Implementing AIX and PowerVM technologies with Oracle Retail Predictive Analysis Server

AIX and PowerVM technologies are tested with Oracle Retail Predictive Analysis Server with Merchandise Financial Planning

Jubal Kohlmeier
IBM STG Oracle Applications Enablement
November 2013
# Table of contents

Abstract ................................................................................................................................. 1  
Introduction .......................................................................................................................... 1  
  Prerequisites ....................................................................................................................... 1  
PowerVM technologies ........................................................................................................ 1  
  POWER7+ hardware components ...................................................................................... 2  
Oracle Retail RPAS MFP Workloads ................................................................................... 4  
  Software components ....................................................................................................... 4  
  Single user environment ................................................................................................. 4  
  Multiple User environment ............................................................................................. 5  
  Virtualized machine LPAR layout ................................................................................. 5  
  Oracle Retail RPAS disk usage and Easy Tier ............................................................... 6  
Active Memory Expansion ................................................................................................. 7  
  Requirements ................................................................................................................. 8  
  Activating AME ............................................................................................................. 8  
  Setting AME expansion factor ....................................................................................... 9  
    Determining the expansion factor ............................................................................. 9  
  Testing AME on the Oracle Retail RPAS MFP Application Server ................................. 10  
    Collecting AME recommendations ........................................................................ 10  
    Enabling 1.2 AME expansion factor ........................................................................ 11  
    Enabling 1.6 AME expansion factor ........................................................................ 12  
  AME best practices ........................................................................................................ 13  
Active Memory Sharing ..................................................................................................... 14  
  AMS requirements ....................................................................................................... 15  
  AMS Activation ............................................................................................................. 16  
  Creating the shared memory pool .............................................................................. 16  
    AMS paging devices ................................................................................................. 17  
    LPAR shared memory enabling ............................................................................. 18  
    Oracle Retail RPAS with AMS ............................................................................... 19  
  Best AMS practices ..................................................................................................... 20  
Simultaneous multithreading ............................................................................................ 20  
  Different values of simultaneous multithreading ....................................................... 20  
  Best SMT practices ..................................................................................................... 21  
Dynamic logical partitioning ............................................................................................. 21  
  Setup ............................................................................................................................ 22  
  Viewing the performance difference ....................................................................... 24  
  DLPAR best practices ............................................................................................. 25  
WPARS and Oracle Retail RPAS ....................................................................................... 26  
  Two multi-user environments ................................................................................... 26  
  Creating the system WPAR ..................................................................................... 26  
  Results from running on the System WPAR ............................................................ 26  
Live Partition Mobility ...................................................................................................... 27
Abstract

IBM PowerVM® provides the industrial-strength virtualization technologies for IBM Power Systems™ servers and blades. In this paper we explore some PowerVM features, such as EnergyScale™, Live Partitioning Mobility, and others used in conjunction with Oracle Retail Predictive Analysis Server. This paper describes how to leverage AIX and PowerVM technologies and by running simulations of hundreds of users demonstrates the benefits of AIX® and PowerVM technologies when used to run an Oracle Retail Predictive Analysis Server with Merchandise Financial Planning.

Introduction

The Oracle Retail Predictive Application Server (commonly referred to as RPAS) is a configurable software platform for developing, forecasting and planning applications, following a Client/Server OLAP model. The RPAS platform provides capabilities such as a multidimensional database structure, batch and online processing, a configurable slice-and-dice user interface, a sophisticated configurable calculation engine, user security and utility functions such as importing and exporting.

RPAS is the foundation for a significant number of applications that are part of the Oracle Retail solution footprint, such as Oracle Retail Demand Forecasting, Oracle Merchandise Financial Planning (MFP), Oracle Assortment Planning, Oracle Item Planning, Oracle Size Profile Optimization, Oracle Replenishment Optimization and Oracle Advanced Inventory Planning.

In the first quarter 2013, IBM assembled a set of systems, consisting of IBM Power 740 servers with POWER7+™ processors running AIX 6.1 and 7.1 with an IBM Storwize® V7000 storage server. The purpose of the two environments is to provide both virtualized and dedicated (non-virtualized) environments to demonstrate the ability of Oracle RPAS to capitalize on the features of PowerVM. Utilizing these test environments along with a simulated workload we will show that the many features of PowerVM have can yield significant advantages for Oracle RPAS customers.

This paper demonstrates the feasibility and viability of PowerVM technologies with Oracle Retail RPAS with MFP. PowerVM technologies with Oracle RPAS have proven robust and effective for IBM and Oracle customers.

Prerequisites

It is assumed that the reader:

- Is familiar with Oracle Retail RPAS software features and concepts
- Has an understanding of the IBM AIX 6.1 operating system
- Is aware of IBM Power® Systems features and server offerings

PowerVM technologies

IBM PowerVM virtualization technology is a combination of hardware and software that supports and manages the virtual environments on IBM servers with POWER5™, POWER5+™, POWER6®, POWER6+™, POWER7®, or POWER7+ processors.

Available on IBM Power Systems, PureSystems™ and IBM BladeCenter® servers as optional Editions, and supported by the AIX, IBM i, and Linux® operating systems, this set of comprehensive systems
The PowerVM technologies include:

- Virtual Ethernet for all networking usage during the Oracle workload simulations.
- Shared Ethernet Adapter for Oracle Retail LPARs to reduce network resources.
- Disk storage virtualization for all LPARs. N_Port ID Virtualization or NPIV is a Fibre Channel facility allowing multiple N_Port IDs to share a single physical N_Port. Each LPAR has its own disk resources attached to a SAN. No local storage is available.
- IBM Micro-Partitioning® technology using uncapped resources for impulsive Oracle Retail demands.
- Multiple shared-processor pools sharing between Oracle RPAS server and Oracle Fusion Client tiers.
- IBM Active Memory™ Sharing between Oracle Retail LPARs.
- Active Memory Expansion can reduce memory usage on the Oracle Retail LPARS.
- Shared memory pools required for AMS where loaning of memory can save resources.
- Live Partition Mobility for on-the-fly migration of a running Oracle Retail LPAR to another physical machine for performance and planned downtime.
- Dynamic Logical Partitioning can move resources to the particular Oracle Retail LPAR as demand requires.

Additionally, the paper examines technologies such as:

- SMT - Simultaneous Multi-threading of 1, 2 or 4 threads per CPU, identifying best performance for Oracle RPAS.
- Storwize V7000 including IBM System Storage® Easy Tier™ to leverage of the speed of solid-state drives (SSDs) in a storage pool that also contains hard disk drives (HDDs).
- System WPARS for consolidating operating systems within a single LPAR.
- EnergyScale provides functions to help customers understand and control IBM server power and cooling usage and to optimize performance of resources.
- Active System Optimizer and Dynamic System Optimizer automatically adjust some tuning settings to maximize the efficiency of an Oracle Retail LPAR.

This focus of this paper will not be on performance, if performance metrics are provided, it will be as an aside. The stated goal is to demonstrate combined technologies for the benefit of Oracle and IBM customers. Typical performance is expected.

**POWER7+ hardware components**

POWER7+ processor technology includes many design features that contribute to the leadership performance of the Power 740 Express servers. High processor to memory and I/O subsystem bandwidths provide faster movement of data throughout the system. The hardware topology used in the tests is shown in **Error! Reference source not found.**.

There are two client driver machines to generate load. The two IBM System x® x3850 X5 servers running Windows® can run either the single user or multiple-user test load. The load is sent through the private 1GB network (in red) to the awaiting RPAS or RPAS Fusion Client (WebLogic/application) servers. The
choice of client driver is determined by the test load model. The two Power 740 Express servers have dual virtual I/O servers, for redundancy that provide virtualized LPARs for use during our runs. The virtualized environment is amazingly adaptable to change and re-configuration. The virtualization is dependent upon shared SAN storage to provide disks and file systems that are instrumental in the ability to migrate LPAR’s during Live Partition Mobility. All in all, the configuration should be of interest to customers that want a wide range of technology features available.

Figure 1. Hardware test environment layout
Oracle Retail RPAS MFP Workloads

The Oracle Retail RPAS workload environments consist of a single user workload and a multi-user workload. The single user workload uses the classical architecture of two tiers while the multi-user workload adds in a middle tier consisting of Oracle Weblogic Server with RPAS Fusion Client.

Software components

The core software components used to complete the Oracle Retail workload environment were as follows:

- Oracle Retail RPAS Release 13.3.0.14
- IBM AIX 6.1, and AIX 7.1 (RPAS 13.3 is not officially certified on AIX 7.1)
- Oracle Retail RPAS MFP domain (multi-dimensional database)
- Oracle Application Test Suite (OATS) Retail RPAS MFP Scripts
- Oracle RPAS Weblogic Client for Oracle Retail running on Oracle WebLogic Server 10.3.

Single user environment

The single user environment is best used for testing the performance of a particular PowerVM feature or to identify the difference between tuning parameters. The single user environment has a two-tier architecture with an application server and a client front end driver.

Single User Two Tier Workload

Figure 2. Single user two tier workload

The test driver, shown above as RTDTestRunner.exe, was implemented using the RPAS Test Driver library (RTD) that is shipped as part of RPAS.

The tool allowed scripted single-user RPAS RTD commands to be sent to the application from an easy to use GUI. The script determines the work that is sent to the application server. Fourteen MFP scripts were made available to IBM, covering a variety of application scenarios, however most performance measurements were taken with the following scripts:
Multiple User environment

The multiple-user environment consists of a three-tier model and is shown in the picture below. The Oracle Application Testing Suite (OATS) generates browser-based simulated users sending request to the Oracle Weblogic Server. The Weblogic server sends the requests for each browser user to the Oracle Retail RPAS Server. The multi-user environment provides the ability to simulate up to 400 users.

![Multi-user three tier workload](image)

**Figure 3. Multi-user three tier workload**

The Oracle Application Testing Suite is a comprehensive, integrated testing environment that ensures the quality, scalability, and availability of your Web applications, Web Services, packaged Oracle Applications and Oracle Databases, and in this case Oracle Retail RPAS.

Application Testing Suite is comprised of the following tightly integrated products:

- Oracle Functional Testing - automated functional and regression testing of web applications
- Oracle Functional Testing Suite for Oracle Applications - functional and regression testing of Oracle packaged applications
- Oracle Load Testing - scalability, performance and load testing of web applications
- Oracle Load Testing Suite for Oracle Applications - scalability, performance and load testing of Oracle packaged applications
- Oracle Test Manager - test process management, including test requirements management, test management, test execution and defect tracking

Virtualized machine LPAR layout

The machine environments for the two identical Power 740 Express 16-way servers were fully virtualized. The environment is specifically configured for testing the PowerVM features and functionality. All LPARs within the virtual environment are fully virtualized, meaning that disk access used virtual SCSI and networking used virtual Ethernet. In addition, all LPARS met requirements for Live Partition Mobility. VIO servers provide the resources for virtualization. Redundant VIO servers are defined and used for the
purposes of availability. A Power 740 Express server fully supports the complete Oracle Retail RPAS environment, residing upon a single server if desired.

Figure 4 below shows the Oracle Retail RPAS LPARs and there is ample room on each machine for testing LPM to the second Power 740 server. Notice that both machines have a full Workload stack, duplicating LPAR functionality. During testing only one Oracle Retail RPAS Server and one Weblogic Server (if using the multi-user workload) is chosen for testing.

Figure 4. Multi-machine virtualized LPARs

Oracle Retail RPAS disk usage and Easy Tier

The Oracle Database disks are housed on the Storwize V7000. The Storwize V7000 is a mid-range SAN disk array with some very high-end features. The V7000 includes a feature called IBM System Storage Easy Tier. Easy Tier provides the ability to store very active disk pages on solid state drives. In order to understand how to enable Easy Tier, a few concepts need defining. Within the V7000, a storage pool is a collection of managed disks (MDisks) that jointly contain all of the data for a specified set of volumes. A managed disk (MDisk) is a logical unit of physical storage. Think of MDisks as a raid set, or mirrored set of disks, that act as a single disk of storage. By creating the storage pool (managed disk group) with both generic SSD MDisks and generic HDD MDisks, Easy Tier is automatically turned on for pools with both SSD MDisks and HDD MDisks. For the Oracle Retail RPAS Server, as soon as the SSD MDisk, mdisk0, was added to the storage pool, Easy Tier became active. See Figure 5: showing that Easy Tier is active.
Once Easy Tier becomes active, hot pages begin migrating to the SSD Mdisk, from the HDD Mdisks, mdisk4 and mdisk7 respectively. The system automatically and non-disruptively moves frequently accessed data from HDD MDisks to SSD MDisks, thus placing the data in a faster tier of storage.

**Active Memory Expansion**

Active Memory Expansion (AME) is the ability to define an area of memory for compressing infrequently used pages and reducing the memory requirements of the system as a whole. The ability to save memory helps customers reduce costs and increase overall memory resource usage.

Oracle Retail RPAS application servers can contain large memory footprints. With AME, an Oracle Retail RPAS server can potentially reduce the amount of required memory, with minimal CPU usage increases. AME can also extend the usage of current memory, assuming CPU resources are available for the additional users. The team did not test extending usage by increasing client loads, but it is assumed that if memory is saved, then that memory can be put to use for more clients.
**Requirements**

In order to use AME the system requires the following:

- AIX 6.1 TL04 SP2+
- AMS Activation
- Disabling 64KB pages

**Activating AME**

AME must be activated on the server prior to usage. To verify activation, using the HMC, go to Systems Management ->Servers within the left panel. In the right panel check the target server, Click on Task->Properties. When the window opens up, click on the Capabilities tab. Find ‘Active Memory Expansion Capable’ row and validate the value is true. If AME is not enabled, purchasing the AME license and
providing the key enables the AME feature. The figure below shows that AME is enabled on one of the Power 740 servers.

![Figure 7. Active Memory Expansion enabled](https://129.40.78.11/hmc/T645.png)

**Setting AME expansion factor**

Once AME is available on the server, each LPAR has an AME option. By default AME is disabled within the LPAR. To enable AME within the LPAR, using the HMC, navigate to the LPAR profile. Click on the memory tab, and at the bottom, click the checkbox to enable AME. Once enabled, the AME expansion factor can be modified. Be conservative in testing AME expansion factors. Aggressive expansion factor values can degrade performance of the system.

**Determining the expansion factor**

The AIX `amepat` command is a helpful tool in analyzing and determining the best possible AME expansion factor. The default `amepat` command output values for monitoring and recommending values is based on the physically defined memory within the LPAR. The physically defined memory is best used when extending the workload on the current system. The default command is:

```
amepat [-P recordfile ] [monitor_time_in_minutes]
```

For capturing recommendations for reducing memory usage within an LPAR, the AIX `amepat` command requires:

- Configuring memory so that there is little or no paging space usage
- Having AME disabled during the monitoring runs
Using the AIX `amepat` `-a` option

The `-a` option of the AIX `amepat` command bases memory estimations on the current in-use memory, rather than the physically configured memory. By discovering the ‘in-use’ memory for the workload, fine-tuning the memory footprint is then viable.

Testing AME on the Oracle Retail RPAS MFP Application Server

The Oracle Retail RPAS Application Server is the most CPU and memory intensive tier within the multi-user workload simulation. Multiple runs over the Oracle Retail RPAS application server provided the means to identify whether AME saved memory.

Collecting AME recommendations.

To set the stage, the multi-user workload generated a sufficient usage of 60% CPU activity and around 23 GB of memory usage. During steady state, the AIX `amepat` command was executed:

```
/usr/bin/amepat -e 1.10:3.00:0.3 -a 05
```

![Figure 8. Usage of the AIX amepat command tool](image)

The circled areas of the output from the AIX `amepat` command are shown above. Noted that we have taken a sample for 05 minutes (`-a 05`), but the actual time to get the sample was 6 minutes and 23 seconds. This is typical when using the tool, the sampling period may run longer than the requested timeframe. Notice also that AME is disabled and the actual memory used can be determined by the ‘Virtual Memory Size’ statistic, which is 21778 MB.
Figure 9. AIX `amepat` command tool recommendations

The continuation of the AIX `amepat` command results recommend a 1.12 AME expansion factor, stating that the average compression ratio is 2.34. The amount of modeled True Memory Size consumed memory is 23.75GB, and if the 1.12 expansion factor is true, would suggest an expanded memory size of 26.50GB.

**Enabling 1.2 AME expansion factor**

Starting conservatively, the next run occurred with AME expansion factor of 1.20. Recall the purpose is to reduce memory footprint while achieving equal performance. To enable AME with 1.2 expansion factor, the LPAR was changed as follows:

1. The LPAR was shut down prior to modifying the LPAR profile.
2. The LPAR profile enabled Active Memory Expansion. The AME expansion factor was set to 1.2.
3. The LPAR was rebooted with the new changes to AME.

Validation of the changes is just a matter of running the AIX `amepat` command without arguments. We collected more information using the AIX `amepat` command. The pertinent section is shown below:

---

1 We tested many expansion factors, all the way up to 2.4, and found no compression.
Within the figure, the AME expansion factor, and the “Modeled Expanded Memory Size” is validated. Statistics include the actual amount of virtual memory consumed, along with the actual size of memory configured (True Memory In-Use). The important statistic is the “compressed memory” value. If AME can be a benefit, “compressed memory” indicates the amount of memory that was found to be compressible. If the compressed memory was zero or zero per cent, then this workload would not be a candidate for AME.

Enabling 1.6 AME expansion factor

The expansion factor was increased to 1.6, and the test re-run. As can be seen from the AIX amepat command output in Figure 11, the situation hasn’t changed and there is still 0% ‘compressed memory’ shown.
AME best practices

The technical experiences that emerged from the work with AME yielded the following insights:

- Pay attention to the ‘virtual memory size’ field from the AIX amepat command output. This is the best indication of how much memory the applications are using for the system.
- Make sure that there is little paging, by making sure there is ample memory within the system prior to starting. Paging can affect how well the AME compression behaves.
- If the target is to reduce memory, then use the ‘-a’ option with the AIX amepat command. The ‘-a’ option uses the actual memory used for estimating memory recommendations.
- The ‘-e’ option is useful for fine tuning and cutting down on the number of simulations. This option defines how the recommendation estimates are delivered.
- When using the monitoring mode of the AIX amepat command while AME is disabled, be aware of the time that the monitor actually takes. The actual execution of the AIX amepat command typically takes much longer than the time specified as a command line argument (example 5 minutes). There have been cases on busy systems where the monitoring form of AME can take 10x or longer. Make sure your steady state environment can last the time for the AIX amepat command to finish.
- Take your ‘collection’ measurements with AME disabled. The team found that the results yielded better resultant estimates for AME.
The simulations in this paper had very tight memory constraints. As a customer, leave yourself extra memory to handle volatile CPU peaks or memory peaks. In our experience, an extra 10% additional memory can take care of 25% user load volatility, and provided room for additional users.

AME requires re-tuning if the load increases dramatically. Changing the load from 200 to 300 simulated users required additional memory.

Active Memory Sharing

Active Memory Sharing (AMS) intelligently flows memory from one logical partition (LPAR) to another for increased utilization and flexibility of memory. The concept is very similar to that of the shared processor pool and micro-partitions. It allows you to oversubscribe memory on a POWER6 (or more recent) server and let the system (the POWER Hypervisor™) allocate memory where it is needed. In situations where memory resources are tight and the environment provides the luxury of sharing memory resources, AMS is the perfect vehicle. One example might be a server that is fully configured with lower capacity memory DIMMs were adding more memory would cause replacement of all DIMMs to larger sizes. Another example is in a development environment and memory usage is well understood, there could be a need to squeeze in another LPAR for nightly compilations across global geographical distances.

AMS is best utilized with complementary LPARs where sharing of memory is not contentious. An example of two complementary LPARs would be an LPAR that works at night (batch loads) while another LPAR works during the day (business ledger LPAR). The Oracle Retail RPAS MFP workload simulation does not provide an LPAR that is complementary to another LPAR. However, for the sake of demonstrating AMS within a live Oracle Retail environment, two LPARS with a shared memory pool are chosen and configured.
Figure 12. Active Memory Sharing example

Figure 12 is a hypothetical example of three partitions sharing a shared memory pool of 16.25GB. The functionality is controlled by the hypervisor which relies on VIOS paging spaces to support the AME features.

AMS requirements

The Active Memory Sharing documentation is very good in defining the AMS requirements and for any question please defer to the standard documentation. For brevity, in order to use AMS, the system requires the following:

- AIX 6.1 TL 3 or later
- Machine Firmware 710_43 or better
- HMC Management Console 7.7.2 SP1 or better
- Virtual I/O Server version 2.1.3.10-FP23 or better
- AMS Activation
Virtualized Ethernet and virtualized disk I/O. (All I/O)
AMS paging spaces defined on the VIOS servers.

AMS Activation

To verify if AMS is activated, using the HMC, go to Systems Management -> Servers within the left panel. In the right panel check the target server, Click on Task->Properties. When the window opens up, click on the Capabilities tab. Find ‘Active Memory Sharing Capable’ row and validate the value is true. If not, purchasing the AMS license and providing the key will be required to enable the feature.

Creating the shared memory pool

Creating the shared memory pool is accomplished through the Hardware Management Console (HMC). Figure 13 shows an example screen and how to change the value to 28GB using the Shared Memory Pool wizard. It also is possible to configure manually without the wizard as shown in Figure 14.

Figure 13. Creating a shared memory pool using the Share Memory Pool wizard
The hardest task during AMS setup is specifying, allocating and defining the required paging space devices. The number of required paging devices is based on the total number of LPARs that will be sharing the single shared memory pool. There is only one shared memory pool per server. If three LPARS will share the pool, then three paging devices are required. The size of all paging devices must be at least as large as the largest LPAR’s memory footprint. For example, if there are three LPARS with memory sizes of 16 GB, 24 GB and 48 GB, then three paging devices must be defined that are at least 48 GB in size. It is prudent to think ahead and make paging devices as big as the potential growth of the largest LPAR memory footprint.

The lab technician created three paging devices per VIOS, each paging device of 48 GB in size for the three LPARS participating in AMS.

The Active Memory Sharing documentation has ample description of the process to setup the shared memory pool via the HMC, once the paging spaces are available on the VIO servers. Without the paging devices, Active Memory Sharing will be disabled within the OS. In Figure 15, three paging devices can be seen, page212, page213, and page214.
Figure 15. Example of AMS paging spaces

**LPAR shared memory enabling**

Via the HMC, enabling AMS within an LPAR is simple. Within the LPAR profile, click on the memory tab, check "shared memory". The example below shows the values for the web LPAR within the Oracle Retail environment. In the example in Figure 16, the memory pool is listed as 24 GB and the application server LPAR has 8 GB of memory (Desired Memory) defined.
Implementing AIX and PowerVM technologies with Oracle Retail Predictive Analysis Server
http://www.ibm.com/support/techdocs
© Copyright 2013, IBM Corporation

Figure 16. Example of LPAR sharing the 24 GB shared memory pool

Oracle Retail RPAS with AMS

Figure 17. `vmstat` output with `-h` option showing the loan field
Figure 17 was taken from the application server and shows the AIX `vmstat -h` command output (some fields were removed for simplicity sake). The AIX `vmstat pmem` (physical memory) field shows the application server consuming 4.05 GB of memory. The loan field starts off at .50 GB loaned away at the first interval, and ends with 1.10 GB loaned away at the end of the fifth interval. This means that the application server has loaned out 1.10GB to other LPARS within the shared memory pool. The pmem field changes when the memory returns, except that of the 14 GB memory defined for this LPAR, only 4.05 is being used, therefore plenty is unused and can be made available for other LPAR's to borrow.

**Best AMS practices**

The Oracle Retail simulation stressed the system further than a typical customer environment. The technical experiences gained by working with AMS yielded the following insights:

- It’s very important to know your workload for the LPARs using the shared memory pool of AMS. Production systems are not the place to test AMS.
- AMS works best when LPARs are not competing for the same memory resources.
- There were times when AMS borrowed memory, even when memory was still available within the shared pool. Keep track of memory usage during non-critical periods.
- Understand the AIX `vmstat` command output for AMS. Know how HPI and HPIT influence memory loaning and keep HPI numbers small.
- When combining Live Partition Mobility with AMS, the memory situation can improve. If the LPAR is migrating to another server, the memory frees up on the source server while resources on the target are taxed. Care must be planned such that migrating an LPAR with AMS doesn’t over-tax the size of the shared memory pool on the destination server. Live Partition Mobility is covered later in the document.
- AMS can co-exist with AME. Be conservative when configuring AME with AMS. Keeping the shared memory pool bigger than the sum total LPAR memory requirements is best.

**Simultaneous multithreading**

Simultaneous multithreading (SMT) is the ability of a single physical processor to simultaneously dispatch instructions from more than one hardware thread context. Because there are four hardware threads per physical processor on the POWER7 processor, additional instructions can run at the same time.

The AIX `smtctl` command controls the enabling and disabling of processor simultaneous multithreading modes. By setting different values using the AIX `smtctl` command, the team configured the underlying system to test SMT4 (four threads), SMT2 (two threads) and SMT1 (1 thread and disabled). Modifying SMT value is dynamic and changes can be seen in real time with monitoring tools like the AIX `nmon` utility.

**Different values of simultaneous multithreading**

Changing the value of the SMT occurs using the AIX `smtctl` command utility.

For example, to change the value to 2 threads per CPU run: `smtctl -t 2`.

The IBM team ran three different multi-user workload runs, one for each of the different SMT thread counts. The chart below describes the results. From the Oracle Retail RPAS MFP workload, the best
results occurred with the 4-threaded SMT thread value. The average response time is consolidated response times for the 300 users running over 30 minutes. The average CPU percentage drops as expected when the thread count increases.

<table>
<thead>
<tr>
<th>SMT Thread Value</th>
<th>Average CPU %</th>
<th>Avg Response Time (secs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>85.8</td>
<td>85</td>
</tr>
<tr>
<td>2</td>
<td>66.6</td>
<td>75</td>
</tr>
<tr>
<td>4</td>
<td>57.4</td>
<td>73</td>
</tr>
</tbody>
</table>

Table 1. SMT performance results

**Best SMT practices**

In most situations, it is best to leave the SMT value set to the default for performance. For POWER7+ processors, the default SMT value is SMT4. Testing the different values of SMT with Oracle Retail RPAS confirmed that the default works best with this workload. Here are some points to remember:

- Take caution, when dynamically modifying the SMT value on a highly loaded system. In some testing, the system backed up and delayed client response times due to changing SMT values.
- Wait for the SMT change to complete and stabilize prior to taking measurements. The time it takes to change from one SMT value to another is dependent on the load. The AIX `nmon` command is useful for providing real-time CPU de-allocation or allocation.
- When dynamically testing the SMT value, it is safer and easier to start from one thread and increase upwards to four threads; additional CPU resources become available.

**Dynamic logical partitioning**

Dynamic LPAR allows hardware resources to be moved between logical partitions. The introduction of logical partitioning has expanded the options for deploying applications and workloads. Logical partitioning provides more flexibility by making it possible to run multiple, independent operating system images concurrently on a single server. Dynamic LPAR allows partition resources to be moved from one logical partition to another without requiring a reboot of the system or affected partitions. The following are some example situations in which dynamic LPAR is useful:

- Move processors from a test partition to a production partition in periods of peak demand, then move them back again as demand decreases.
- Move memory to a partition that is doing excessive paging.
- Move an infrequently used I/O device between partitions, such as a CD-ROM for installations, or a tape drive for backups.
- Release a set of processor, memory, and I/O resources into the free pool, so that a new partition can be created from those resources.
Configure a set of minimal logical partitions to act as backup to primary logical partitions, while also keeping some set of resources available. If one of the primary logical partitions fails, you can assign available resources to that backup logical partition so that it can assume the workload.

The use of dynamic LPAR with Oracle Retail RPAS will demonstrate that Oracle Retail RPAS can take advantage of additional resources that will improve the performance of a constrained application server.

**Setup**

Oracle Retail RPAS multi-user workload was configured to run a heavy MFP workload on a constrained LPAR. Once the workload is underway, the team then added CPU and memory resources to the Oracle Retail RPAS application server. The expectation is to see how the system behaves during and after the additional resources are added in. If all goes well, performance and activity from the client’s perspective should improve.

The system was setup with 2 POWER7 3.54 GHz cores and 32 GB of memory. See Figure 18.

![Hardware Management Console](https://i.imgur.com/1294078134/hmc/connections/manuiFrameset.jsp)

**Figure 18. LPAR with 2 cores and 32 GB memory prior to DLPAR**

A load of 400 users with a ramp up rate of 1 user every 6 seconds generated a load upon the system. When steady state was achieved, it was time to add the additional 2 cores and additional 64 GB of memory. The screenshots below show the processors with additional cores for a total of 4 cores and additional memory up to 96 GB.
Figure 19. DLPAR HMC page for changing number of cores

Figure 20. Adding an additional 64 GB for a total of 96 GB with DLPAR
After the changes, it can be seen from the Hardware Management Console, that the LPAR is now running with 4 cores and 96GB of memory.

[Image: Hardware Management Console]

Figure 21. LPAR now has 4 cores and 96 GB of memory

**Viewing the performance difference**

The additional CPU and memory resources had a dramatic affect upon the workload. As can be seen in Figure 22, the system is running along at a constrained 97% average CPU percentage. At about the middle of the run statistics capture, the CPU % drops as additional CPU’s come available, along with memory. The CPU % then drops to an average of 60% for the second half of the run.

[Image: System Summary DLPAR Application Server]

CPU and Memory Resources Added.

Figure 22. The nmon CPU % output during the test run

Looking at another chart, average write I/O per second shows a dramatic improvement with the inclusion of additional cores and memory.
Figure 23. The nmon average writes per second graph

These test demonstrate the value dynamic LPAR - being able to providing on-the-fly resources to environments that could use additional resources. By adding additional cores and memory, the CPU percentage dropped to an average of 60% and also increased the I/O writes per second. The table below shows the improvements gained by adding additional resources.

<table>
<thead>
<tr>
<th>Physical CPU's</th>
<th>MEM (GB)</th>
<th>CPU Avg</th>
<th>Avg tps</th>
<th>Total Mbytes Written</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>32</td>
<td>97%</td>
<td>3,141</td>
<td>14,673</td>
</tr>
<tr>
<td>4</td>
<td>96</td>
<td>60%</td>
<td>5,309</td>
<td>25,812</td>
</tr>
</tbody>
</table>

Table 2. Performance improvements enabled by DLPAR

**DLPAR best practices**

There are some practices that help with making dynamically added resources smoother. They are:

- Before actually applying dynamic LPAR functionality to a live environment, it is best to test the full configuration DLPAR size, by booting up with the full configuration. This will avoid surprises later if machine resources are not available.
- If it is important to permanently add the additional CPU and memory resource, remember to update the HMC profile for the LPAR since dynamically adding CPU and memory resources does not update the LPAR profile.
- Adding CPU and memory resources are limited by the minimum and maximum CPU and memory values defined within the profile. Check the values prior to testing the use of dynamic LPAR capabilities.
- Dynamically removing CPU and memory may affect the performance of currently running applications. Use care and test when considering removing resources dynamically.
WPARS and Oracle Retail RPAS

Workload partitions (WPARs) are virtualized operating system environments within a single instance of the AIX operating system. WPARs secure and isolate the environment for the processes and signals that are used by enterprise applications.

System WPARs are autonomous virtual system environments with their own private file systems, users and groups, login, network space and administrative domain. Most traditional system services are virtualized at the WPAR level, and they can be independently used and managed within each WPAR. While the system WPAR environment is largely partitioned and isolated, read-only file systems might be shared between WPARs to facilitate the sharing of application data and text.

Two multi-user environments

Oracle Retail RPAS multi-user workload was set up under two different environments within the same operating system. The first environment had installed the multi-user workload environment within the global operating system. The second environment was installed under a system WPAR. The environments were identical otherwise.

Creating the system WPAR

Please refer to the WPAR documentation for creation specifics. Here is a list of steps taken to provide the environment:

- Specified a unique IP address for the system WPAR.
- Use of the AIX `mkwpar` command for creation of the WPAR.
- Use of the AIX `startwpar` command to bring up the WPAR.
- Use of the AIX `clogin` command to login to the system WPAR.
- Creation of a 200 GB multi-user workload file system to be used within the system WPAR. The file system is cloned to be identical to the global environment file system. This file system contained the RPAS user data files and binaries. This file system was not shared between the two environments.

Results from running on the System WPAR

The multi-user workload was run multiple times to gather the data for Table 3 below. There were two runs within the system WPAR and two runs without the use of system WPARs. Measurements of the system were taken within the global environment. When the RPAS multi-user workload ran within the system WPAR, the multi-user workload performed identical, according to the Client driver. A 1% cost occurred when operating within the System WPARs for both the RPAS server and the application server. This basically can be considered the cost of using a system WPAR. Response times and throughput remained excellent during the system WPAR test as compared with the global OS environment test.
Table 3. Overhead of using system WPAR

<table>
<thead>
<tr>
<th>Tier</th>
<th>CPU's</th>
<th>Memory (GB)</th>
<th>CPU %</th>
<th>Within WPAR</th>
<th>Wpar Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>App Svr</td>
<td>4</td>
<td>48</td>
<td>57.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>48</td>
<td>58.8</td>
<td>Yes</td>
<td>1.3%</td>
</tr>
<tr>
<td>Web Svr</td>
<td>4</td>
<td>64</td>
<td>75.6</td>
<td></td>
<td>1.2%</td>
</tr>
</tbody>
</table>

Live Partition Mobility

Live Partition Mobility, a component of the PowerVM Enterprise Edition hardware feature, provides the ability to move AIX, IBM i, and Linux logical partitions from one system to another. The mobility process transfers the system environment including the processor state, memory, attached virtual devices, and connected users to another machine.

Benefits of partition mobility

Partition mobility provides systems management flexibility and is designed to improve system availability. For example:

- You can avoid planned outages for hardware or firmware maintenance by moving logical partitions to another server and then performing the maintenance. Partition mobility can help because you can use it to work around scheduled maintenance activities.
- You can avoid downtime for a server upgrade by moving logical partitions to another server and then performing the upgrade. This allows you to continue your work without disruption.
- If a server indicates a potential failure, you can move its logical partitions to another server before the failure occurs. Partition mobility can help avoid unplanned downtime.
- You can consolidate workloads running on several small, under used servers onto a single large server.
- You can move workloads from server to server to optimize resource use and workload performance within your computing environment. With Live Partition Mobility, you can manage workloads with minimal downtime.
- For some systems, you can move applications from one server to an upgraded server by using IBM PowerVM Editions Live Partition Mobility or the AIX Live Application Mobility software, without affecting availability of the applications.

However, while partition mobility provides many benefits, it does not do the following functions:

- Partition mobility does not provide automatic workload balancing.
- Partition mobility does not provide a bridge to new functions. Logical partitions must be restarted and possibly reinstalled to take advantage of new features.
Requirements for LPM

The main requirement for LPM is that the LPARs be virtualized. This means that network and disks be virtualized. Once the setup with the virtualized environment is complete, testing the migration of a newly installed OS is an obvious first step. Please follow the LPM documentation for requirements for the VIO Servers and the VIO clients.

Migrating an Oracle Retail RPAS MFP multi-user workload

In order to understand how LPM works well with Oracle Retail RPAS application server, the team ran two scenarios of migration.

First, an idle Oracle Retail RPAS application server was migrated. LPM works by moving active memory from one system to another. The speed of the migration depends on the network. A private 10Gb Ethernet LAN provided the means of transferring the data between VIO Servers.

Second, a full load was placed on the Oracle RPAS application server. We will compare the differences between the two types of migrations.

Idle migration

During the migration of an idle LPAR, the following were true:

- The system was idle. We migrated the same CPU and memory sizes across machines.
- There was nothing to measure (CPU is at 0 %) except on the network, which shows 2 migrations completing a round trip.
- The total time for one migration to complete was 3:30 minutes.
- The LPAR was configured with 4 cores and 48 GB memory.
- The VIO servers were configured with 2 cores and 8 GB memory, with 10Gb Ethernet LAN.

Figure 24 shows the network IO activity of one of the two VIO servers. The graph shows two migrations of an idle system, one going to another machine, then followed by the return trip back to the original machine.

![Network I/O](image)

*Figure 24. Example of network activity during the LPM migration*
An active application server migration

The next run was made using the active multi-user workload environment. Three hundred users generated a load over an hour time period. The multi-user workload is comprised of an Oracle Retail RPAS server and RPAS Fusion Client application server. The system CPU busy graphs below show the CPU performance while the migration was happened on a live system.

Live Partition Mobility begins by copying over the environment to another machine, first moving memory that isn’t active and when enough of the environment is migrated, begins running the LPAR on the migrated machine. Looking at the CPU performance graphs seen below, the RPAS server (migrating) and the application server (not migrating) show their performance patterns during the migration. A reverse spike typically occurs when the ‘cutover’ to the target machine occurs. This is not apparent on the application server, which is the target migration LPAR. The application server shows the reverse spike seen during the cutover. Quickly the systems recover and rebalance their workload.

Figure 25. nmon RPAS application and RPAS Fusion Client CPU busy during the migration
While the migration occurs, work is still executing and the performance characteristics can be seen from the charts regarding transaction times. During the migration, the transaction average steadily rises while the work is migrating to the target machine. After the migration it takes some time before the system settles back to the performance levels of the non-migration status.

![Graph of Avg Txn Completion Time (Secs)](#)

![Graph of Avg Client Txns Per Second](#)

*Figure 26. Client-side performance indicators during RPAS LPAR migration*
Notes on using Live Partition Mobility for the RPAS application server migration:

- The total time to complete migration was: 4 minutes and 40 seconds, an additional 70 seconds over the idle system migration time.
- CPU busy statistics for the application LPAR shows the system is busy with migration, but this is not readily apparent from the normal CPU usage pattern.
- CPU busy statistics on the Web LPAR give the telltale sign of a migration, a brief period where activity synchronizes with the migration.
- The Avg Txn Completion Time graph shows the slight delay and the system recovering quickly.
- The Avg Client Txns Per Second show that during the migration the average increases from .14 seconds to .16 seconds, a 13% response time adjustment, which settles back down to the .14 average again.

**An active RPAS Fusion Client migration**

Another migration run challenged the ability of the Oracle RPAS Fusion Client (WebLogic) during Live Partition Migration. The same multi-user workload scenario was repeated, but this time the application server stayed local while the Oracle RPAS Fusion Client migrated to another machine.

Similar patterns can be seen here as well. The Oracle Retail RPAS workload continues to perform well while the migration of the LPAR continued. The Oracle RPAS Fusion Client and the RPAS server both show the reverse spike during the cutover to the target machine.

![Figure 27. nmon CPU performance indicators during RPAS server migration](image-url)
The transaction response times were similar to the times shown for the Oracle RPAS application server migration, with some demand build up during the cutover, and then the system settles back to standard workload behavior.

**Best practices**

The following best practices were observed during the migration tests:

- Consider lengthening timeouts on Client browsers to provide room for slight response time lengths during migration. Typically this is not necessary.
- Use 10Gbit network LAN, with jumbo frames on the VIO Servers. Adding jumbo frames improved performance significantly.
- Validate network performance, and tune as necessary; the speed of the migration is directly related to the network performance.
- Impact to client response times can be mitigated if migrations are performed during non-peak hours.
- Additional CPU resources recommended above what is required to support virtual I/O traffic through the VIOS when using 10Gig Ethernet. If using 1Gig Ethernet, only 1 additional CPU or vCPU is recommended regardless of the number of concurrent migrations.
For the 10Gig Ethernet network, large send/receive and LRO should be enabled at all interfaces, e.g., physical devices and SEA. Set tcp_sendspace=524288 and tcp_recvspace=524288 for the 10Gig Ethernet network. Enable jumbo frames, if the environment supports it, this will increase performance. Lastly, test the 10Gig Ethernet network performance with a tool like ftp and validate 90% of theoretical throughput.

**EnergyScale**

The energy required to power and cool computers can be a significant cost to a business – reducing profit margins and consuming resources. In addition, the cost of creating power and cooling infrastructure can be prohibitive to business growth. In response to these challenges, IBM developed EnergyScale Technology for IBM Power Systems servers. EnergyScale provides functions that help the user to understand and control IBM server power and cooling usage. This enables better facility planning, provides energy and cost savings, enables peak energy usage control, and increases system availability. Administrators may leverage EnergyScale capabilities to control the power consumption and performance of POWER® processor-based systems to meet their particular data center needs.

EnergyScale provides the ability to choose from a variety of pre-defined power management modes as well as the option to customize tuning parameters for some modes. The pre-defined modes include Static Power Saver, Dynamic Power Saver (favor power), and Dynamic Power Saver (favor performance) modes. These modes can be selected in the Power Management section of the Advanced System Management (ASM) web interface, which is accessible from the HMC. Figure 30 below shows a portion of this interface and the selection of the Dynamic Power Saver (favor performance) mode.

![Figure 30. Sample Advanced System Management screen](image_url)

The Dynamic Power Saver (favor performance) mode is suitable in environments where high performance is desired at the expense of dynamic power savings. This mode runs processor cores at Turbo frequency except when all cores are lowly utilized. By default, this mode runs all cores within a partition at the same frequency. However, it can select a different frequency for each partition. In other words, the default Dynamic Power Saver (favor performance) mode is a partition-level mode.

The Power Management section of the ASM interface has a Tuning Parameters sub-section that can be used to customize the dynamic power saver modes. Figure 31 below shows the Tuning Parameters that were selected to improve the energy efficiency for the Oracle Retail RPAS test. This whitepaper will refer to this customized mode as Dynamic Power Savings (performance-power tuned). The tuning parameters...
used for this mode select an algorithm that adjusts frequency on a per-core basis in response to the characteristics of the workload running on each core.

<table>
<thead>
<tr>
<th>Tuning Parameters</th>
<th>Current value</th>
<th>New value</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Utilization threshold for increasing frequency</td>
<td>99.9%</td>
<td>25</td>
<td>MinVal-0% MaxVal-100.0%</td>
</tr>
<tr>
<td>Utilization threshold for decreasing frequency</td>
<td>99.9%</td>
<td>25</td>
<td>MinVal-0% MaxVal-100.0%</td>
</tr>
<tr>
<td>Delta percentage for determining active cores</td>
<td>18%</td>
<td>10</td>
<td>MinVal-0% MaxVal-100%</td>
</tr>
<tr>
<td>8-bit scale factor for per-thread counter scaling</td>
<td>128</td>
<td>128</td>
<td>MinVal-1 MaxVal-255</td>
</tr>
<tr>
<td>Number of samples for computing utilization statistics</td>
<td>4</td>
<td>4</td>
<td>MinVal-1 MaxVal-1024</td>
</tr>
<tr>
<td>Algorithm Selector</td>
<td>1</td>
<td>3</td>
<td>MinVal-1 MaxVal-4</td>
</tr>
<tr>
<td>Step size for going up in frequency</td>
<td>0.8%</td>
<td>4.8</td>
<td>MinVal-0.1% MaxVal-100.0%</td>
</tr>
<tr>
<td>Step size for going down in frequency</td>
<td>0.8%</td>
<td>0.8</td>
<td>MinVal-0.1% MaxVal-100.0%</td>
</tr>
<tr>
<td>Utilization threshold to determine active cores with slack</td>
<td>98.0%</td>
<td>90</td>
<td>MinVal-0.0% MaxVal-100.0%</td>
</tr>
</tbody>
</table>

*Figure 31. ASM tuning parameters*
**Single User environment**

The effect of EnergyScale modes in the single user environment is summarized in Table 4 below.

<table>
<thead>
<tr>
<th>Power management mode</th>
<th>Average execution time (Secs.)</th>
<th>Average power (Watts)</th>
<th>Performance/watt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal</td>
<td>393.8</td>
<td>432.4</td>
<td>5.87E-06</td>
</tr>
<tr>
<td>Dynamic Power Saver (favor performance)</td>
<td>367.5</td>
<td>388.9</td>
<td>7.00E-06</td>
</tr>
<tr>
<td>Dynamic Power Saver (favor performance-power tuned)</td>
<td>382.4</td>
<td>365.9</td>
<td>7.15E-06</td>
</tr>
</tbody>
</table>

*Table 4. Single user environment EnergyScale performance effect*

Dynamic Power Saver (favor performance) runs active cores at Turbo frequency, thereby achieving better performance than nominal mode. This is seen in terms of lower execution time. In spite of this, the server power consumption is less than nominal mode. This is because Dynamic Power Saver (favor performance) mode reduces the frequency of VIOS and any other low-utilization partitions to minimum frequency. On the other hand, all cores in the server run at nominal frequency when the server operates in nominal mode. Dynamic Power Saver (performance-power tuned) mode is seen to achieve a better performance/watt metric, which is expressed here as the inverse of average execution time divided by the average power. For the Oracle Retail RPAS workload, this mode yielded performance equivalent to or better than nominal mode performance while consuming less power than Dynamic Power Saver (favor performance) mode.

**Multiple user environment**

The impact of EnergyScale modes on performance and power use in the multi-user environment was studied to identify a suitable mode for the Oracle Retail RPAS platform. This test was run with a total of 300 users with a ramp-up rate of 1 user every 7 seconds. The average execution latency for users completing in a 10 minute window in the middle of the entire test is taken as the performance metric. Of these, 5% of the reported latencies at the lower and higher ends of the latency range are discarded to avoid outliers. The average power is noted over a 20 minute window in the middle of the test where the average power was observed to be in a steady state. As in the case of the single-user environment, Dynamic Power Saver (favor performance) mode improves performance and has a higher performance per watt metric compared to nominal mode. However, unlike the single user environment, server power for the multi-user environment is higher with Dynamic Power Saver (favor performance) mode since more cores in the application server partition are active and running at Turbo frequency in order to run 300 users. Here too, Dynamic Power Saver (performance-power tuned) mode improves performance compared to nominal mode, while saving power. As a result, this mode achieves better power efficiency, expressed here as the inverse of average execution time divided by the average power.
Table 5. Multiple user environment EnergyScale performance effect

The reason for the improved power efficiency when using the Dynamic Power Saver (favor performance) mode can be understood by looking at the core frequencies. Figure 32 below shows that the core frequencies on the processor hosting the RPAS Server and RPAS Fusion Client Server (application) partitions remain close to Turbo for most of the multi-user run.

![Core Frequency](image)

Figure 32. Measured core frequencies for run using Dynamic Power Saver (favor performance) mode

Figure 33 shows core frequencies when the Dynamic Power Saver (performance-power tuned) mode. The tuning parameters have been chosen such that it increases frequency more rapidly when the load on the cores requires it, but decreases frequency more gradually. It can be seen that this mode takes cores to Turbo frequency when the load demands, but is also capable of reducing frequency when possible if the workload characteristics allow.
Best practices

The following best practices were observed during the EnergyScale tests:

- When best performance is required, use the Dynamic Power Saver (favor performance) EnergyScale mode.
- When power saving is desired while at the same time maintaining performance at nominal mode levels, use Dynamic Power Saver (performance-power tuned). The tuning parameters shown in this whitepaper are suitable for the Oracle Retail RPAS application. Additional tuning of threshold parameters may be required for other workloads.
- In an environment with a mix of partitions with different utilization levels, use one of the Dynamic Power Saver modes to reduce power. With proper tuning of the parameters, performance can be maintained at levels equivalent to or better than in nominal mode.

AIX Dynamic System Optimizer

IT environments are becoming vastly more complex as technology changes and business demands evolve. Workloads continue to change to reflect new innovations in technology and the demand for integration creates more complex interdependencies. System core counts increase and with that, threading options multiply. Virtualization and workload mobility expand the complexities even more. Tuning such an environment for optimal performance is not trivial and traditional techniques are static in nature and require constant adjustment to reflect changes in workload characteristics. Meanwhile, user expectations have evolved to the point that everything "should just work". All these pressures drive more focus on IT operations, raising expectations, and eventually, something has to give.

The IBM AIX Dynamic System Optimizer (DSO) is designed to address this challenge. It leverages the performance of POWER7 processors and incorporates IBM experience and expertise to analyze your workloads and make adjustments to your partition settings. It does this dynamically and efficiently while the application is running and works to optimize performance, all without additional time or expertise from the systems administrator. All you have to do is turn it on. Then sit back and relax, knowing that the system is acting on your behalf, using IBM's built-in expertise to constantly monitor and tune itself for
improvements in performance. The DSO is built on the Active System Optimizer (ASO) framework introduced in AIX 7.1

**Requirements for ASO/DSO**

For the ability to actively use ASO/DSO features, there are requirements that must be met:

- The POWER7 chips must be in native mode, (P7 compatible)
- The versions of Cache affinity/Memory Affinity must be at least: AIX 7.1 TL1 SP1, AIX 6.1 TL8
- For Large Page Opt/Data Stream Prefetch, versions must be at least AIX 7.1 TL2 SP1, AIX 6 TL8 SP1
- Within a dedicated LPAR, processor core folding must be disabled (default)
- Power Management disables ASO, therefore no power management can be in affect
- If Active Memory Sharing is enabled, ASO/DSO disables all except the data stream prefetch optimization
- Within shared core LPARs, system entitlement must have a minimum of two cores
- For large page optimization, the system must have a minimum of 16 GB of system memory

The goal here is to showcase ASO/DSO with Oracle Retail RPAS application server and to get some experience with how ASO/DSO operates. Within the Oracle Retail RPAS environment, and while running ASO/DSO we have established that:

- AME, AMS, Power Management must be disabled
- The LPAR is configured with dedicated cores
- No large page optimization is configured

**Enabling ASO/DSO with Oracle Retail RPAS**

It is very simple to enable ASO/DSO, using the AIX `asoo` command. A value of 1 enables the feature while a value of 0 disables ASO/DSO activity. Enable ASO/DSO by executing:

```
asoo -o aso_active=1
```

![Figure 34. Output of the `asoo -L` command](image)

Once enabled, no reboot is required, ASO/DSO is active. Logging of ASO/DSO files provides the means to examine what actions are occurring. The three logging files and their locations are:

- `aso.notice` /var/log/aso/aso.log rotate size 1m files 8 compress
- `aso.info` /var/log/aso/aso_process.log rotate size 1m files 8 compress
- `aso.debug` /var/log/aso/aso_debug.log rotate size 32m files 8 compress
ASO/DSO performance with Oracle Retail RPAS

To test the performance of ASO/DSO with Oracle Retail RPAS, a multi-user workload was used to generate a sustained load. The workload was tested with and without ASO/DSO features on AIX 7.1. The chart below shows that enabling ASO/DSO was transparent to the system and that performance was maintained.

Overall, ASO/DSO is designed for applications that share memory. There are two areas where ASO/DSO excels in performance improvement:

- Applications that have a high level of multi-threaded processes (shared memory among threads)
- Applications that use AIX shared memory segments between processes

Oracle Retail RPAS workload environment on the application server consists of single threaded processes, which do not share memory, and therefore there was little benefit found with ASO/DSO. Table 6 below provides CPU statistics which show that ASO/DSO does not effect performance.

<table>
<thead>
<tr>
<th>DSO Enabled?</th>
<th>Run Time</th>
<th>Status Time</th>
<th>Avg CPU (%)</th>
<th>Mem Used (GB)</th>
<th>Avg (%)</th>
<th>PCPU</th>
<th>AVG Users</th>
<th>Avg Tns/sec</th>
<th>Tns Errs</th>
<th>Avg Tps</th>
<th>Avg KPS</th>
<th>BU pc90</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>R148</td>
<td>0:10:00</td>
<td>73.84</td>
<td>15.2</td>
<td>16</td>
<td>91.0</td>
<td>107.9</td>
<td>0</td>
<td>0.161</td>
<td>3144.8</td>
<td>34.5</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>R152</td>
<td>0:10:00</td>
<td>75.19</td>
<td>15.3</td>
<td>16</td>
<td>93.5</td>
<td>124.8</td>
<td>1</td>
<td>0.191</td>
<td>3145.7</td>
<td>34.7</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>R149</td>
<td>0:10:00</td>
<td>74.25</td>
<td>15.3</td>
<td>16</td>
<td>91.1</td>
<td>110.6</td>
<td>0</td>
<td>0.19</td>
<td>3155.7</td>
<td>77.8</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>R151</td>
<td>0:10:00</td>
<td>75.4</td>
<td>15.2</td>
<td>16</td>
<td>94.4</td>
<td>126.1</td>
<td>1</td>
<td>0.183</td>
<td>3168.9</td>
<td>33.6</td>
<td></td>
</tr>
</tbody>
</table>

Table 6. ASO/DSO performance indicators

ASO/DSO logging with Oracle Retail RPAS application server

To understand further, let us review the ASO/DSO log files on the Oracle Retail RPAS application server. As shown in Figure 35, ASO/DSO is creating jobs and deleting jobs. ASO also shows the monitoring of workloads. Watching the log files helps to identify if ASO/DSO is working and what benefits are being described. In this example very little activity within the application server is being shown.

May 7 15:40:05 p134n205 asc:notice aso[5374150]: [monitor][8] StabilityMonitorBasic resetting job: 4 variable samples
May 7 15:40:40 p134n205 asc:notice aso[5374150]: [J05][8] Deleting job
May 7 15:54:40 p134n205 asc:notice aso[5374150]: Created MultiThread job with id 9
May 7 15:54:40 p134n205 asc:notice aso[5374150]: Adding pid 12320784 [java] to MultiThread job 9
May 7 15:54:40 p134n205 asc:notice aso[5374150]: [perf_info] ASO monitored workloads: total utilization 0.00; total load 0.34
May 7 15:54:45 p134n205 asc:notice aso[5374150]: [J05][9] Deleting job

Figure 35. Reviewing /var/log/aso/aso.log file
ASO/DSO with Oracle RPAS Fusion Client application server (WebLogic)

Recall that sharing of memory is the strength of ASO/DSO. The Oracle Retail RPAS application server did not show benefit from ASO/DSO because it doesn’t use shared memory. However, within the multi-user workload environment the RPAS Fusion Client application server and its Java™ thread are multi-threaded. Enabling the ASO/DSO for this application server, we see different log results as shown below:

```
May 15 19:34:49 p184n205 aso-info aso[3211472]: [perf_info] ASO monitored workloads; total utilization 1.74; total load 2.96
May 15 19:34:49 p184n205 aso-info aso[3211472]: [SG]0 Considering for optimization
utilization 1.00 preferred=00 attaching PredictorStrategy
May 15 19:34:49 p184n205 aso-info aso[3211472]: [EF]0 attaching strategy PredictorStrategy
May 15 19:34:49 p184n205 aso-info aso[3211472]: [SG]0 Considering for optimization
utilization 1.00 preferred=00 attaching ExperimentStrategy
May 15 19:34:49 p184n205 aso-info aso[3211472]: [EF]0 attaching strategy ExperimentStrategy
May 15 19:34:49 p184n205 aso-info aso[3211472]: [EXP] Allowing domain SRAB
May 15 19:34:49 p184n205 aso-info aso[3211472]: [FRED]0 SubSRAD [2]: -Cross: 0.10
May 15 19:34:49 p184n205 aso-info aso[3211472]: [FRED]0 SubSRAD [2]: -Cross: 0.10
May 15 19:34:49 p184n205 aso-info aso[3211472]: [FRED]0 Recommending max domain None of minimum size 68
May 15 19:34:54 p184n205 aso-info aso[3211472]: [EXP]0 Predictor recommends trying None [48]
May 15 19:34:54 p184n205 aso-info aso[3211472]: [EXP]0 giving up experimenting because only 1 domains allowed.
May 15 19:34:54 p184n205 aso-info aso[3211472]: [EXP]0 Detaching without recommendation.
```

Figure 37. Reviewing /var/log/aso/aso_debug.log file

Unfortunately, the team was not able to run the ASO/DSO experiment long enough to identify if there was performance improvement, due to length of the multi-user workload. ASO/DSO requires at least 30 minutes of sustained runtime for performance analysis algorithms to produce results, with some features taking as long as 24 hours to collect and generate ASO/DSO performance analysis and improvements. However, it can be seen that ASO/DSO is working as expected from the log files.

Conclusion

The combination of PowerVM with Oracle Retail RPAS with MFP applications for Oracle and IBM customers provides a cornucopia of features for gaining the best utilization of Power Systems hardware and resources. The Active Memory Expansion, Active Memory Sharing, Live Partition Mobility, Energyscale, and ASO/DSO features build on top of the strong virtualization foundation of PowerVM and can provide Oracle and IBM customers with newly reclaimed CPU and memory resources. With these results, customers using Oracle RPAS with MFP on IBM Power Systems servers can be confident they can take advantage of AIX, PowerVM and Power Systems technologies in their environments, and can obtain cost of ownership gains through consolidation using these virtualization technologies.
Resources

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

- Understanding Processor Utilization on POWER Systems - AIX

- IBM eServer pSeries [System p] Information Center
  http://publib.boulder.ibm.com/infocenter/pseries/index.jsp

- IBM Publications Center

- IBM Redbooks
  www.redbooks.ibm.com/

- IBM PowerVM Live Partition Mobility
  http://www.redbooks.ibm.com/abstracts/sg247460.html?Open

- Active Memory Expansion Wiki
  https://www.ibm.com/developerworks/wikis/display/WikiPtype/IBM+Active+Memory+Expansion

- Active Memory Sharing
  http://www.redbooks.ibm.com/abstracts/redp4470.html

- WPAR’s
  http://www.aixmind.com/?p=2609
  http://www.tutorialsto.com/os/aix/aix-6-workload-partitions-wpar-frequently-asked-questions.html

- Active Memory Sharing
  http://www.redbooks.ibm.com/abstracts/redp4470.html

- ASO/DSO:
About the author

Jubal Kohlmeier is an advisory software engineer in the IBM Systems & Technology Group (STG), Business Systems Division. Jubal has more than 30 years of industry experience in the computing industry and more than 15 years with Oracle products. Jubal has worked with customers using Oracle applications and Oracle Databases to help them achieve high volume mission critical solutions on IBM Power Systems and System x servers.

Acknowledgements

Assistance from Srinivasan Ramani was greatly appreciated. He was instrumental in providing instruction, his time and effort, and for providing the section on EnergyScale.
Trademarks and special notices

© Copyright. IBM Corporation 2013. 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.

Active Memory, AIX, BladeCenter, Easy Tier, EnergyScale, IBM, the IBM logo, Micro-Partitioning, Power, POWER, POWER Hypervisor, PowerVM, POWER5, POWER5+, POWER6, POWER6+, POWER7, POWER7+, Power Systems, PureSystems, Storwize, System Storage, and System x are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both:

Oracle and Java are registered trademarks of Oracle and/or its affiliates.

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

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.

The information provided in this document is distributed “AS IS” without any warranty, either express or implied.

The information in this document may include technical inaccuracies or typographical errors.

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.

All statements regarding IBM future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only. Contact your local IBM office or IBM authorized reseller for the full text of the specific Statement of Direction.

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.

Photographs shown are of engineering prototypes. Changes may be incorporated in production models.

Implementing AIX and PowerVM technologies with Oracle Retail Predictive Analysis Server
http://www.ibm.com/support/techdocs
© Copyright 2013, IBM Corporation
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.