.
- The mobile partition cannot use barrier synchronization register (BSR) arrays.
- The mobile partition cannot use huge (16GB) memory pages. The 64KB and 16MB AIX page sizes are allowed, however.
- The mobile partition does not have physical or dedicated I/O adapters and devices.
- The mobile partition cannot be configured with any logical host Ethernet adapter (LHEA) devices.  
A host Ethernet adapter (HEA) is a physical Ethernet adapter that is

integrated directly into the system (GX+ bus) on a managed system. HEAs offer high throughput, low latency, and virtualization support for Ethernet connections. HEAs have the same uses as other types of physical Ethernet adapters. For example, you can use an HEA to establish a console connection to a logical partition.

## External storage

- The same Storage Area Network (SAN) disks used as virtual disks by the mobile partition must be assigned to the source and destination Virtual I/O logical partitions.
- `reserve_policy` attributes on the shared physical volumes must be set to `no_reserve` on the VIOS.
- The physical volumes must have the same unique identifier, physical identifier, or an IEEE volume attribute. This can be verified using the AIX `lsattr` command.

The VIOS must be able to accurately identify a physical volume each time it boots, even if an event such as a SAN reconfiguration or adapter change has taken place. Physical volume attributes, such as the name, address, and location, might change after the system reboots due to SAN reconfiguration. However, the VIOS must be able to recognize that this is the same device and update the virtual device mappings.

To export a physical volume as a virtual device, the physical volume must have either a unique identifier (UDID), a physical identifier (PVID), or an IEEE volume attribute.

- The mobile partition must have access to the storage through the VIOS.
- The destination VIOS must have sufficient free virtual slots to create the virtual SCSI adapters required to host the mobile partition.
- The mobile partition must have access to the same physical storage on the SAN from both the source and destination environments.

## Network considerations

- A shared Ethernet adapter has to be configured on both the source and destination VIOS.
- A virtual Ethernet adapter has to be configured on the mobile partition.

## Mobility in action

From the end user's perspective, an active migration is simply a few clicks on the HMC graphical user interface. Behind the scenes, a lot of work is involved in transferring the migrating LPAR state and configuration information from the source to the destination system (including many gigabytes of memory configured in the migrating LPAR), while keeping all services of the LPAR functional, and maintaining coherency between the source and destination systems. An active migration involves:

1. A validation phase to make sure migration criteria are met. The validation process essentially verifies the prerequisites listed above.
2. New LPAR creation on the destination system.
3. MSP setup on both VIO servers.
4. Virtual SCSI adapter setup on the destination VIO server.
5. Memory copy from the source to the destination system.
6. Suspend the LPAR on the source system and resume it on the destination system.
7. Virtual SCSI removal from the source VIO server.
8. Notification of completion to the mobile partition and VIO servers.
9. Removal of the LPAR from the source system.

## Test plan

The test plan outlines the test environment and test cases.

### Test environment

The following table summarizes the system, storage, and software used for the proof of concept.

**Table 1. Environment**

<b>Hardware</b>
<b>Software</b>
<b>Workload</b>

## Test cases

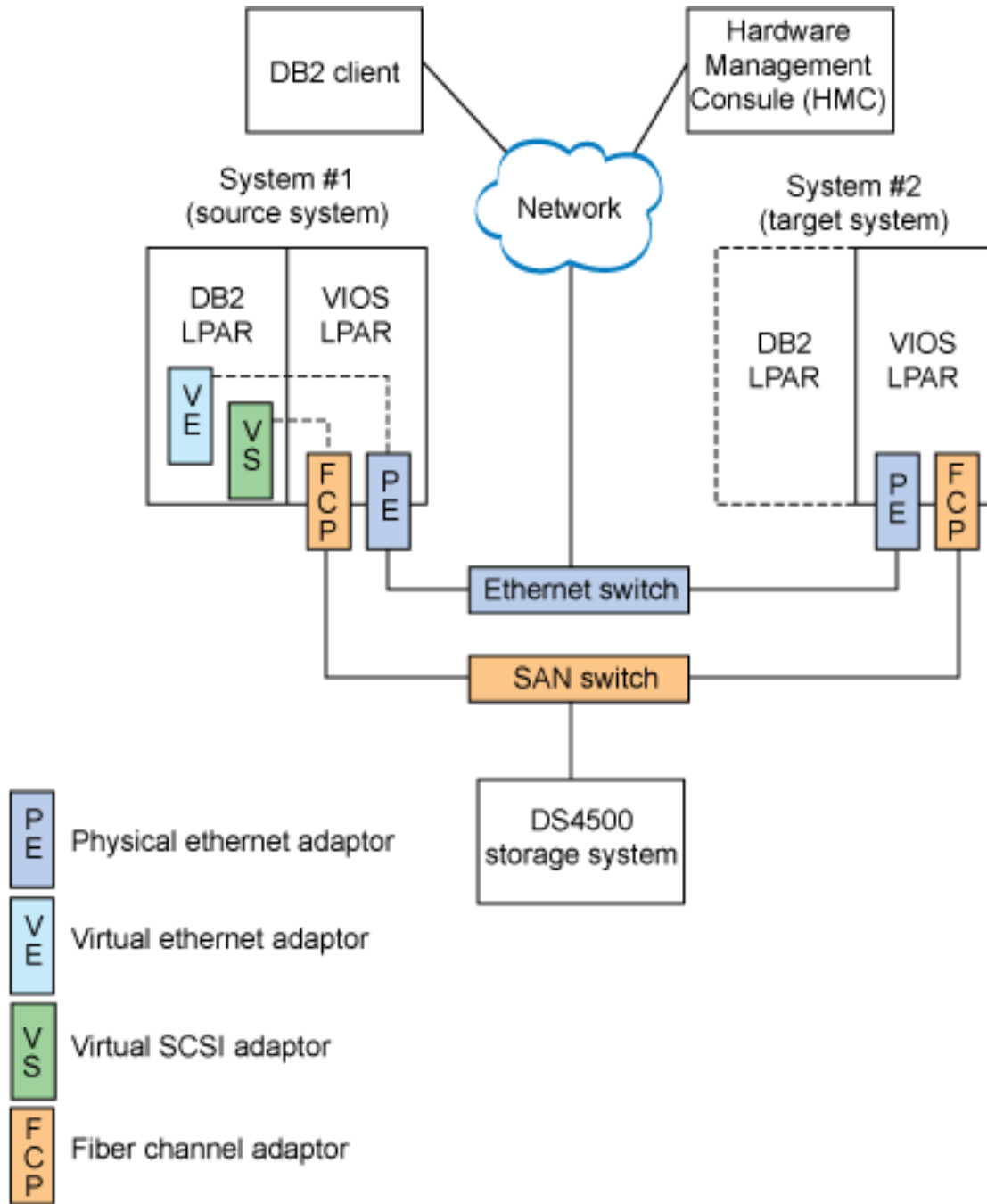
The test cases are designed to understand the following characteristics of the DB2 performance under active migration. The DB2 server instance continues to serve clients at the time of migration.

- Performance impact at various phases of the migration.
- Observe client or end user perceived throughput during active migration.
- How soon the original throughput is reestablished on the destination server.
- Effect of load on the migration time.

## Setup for mobility

Figure 4 shows the setup used for the DB2 and Live Partition Mobility performance characterization. Throughout the test, the setup evolved around best practices to ensure ease in management and higher performance throughput. As mentioned, preparation for mobility requires careful storage and network planning.

### **Figure 4. DB2 live partition mobility hardware setup**



**Hardware setup**

This section describes the system, storage, and network setup required to enable the active migration.

System #1 in Figure 4, above, is considered the source system. The LPAR (running a DB2 server instance) was originally defined on this system. System #2 was set up to receive a migrating DB2 server LPAR. The DB2 server LPAR does not have any

physical resource (including physical adapter, or local disk) belonging to it. This is one of the major requirements for mobility to work. The required network and storage interfaces are a shared Ethernet adapter (SEA), and a virtual SCSI, respectively. The DB2 network client is connected to the DB2 server by an SEA, whereas DB2 server data is hosted off the DS4500 using a virtual SCSI adapter.

The storage system is hooked up using a SAN switch to be able to connect to both the systems. The important step in configuring SAN storage is to ensure that worldwide port numbers (WWPN) are appropriately declared so both systems can access data. (At any point in time, though, only one system is accessing the storage system.) [Configuration to enable Live Partition Mobility](#) has detailed setup and configuration steps.

PowerVM Enterprise edition, a separately licensed, paid feature, must be enabled on both the systems to enable partition mobility. The minimum required AIX operating system is version 5.3 technology level 07. Specific minimum HMC and firmware versions are also required. Complete details about the environment are in [Test environment](#).

It is mandatory that you install the AIX operating system on virtual storage backed by the SAN storage device backing. It is required so that the AIX install media (specifically, rootvg) will also move along with the migrating partition.

To minimize migration latency time, high speed network infrastructure such as Gigabit or 10 Gigabit interfaces are recommended at both ends. To improve network throughput, if possible connect both the systems and HMC with the same high speed Ethernet switch, as shown in [Figure 4](#).

## DB2 setup

Just as you need an AIX installation on SAN storage, your DB2 installation location, DB2 instance directory, DB2 catalog, and all other storage (including DB2 tablespaces containers) also must be on SAN storage. The DB2 database cannot use local storage devices.

Be sure that the DB2 instance is not using any physical network adapters, disk adapters, or devices. The migration is not possible with physical adapters belonging to a migrating LPAR. The physical adapters can be removed using the DLPAR operation *before* the migration. If DB2 is accessing such devices, then the DLPAR operation to remove them will fail, resulting in a migration validation failure. The best way to avoid this situation is to follow the rule that there should be no physical adapters belonging to the LPAR. All adapters (network, storage) must be virtual adapters.

*PowerVM Virtualization on IBM System p Introduction and Configuration* (see [Resources](#)) has details about System p virtualization setup and configuration.

Partition mobility is not supported if the LPAR is using huge AIX pages (16GB page size). It's important to ensure that DB2 is not using AIX with huge pages configuration; for example, using DB2 registry variable `DB2_LARGE_PAGE_MEM=DB:16GB`. The AIX 4KB, 64KB, and 16MB page sizes, which DB2 supports, are supported for use with the partition mobility.

The DB2 database manager configuration parameter `INSTANCE_MEMORY` controls the amount of memory used by the DB2 instance. If the purpose of the migration is to assign more memory (by performing a DLPAR operation after the migration to add more memory to the LPAR hosting DB2) for the DB2 instance on the target system, the `INSTANCE_MEMORY` configuration parameter may require attention after the migration to be able to use additional memory on the target system. There is no need for any other change in the DB2 configuration.

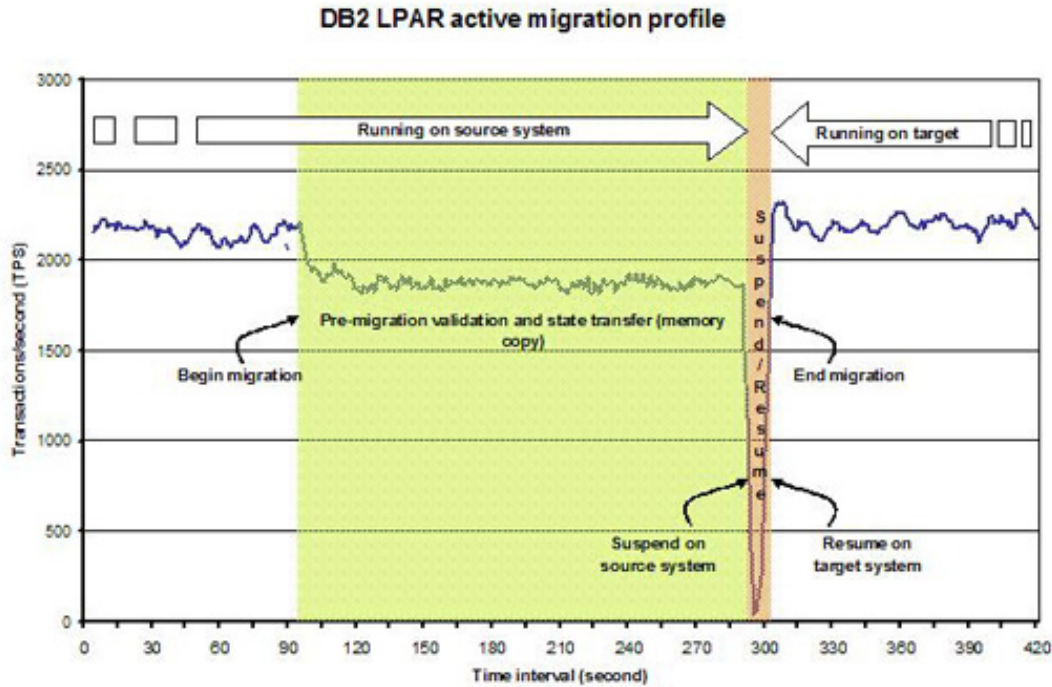
## Migration analysis

This section discusses migration performance, LPAR suspend/resume duration, effect of workload intensity, and resource requirements.

### Active migration performance profile

The shaded area in Chart 1 below shows when the mobile LPAR is migrating. Execution of the mobile LPAR on the source system ends with a suspension (in green). It's followed by a resume operation (the red shaded area) that is observed between suspension on the source system and resumption of the target system.

### Chart 1. Performance profile of DB2 active migration



Notice that the memory copy goes beyond the point of resumption of execution of the mobile LPAR on the destination system. When the LPAR is suspended, some memory pages that were recently modified have not been copied over to the destination system. When processes access memory pages that are still on the source system, the memory pages are demand-paged over from the source system. Therefore, processes can continue execution on the destination without waiting for all the remaining memory pages to be copied over.

Chart 1 also shows that database transactions continue once the mobile LPAR is resumed, and they return to full performance within a few seconds. This timing is workload dependent.

Table 2 below summarizes elapsed time at various stages of the active migration. The times are for example purposes only, and not representative. Times depend entirely on the workload, workload type, DB2 memory consumption, network adapter speed, network topology, amount of memory, storage system, and so forth.

**Table 2. DB2 live partition mobility summary**

Event	Duration (mm:ss)
Pre-migration validation and state transfer (memory copy)	03:15

<b>Total migration time (elapsed time between the “Begin migration” and “End migration” markers)</b>	03:24
--	-------

The whole migration scenario took only three minutes and 24 seconds. This time is a function of the amount of memory in an LPAR, and how frequently memory pages are updated during the migration. Further breakdown of the total migration time can be divided into two major events: state transfer took 3 minutes and 15 seconds, followed by around 9 seconds of suspend/resume duration. The original throughput (the throughput at or before the “Begin migration” marker) is established almost instantly on the target system.

As seen in [Chart 1](#), transaction throughput is affected during the state transfer phase. The average throughput dropped by about 12% between the markers “Begin migration” and “Suspend on source system.”

It is worth reemphasizing that all the performance analysis metrics and elapsed times mentioned in this discussion are examples only. They will vary widely from environment to environment.

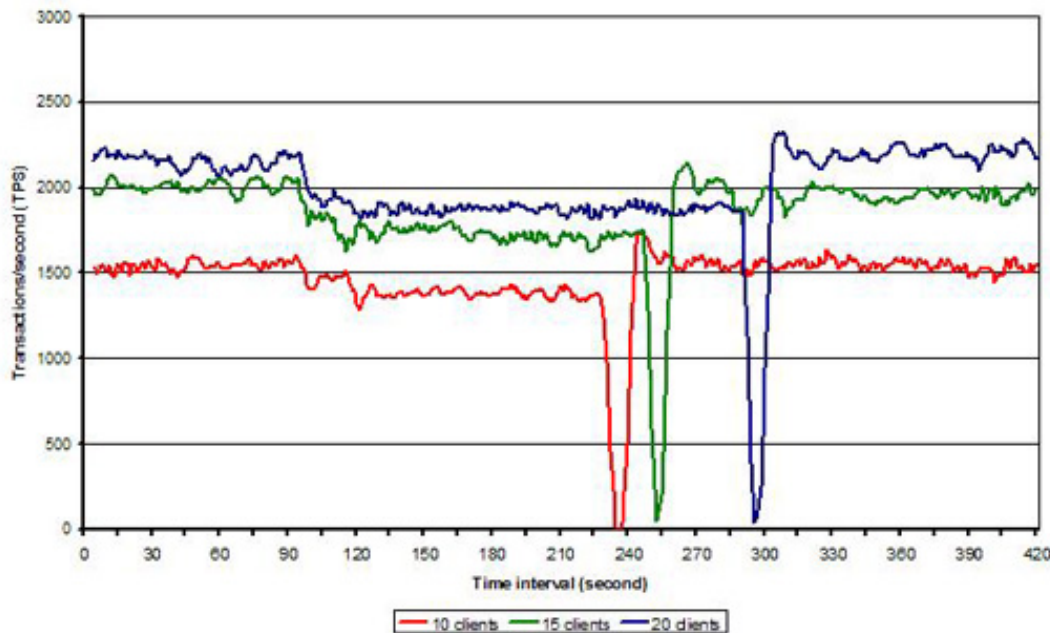
The next section discusses test results with different levels of load on the mobile LPAR in combination with different VIOS configurations.

## Load scenarios

The next set of tests examines the effect of LPAR load on the mobility characteristics. There are three load levels synthesized by a varying number of DB2 clients. The resulting average DB2 transaction throughput is approximately 1500, 1800, and 2200 transactions/second (TPS), as shown in [Chart 2](#).

### Chart 2. Effect of partition load on mobility profile

**DB2 LPAR active migration: different load scenarios**



There were not any other changes in the LPAR, AIX, or DB2 configuration. It is clear that the amount of load has apparent effect on the mobility metrics, including total migration time, state transfer time, and suspend/resume duration. The detailed analysis of exact metrics is shown in Table 3.

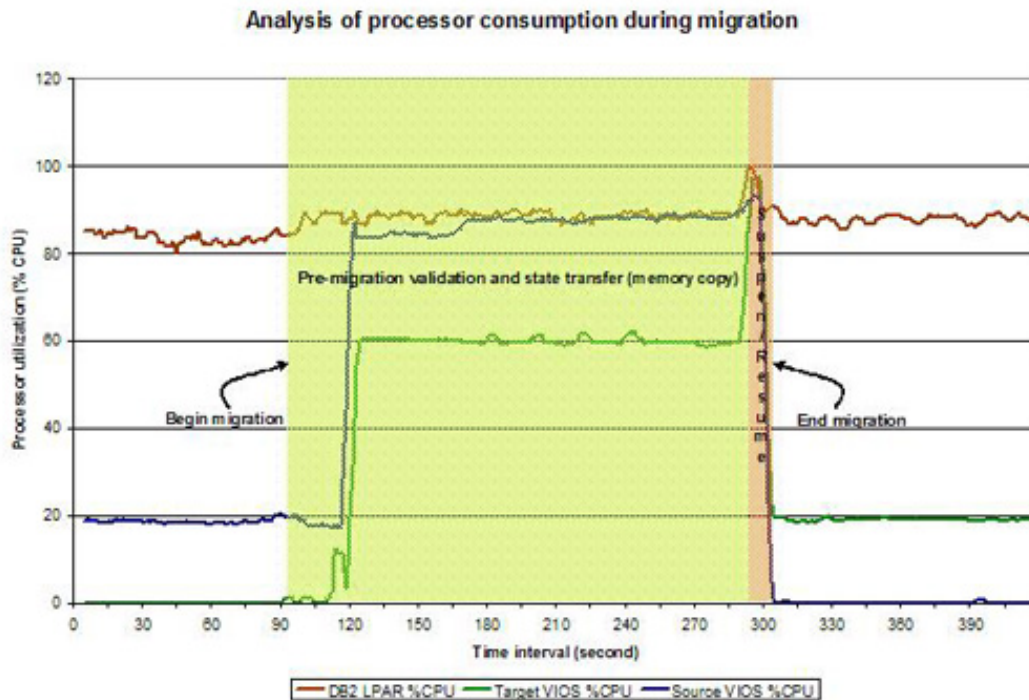
**Table 3. Effect of workload intensity on live partition mobility**

Event	Low	Medium	High
Pre-migration validation and state transfer (memory copy)	02:16	02:33	03:15
Suspend/resume duration	00:09	00:09	00:09
Total migration time (elapsed time between the “Begin migration” and “End migration” markers)	02:25	02:42	03:24

### Processor consumption

This section reviews processor consumption behavior during the various stages of active migration operation. Chart 3 below represents the same run scenario as in [Chart 1](#), but this time showing processor utilization for each of the involved partitions. Each VIOS (on the source and target system) are configured as a dedicated LPAR, with one processor, having 512MB of physical memory. The DB2 LPAR is configured as a dedicated LPAR with two processors, having 9GB of physical memory.

**Chart 3. Processor utilization of VIOS and DB2 LPAR during active migration operation**



Before the “Begin migration” event, DB2 LPAR processor utilization is around 85%. Right after the migration is triggered, DB2 LPAR processor utilization slightly increases to around 87%, and stays there until the “End migration” marker. The slight DB2 partition processor utilization increase is attributed to the DB2 partition participating in the state transfer. Compared it with [Chart 1](#), where at the same time DB2 throughput starts gradually declining up to the time marker 120 in the charts. After the marker 120, DB2 throughput stays at the same level (down by around 12% as compared to pre-migration level).

Before the migration, the source VIOS is servicing DB2 disk and network I/O requests, so it has processor utilization of around 18%. Since there's no other activity on the target system, the VIOS processor utilization is 0%. After the 120 marker in the chart, both VIOS processor utilization increases significantly because now both VIOS are conducting the state transfer operation. The source VIOS still continues servicing the DB2 disk and network I/O, in addition to performing the state transfer operation, causing processor utilization to jump to as high as 90%.

Around the same time, the target VIOS processor increases from 0% to around 60% and stays at that level until the end of the state transfer operation. Right before the suspend/resume operation, the target system VIOS processor utilization increases to as high as 95%, the source VIOS shows a similar trend, and so does the DB2 LPAR.

Observe that after the “End migration” event, the VIOS processor utilizations are reversed. The target system VIOS processor utilization is at the same level as the source system VIOS processor utilization, since now the target system is servicing the DB2 client, and vice versa.

To minimize total active migration duration, it is essential to size the VIOS appropriately. It is recommended that you use a partition type of shared dedicated capacity, or an uncapped shared-processor VIOS partition. To derive the initial amount of processing capacity, set the VIOS partition type as an uncapped shared-processor partition. Do a trial run with a peak workload, perform the active migration, and record the VIOS entitled capacity utilization (using the VIOS `viostat` command, for example). As a safeguard, add 10% spare capacity to the recorded peak entitled capacity utilization. Round up this value in the case of the shared dedicated partition. With an uncapped shared-processor partition type for VIOS, set the VIOS uncapped weight to the highest value, with entitled capacity calculated in the above step.

## Summary

Live Partition Mobility gives administrators greater control over the usage of resources in your data center. The migration process can be performed either with a partition in powered off state, or with a live partition. There are two available migration types. With inactive migration, the logical partition is powered off and moved to the destination system. With active migration, the migration of the partition is performed while service is provided, without disrupting user activities.

Maintenance activities that required down-time in the past can now be performed with no disruption to services. These activities include, but are not limited to, preventive hardware maintenance, hardware upgrades, and server consolidation that involves reallocation of partitions among servers.

The separately licensed product PowerVM Enterprise edition is required to enable the mobility. The PowerVM Enterprise edition includes Live Partition Mobility, Shared Dedicated Capacity, Micro-partitioning, VIOS, and more.

This article provided detailed analysis for DB2 9.5 running with the PowerVM Enterprise edition Live Partition Mobility feature. The DB2 9.5 instance is hosting an OLTP workload. The DB2 instance is servicing a multitude of clients, as more than 2000 transaction/second (TPS) throughput is migrated to another system. The client network connections remained intact, and the application observed a blip for only a few seconds.

### **Acknowledgements**

The authors would like to thank Horace Tong for allowing them to derive some of the content for this article from his previous work; Sunil Kamath and Peter Kokosielis for their help with DB2 setup and configuration; Rich Bassemir for review comments; and Pete Jordan for assistance with lab support and administrative matters related to the system and storage used.

# Resources

## Learn

- The [IBM System i and i5/OS Information Center](#) is your source for technical information about hardware and software solutions for the IBM System i™ platform.
- The [IBM System p and AIX Information Center](#) is your source for technical information about IBM System p and AIX.
- Go to the [IBM Publications Center](#) for many publications online, free of charge. You may view or download titles without having to place them in a shopping cart and without providing any information about yourself.
- The IBM Redbook [PowerVM Live Partition Mobility on IBM System p](#) (SG24-7460-00) explains in detail how Live Partition Mobility allows you to migrate running AIX® and Linux® partitions and their hosted applications from one physical server to another without disrupting infrastructure services.
- The IBM Redbook [PowerVM Virtualization on IBM System p: Introduction and Configuration \(Fourth Edition\)](#) (SG24-7940-03) provides an introduction to PowerVM virtualization technologies on IBM System p servers. The Advanced POWER Virtualization features and partitioning and virtualization capabilities of IBM Systems based on the Power Architecture® have been renamed to PowerVM.
- The [AIX and UNIX developerWorks zone](#) provides a wealth of information relating to all aspects of AIX systems administration.
- Find loads of information in other [IBM Redbooks](#).
- Learn more about [IBM PowerVM](#), the virtualization platform for UNIX, Linux and IBM i clients.
- "[DB2 and System p virtualization - Performance and best practices](#)" (developerWorks, Jan 2008) is an abbreviated guide for best practices for deploying IBM DB2® 9 with the IBM System p virtualization technologies.
- Read more on the [System p Virtualization Wiki](#).
- The developerWorks [Information Management](#) zone provides technical resources for IBM Information Management software.
- Browse the [technology bookstore](#) for books on these and other technical topics.

## Get products and technologies

- Download [IBM product evaluation versions](#) and get your hands on application development tools and middleware products from DB2®, Lotus®, Rational®, Tivoli®, and WebSphere®.

## Discuss

- Participate in the AIX and UNIX forums:
  - [AIX Forum](#)
  - [AIX Forum for developers](#)
  - [Cluster Systems Management](#)
  - [IBM Support Assistant Forum](#)
  - [Performance Tools Forum](#)
  - [Virtualization Forum](#)
  - More [AIX and UNIX Forums](#)
- Check out [developerWorks blogs](#) and get involved in the [developerWorks community](#).

## About the authors

### Punit Shah

Punit Shah's primary responsibilities include advanced solution development, technology exploitation, and enabling IBM middleware to use the latest System p technologies. He has worked in the enterprise application development and performance areas. Punit has also authored or coauthored several articles in this area.

---

### Basker Shanmugam

Basker Shanmugam works in the IBM ISV Business Strategy and Enablement organization as a consultant focusing on performance and benchmarking. Basker has worked in technical support and performance related areas involving Windows, Linux, AIX, Oracle, DB2, and server partitioning for over sixteen years.