Deepti Naphade
Richard D. Prewitt
Peter Yocom

Managing Workload Processor Resources on Cloud

There is an increasing trend to consolidate workloads in cloud environments that are highly standardized, virtualized, and automated. Each of these cloud characteristics offers significant cost benefits. For example, virtualized environments generate the most benefit when you use infrastructure with elastic scaling that meets resource requirements, and at the same time, maintains high resource utilization. You can choose a public cloud provider. However, building a private cloud provides the additional benefits of full control of access, availability and recovery, and better security which you might need for regulation compliance.

Processor consumption over time is characterized by sporadic peaks, cyclic peaks, or it is seen to be trending upward or downward. Taking these characteristics in account, this paper describes techniques to manage workload processor consumption on a private cloud environment on IBM® zEnterprise® System (zEnterprise).

This IBM Redpaper™ publication is a reference to various techniques that are available for managing workload processor resources on clouds that are set up using the IBM zEnterprise System. This publication is intended for system administrators, IT executives, and other IT professionals who are looking for ways to improve cost benefits by efficiently using processor resources.

Managing processor resources

The following techniques are described by using an example of a private cloud that runs two types of workloads:

- zEnterprize Platform Performance Manager (PPM) Dynamic CPU Resource Management. *zEnterprise Unified Resource Manager*[1](*zManager*) includes the zEnterprise Platform Performance Manager[2] which allows for the definition of workloads and performance policies. A *workload* is a classification of virtual servers that run similar work, such as IBM WebSphere® transactions or batch work.

  Using policies, you can classify virtual servers in a workload into service classes, then assign each service class performance goals and importance. Based on the performance index, the ratio of actual versus expected performance of a service class in the policy, processor shares, or entitled capacities on IBM AIX®, is shifted. The shift is from a service class that misses its goals (performance index>1 (PI>1)) and runs higher importance work, to a service class that runs less important work or meets or exceeds its goals. This function is supported on zEnterprise on the following operating systems: Linux for IBM System z® guests on z196 IBM z/VM® logical partitions, z196 logical partitions running IBM z/OS®, and AIX logical partitions on p-blades[3] on *IBM zEnterprise BladeCenter® Extension* (*zBX*).

- zEnterprise PPM Load Balancing weight recommendations. This zEnterprise functionality provides weight recommendations to a software load balancer or a hardware load balancer switch. Higher weight recommendations are made for a virtual server with the most available free processor capacity. The weight recommendations consider the free capacity of the logical partition or the blade on which the virtual server is hosted.

  Transaction distribution that is based on dynamic weights improves overall transaction throughput and performance. This functionality is supported on Linux for System z guests on z196 z/VM logical partitions, AIX logical partitions on a p-blade on zBX BladeCenter extension, and Linux on IBM System x® and Windows virtual machines on an x-blade on zBX.

- Workload mix. To determine the best workload mix for your private cloud, examine daily usage patterns for various periods (weekly, monthly, or yearly) and then determine peak values of processor utilization and peak frequency.

- Offloading to a public cloud. When you notice a continuous increase in transaction throughput and a linear increase in the usage of compute resources for workloads, you might have to provision more resources to be prepared for growth. Another possibility is to free resources on your private cloud by moving suitable workloads to public clouds, and at the same time, still comply to security requirements.

- Partition mobility[4]. *Live Partition Mobility* is a solution to meet increased resource requirements of workloads in a cloud environment. When workload processor usage increases, less important work can be moved to different servers, and the newly available server resources can absorb the peak. You might have to move a workload to a more powerful server to fulfill its peak resource requirements during the peak. Live Partition Mobility is an *IBM PowerVM® Enterprise Edition* feature and is not supported on zEnterprise.

---

[1] *ABCs of IBM z/OS System Programming Volume 10*, Chapter 5, Section 5.1: *Unified resource manager introduction*, SG24-6990-04; IBM Redbooks® publication.

[2] *ABCs of z/OS System Programming Volume 10*, Chapter 5, Section 5.20: *zEnterprise platform performance manager,* SG24-6990-04; Redbooks publication.

[3] *ABCs of z/OS System Programming Volume 10*, Chapter 3, Section 3.1: *zEnterprise overview*, SG24-6990-04; Redbooks publication.

[4] *IBM PowerVM Live Partition Mobility*, Chapter 1, Section 1.4: *Live Partition Mobility is the answer*, SG24-7460-01; Redbooks publication.

Workloads on a private cloud

Use the zEnterprise system to combine IBM multiple processor architectures into one integrated pool of virtual resources, all managed under a single management framework. In addition to the core server, zEnterprise offers *IBM zEnterprise BladeCenter Extension* (zBX), which provides integration with IBM blade and accelerator technology, allowing you to build heterogeneous clouds.

For example, we use a private cloud for an insurance company hosted on an IBM zEnterprise System. Here, two types of workloads are run on multiple Linux guests hosted on multiple IBM z/VM LPARs. The Linux for System z guests on z/VM LPARs run batch and transaction workloads, as shown in Figure 1.

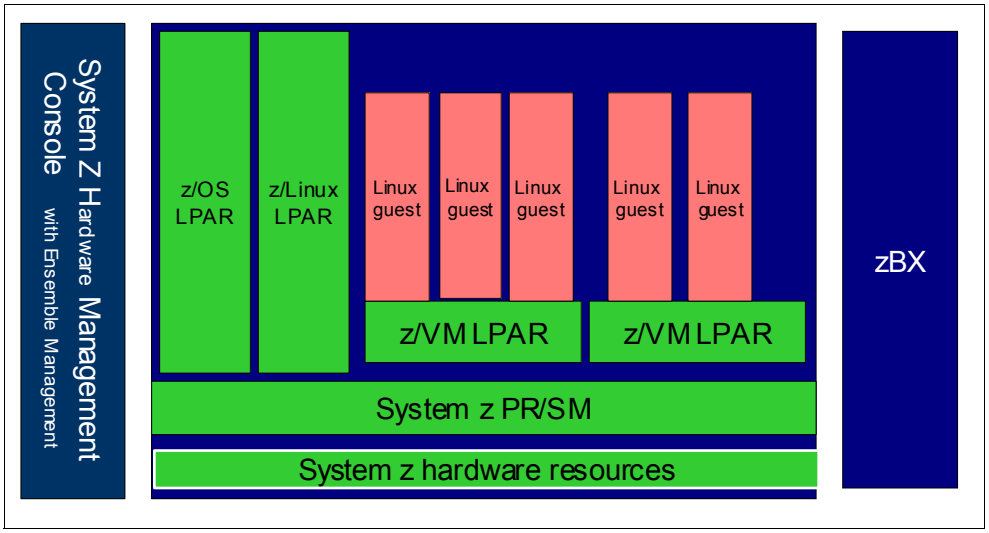

Figure 1 Linux guests (shown in red) on z/VM LPARs running either a transaction or batch workload

There are two types of workloads, but each Linux guest runs only one type:

- A customer relationship management (CRM) application that processes transactions such as setting up new customer accounts, updating customer policies, and deactivating customer accounts. These transactions run between the working hours of 8:00 a.m. - 8:00 p.m., seven days a week. After 8:00 p.m., the insurance company employees use the same application to submit transactions to process insurance claims.
- A batch workload to archive and print claim data is scheduled to begin at midnight each night and completes every day before 10:00 a.m.

Consider z/VM LPAR1. Figure 2 shows the contribution of each workload to the overall processor utilization percentage over time. The transaction workload uses most of the processors during daytime, and that batch workload uses most of the processors during nighttime. The blue line shows the total processor utilization percentage which remains below 100%.

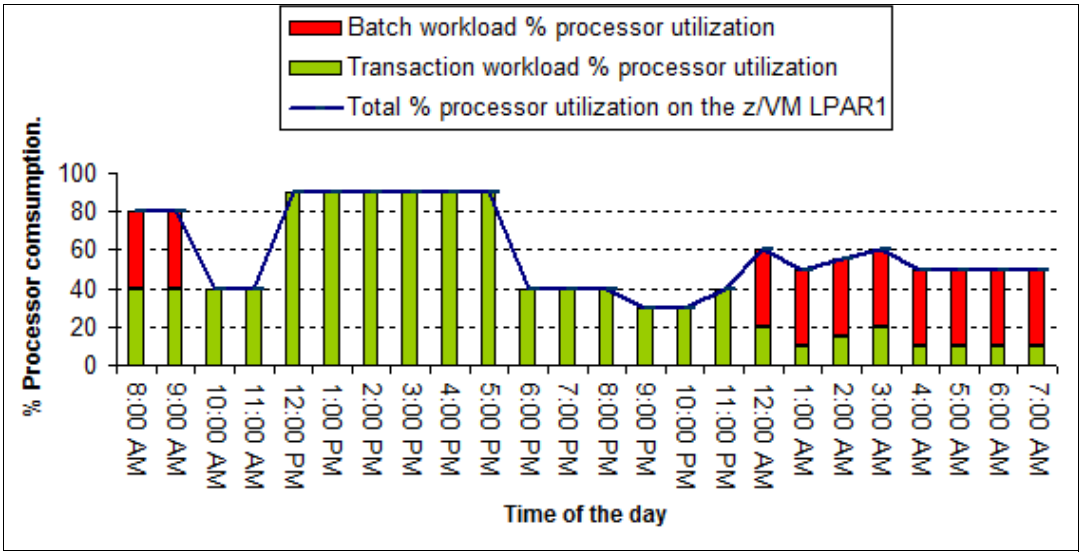

Figure 2 Total percent processor utilization of z/VM LPAR1 versus time of the day

PPM Dynamic CPU Resource Management

The first of five techniques of managing resources is the zEnterprize Platform Performance Manager (PPM) Dynamic CPU Resource Management. By using the example of workloads on the private cloud, zManager is used to define two workloads, *batch* and *transaction* to use for PPM Dynamic CPU Resource Management. See Figure 3 on page 5.

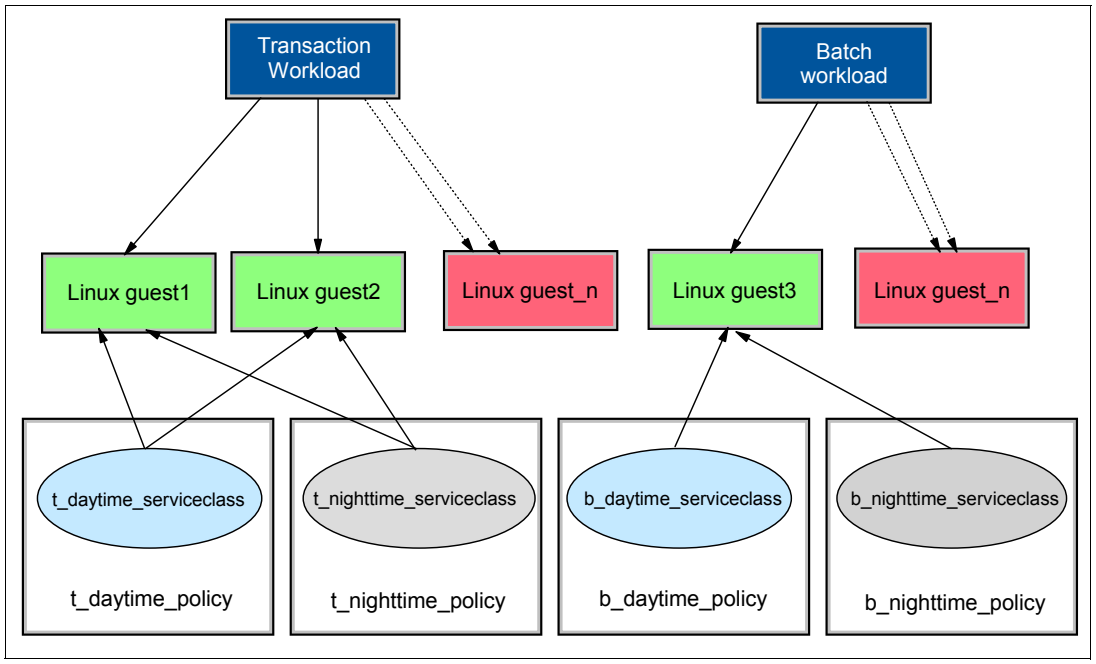

Figure 3 Workload and policies set up using zManager (guests in green are hosted on z/VM LPAR1)

Two unexpected spikes in processor consumption occur on batch and transaction workloads, as shown in Figure 4. In both cases, the demand of processor resources exceeds the available processor units (that is, greater than 100%) on z/VM LPAR1.

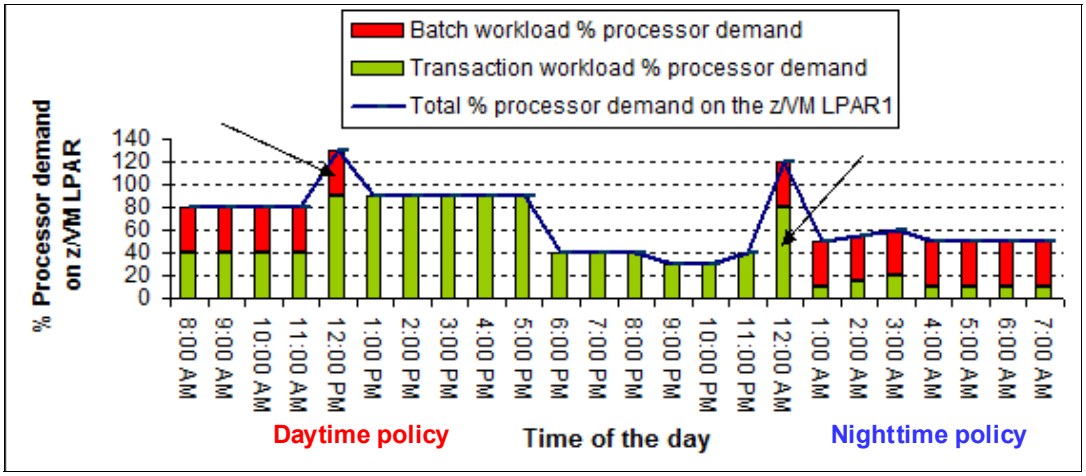

Figure 4 Contention of processor resources at noon and at midnight

The batch workload does not finish by 10:00 a.m. but runs until noon when the transaction workload also picks up. The result is that, at noon, the demand on processor resources exceeds 100% of the available resources on z/VM LPAR1.

If company employees submit many transactions around midnight, another surge in processor consumption occurs and the demand of processor resources again exceeds 100%. These examples are two cases where PPM Dynamic CPU Resource Management can help resolve the contention of processor resources by shifting processor resources to the more important workload.

Use zManager to set up performance policies to help the PPM CPU Resource Manager determine which workloads are important, and to reinforce its performance objectives or

goals. Each workload has different importance and goals during the day and night. Therefore, two performance policies must be defined for each workload, as shown in Figure 3 on page 5. During the day, it is more important for the transaction workload to complete in a timely manner. During the night, the batch workload is more important than the transaction workload and has more aggressive performance goals.

For transaction workload, we define policy *t_daytime_policy* that is active during the day and define a service class *t_daytime_serviceclass*. This service class classifies Linux guest1 and Linux guest2. This service class has an aggressive goal and high importance.

We define another policy, *t_nighttime_policy*, for the transaction workload that is active during nighttime. This policy defines *t_nighttime_serviceclass* that classifies Linux guest1 and Linux guest2. This service class has a lenient goal and low importance.

For batch workload, we define policy *b_daytime_policy* that is active during the day. This policy defines a service class *b_daytime_serviceclass* that classifies Linux guest3. This service class has a lenient goal and low importance.

Another policy, *b_nighttime_policy*, is defined for batch workload that is active during nighttime. This policy defines *b_nighttime_serviceclass* that classifies Linux guest3. This service class has an aggressive goal and high importance.

During the daytime, the *b_daytime_policy* is active for batch workload, and *t_daytime_policy* is active for transaction workload. Around noon, the PPM Dynamic CPU resource manager detects contention for processor resources and a bad performance index (PI>1) for *t_daytime_serviceclass*.

Processor shares are automatically shifted by PPM Dynamic CPU management to Linux guest1 and Linux guest2. These classifications are part of the *t_daytime_serviceclass* from Linux guest3, which is part of the less important *b_daytime_serviceclass* of batch workload. The shares shift until the goal of the *t_daytime_serviceclass* is met and its PI is close to 1.

Figure 5 shows shares shifting from Linux guest3 that is executing batch workload to Linux guest1, and Linux guest2 that is executing transaction workload. Linux guest3 shares the decrease from 200 to 160, and Linux guest1 and Linux guest2 share the increase from 200 to 220.

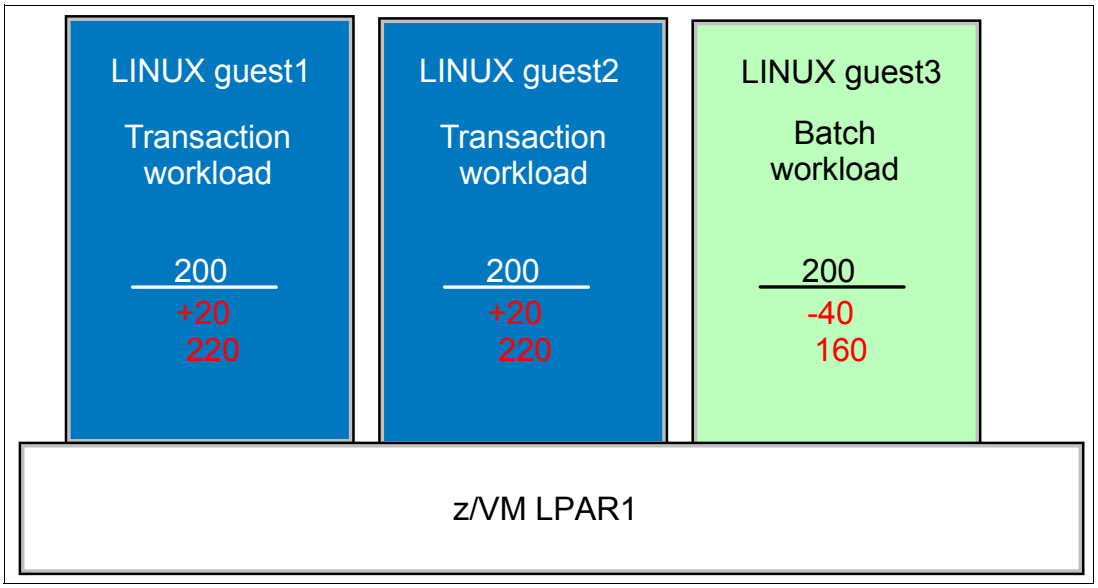

Figure 5 Shares shift until goals of transaction workload are met

Figure 6 shows the PI of the *t_daytime_serviceclass* as transaction workload improves when more shares are shifted to Linux guests that run transaction workload. The shares are shifted back to the donor servers when the extra demand on the processor resources ends.

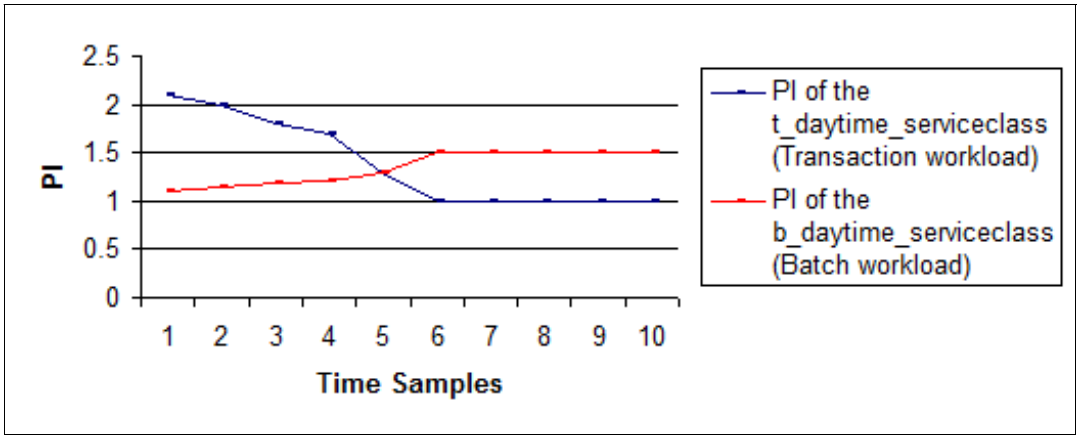

Figure 6 Performance Index (PI) versus time

During nighttime, the *b_nighttime_policy* and the *t_nighttime_policy* are active for each workload. Around midnight, the PPM Dynamic CPU resource manager detects contention for processor resources and a bad performance index (PI>1) for the *b_nighttime_serviceclass*.

Processor shares are automatically shifted by the PPM Dynamic CPU Manager to Linux guest3, which is part of the high importance *b_nighttime_serviceclass* of batch workload from Linux guest1 and Linux guest 2. These classifications are part of the less important *t_nighttime_serviceclass* transaction workload. The shares continue to shift until the goal of the *b_nighttime_serviceclass* is met and its PI is close to 1.

Use this functionality in the following instances:

- When there a sporadic contention for processor resources
- When there is a cyclic contention pattern of processor resources
- When it is alright to set a lenient goal and temporarily shift resources away from a guest that runs less important workload

The processor shares and entitled capacity can be redistributed only between virtual servers that are hosted on the same hypervisor.

PPM Load Balancing weight recommendations

The zEnterprize Platform Performance Manager (PPM) component of the zEnterprise Unified Resource Manager provides a set of application programming interfaces (APIs) that make weight recommendations to dynamically distribute work that is based on weights. You can implement these APIs with a hardware load balancer, for example, an F5[5] switch; or a software load balancer, for example, Co:Z Load Balancer[6]. PPM is not routing work but making the weight recommendations that the switch uses to dynamically distribute the transactions that are proportional to the weight.

To extend the example in Figure 2 on page 4 for guests on two IBM z/VM LPARs, consider that a load balancer switch is routing transactions to four Linux guests that are hosted on two

[5] BIG-IP Local Traffic Manager (LTM), © 2012 F5 Networks, Inc.
[6] Co:Z Load Balancer User's Guide, Chapter 1 Introduction, V 0.1.0 Edition, Published March 2012, Copyright © 2012 Dovetailed Technologies, LLC.

z/VM LPARs: z/VM LPAR1 and z/VM LPAR2, as shown in Figure 7. PPM load balancing makes weight recommendations that are not only based on free processor capacity of a virtual server, but also on free processor capacity of the z/VM LPAR on which the virtual server is running. z/VM logical partitions in this example have identical configurations and have the same number of logical processors. Each z/VM LPAR hosts three Linux guests, all of which have the same configuration in terms of a relative share setting, number of virtual processors, memory, and input/output (I/O). On each z/VM LPAR, two of the three Linux guests run transaction workload that is routed through a hardware load balancer switch. One Linux guest on each z/VM LPAR runs batch workload.

All four virtual servers processing transaction workloads are part of the same load balancing group. The resource consumption patterns of the batch workload running on V3 and V6 might not be identical. If V3 sees increased processor consumption in batch workload processor resource usage at noon, as shown Figure 4 on page 5, guests V4 and V5 get higher PPM-recommended load balancing weight because of more free capacity on z/VM LPAR2.

In Figure 7, V4 and V5 get a higher weight of 72 as compared to V1 and V2, which get a weight of 40.

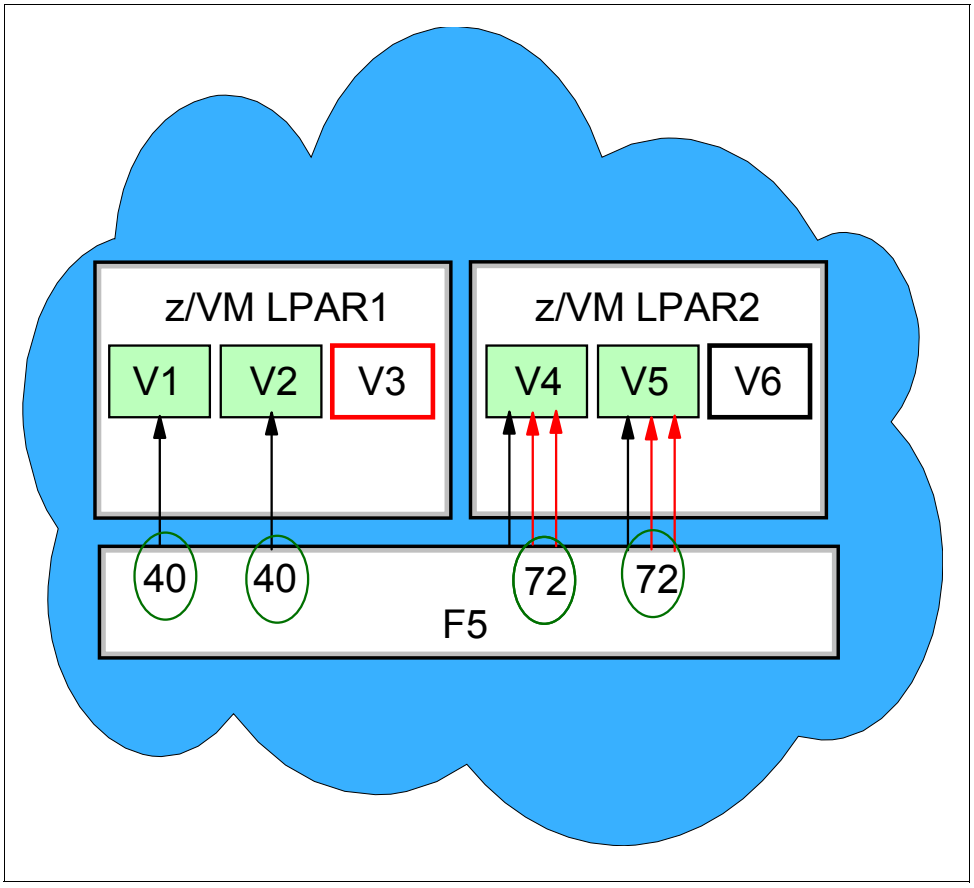

Figure 7 Platform performance manager load balancing weight recommendations

If the batch workload on V6 begins using more processor capacity than V3, more transactions are diverted to guests on z/VM LPAR1. This function helps balance resource consumption across multiple hypervisors. You can use this technique with the PPM Dynamic CPU Resource Management. For example, consider that the goals on the transaction workload cannot be met even after PPM Dynamic CPU Resource Management moves shares from all eligible donor Linux guests to the servers processing high importance service class, transaction workload within a single z/VM LPAR. Then, PPM Load Balancing weight

recommendations divert some transactions to Linux guests on another z/VM LPAR that has free capacity to process more transactions. Both of these approaches react to changes in workload processor consumption in real time.

Correct workload mixture

As you select workloads to be hosted on the cloud, it is important to select those workloads that can be stacked and thus avoid consistently under-utilized and over-utilized systems.

Revisit the example in Figure 2 on page 4. Look at the average processor utilization trend of the two workloads in Figure 8, and notice that the processor utilization is consistently below 50% of total processor capacity, between 6:00 p.m. - 11:00 p.m. Now, imagine a new transaction workload with peak utilization levels lower than 20 - 30% of total processor capacity of the z/VM LPAR. The workload runs only between 6:00 p.m. - 11:00 p.m. At all other times, processor utilization stays below 5%. The usage pattern for all three workloads does not change from day to day, week to week, or month to month.

Therefore, it is an ideal candidate to stack on the existing workloads on this LPAR, as shown in Figure 8.

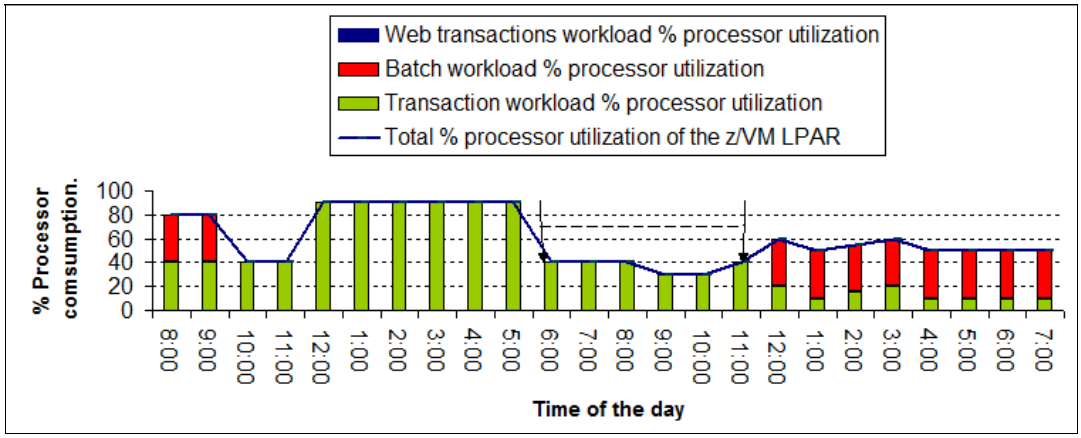

Figure 8 Workloads can be stacked from 6:00 - 11:00 p.m., when average resource consumption is low

Figure 9 shows the workload resource usage pattern after the new workload is stacked.

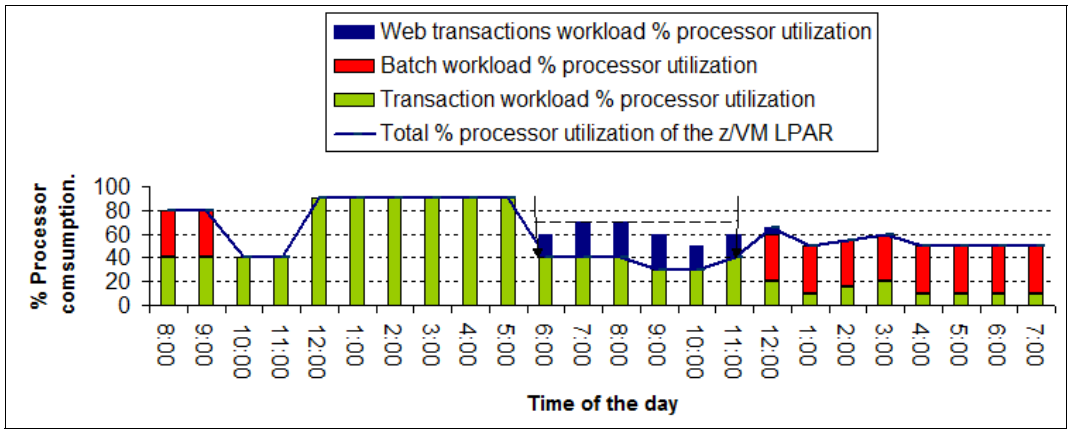

Figure 9 Workload usage pattern after stacking

Before you stack workloads, you have to study processor characteristics in addition to other workload characteristics, for example, memory or I/O. Workloads can be stacked only if they do not attain their peak consumptions at the same time. Stacking such workloads on the cloud ensures that the system has consistently high levels of processor utilization. You can also meet service level agreements and guard against peaks by combining stacking with other techniques, for example, PPM Dynamic CPU Resource Management.

In addition to studying the daily usage pattern of workload consumption, looking at monthly and yearly patterns can identify drops in transaction workloads, for example around holidays. You can use the time frame during drops in usage to run annual batch jobs.

Offloading to a public cloud

You can mine and analyze the data that is gathered from processor usage reports to perform predictive trending. This analysis helps to determine when the increased need for processors exceeds the number of processors that are available on your system. At that point, you can use public clouds to relieve requirements of workloads with less stringent security requirements. Environments that use both private and public clouds are *hybrid clouds*. Public clouds offer added benefits such as high scalability, low initial setup costs, and easy set up and tear down of environments.

Partition mobility

IBM PowerVM Enterprise Edition servers offer LPAR mobility, a function that balances processor resources (not available on zEnterprise servers). LPAR mobility allows you to move the active LPAR and hosted applications to another, more powerful server, without disrupting services or transactions. This technique is useful when the workload processor usage increases to the point that it causes contention of processor resources on a current server. Therefore, this technique is useful to meet the increased workload processor usage of the workloads that caused contention. Alternately, you can transfer LPARs running less important workloads to a new PowerVM server. This process frees system resources on the current server to meet resource demand.

For example, the Wimbledon Tennis Tournament website experiences increased traffic during matches when it posts scores and conducts polls. Workload here, peaks for an expected time frame and then tapers off. The partition mobility function is a good technique to manage such long predictable periods of peak activity.

Conclusion

Consolidating workloads in the cloud environment can result in significant savings if processor resources are sufficient to meet the average utilization levels. However, your environment experiences sporadic and cyclic peaks, and sustained increases or decreases in processor-resource demand. In such cases, you can apply the various techniques described. For example, apply techniques such as PPM Dynamic CPU Resource Management, PPM Load Balancing weight recommendations, correct workload mix, offloading to public clouds, and LPAR mobility. You can also combine techniques for optimal results.

Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this paper.

IBM Redbooks

The following IBM Redbooks publications provide additional information about the topic in this document. Note that some publications referenced in this list might be available in softcopy only:

- *ABCs of IBM z/OS System Programming Volume 10*, SG24-6990-04

  http://www.redbooks.ibm.com/abstracts/sg246990.html?Open
- *IBM PowerVM Live Partition Mobility*, SG24-7460-01

  http://www.redbooks.ibm.com/abstracts/sg247460.html?Open
- *Building an Ensemble Using IBM zEnterprise Unified Resource Manager*, SG24-7921-00

  http://www.redbooks.ibm.com/abstracts/sg247921.html?Open

You can search for, view, or download Redbooks publications, Redpaper publications, Technotes, draft publications and Additional materials, as well as order hardcopy Redbooks publications, at this website:

ibm.com/redbooks

Other publications

These publications are also relevant as further information sources:

- *Building a foundation for a high-performance, low cost private cloud*, The Bathwick Group Ltd, November 2011

  http://www.findwhitepapers.com/content23805
- *Get more out of cloud with a structured workload analysis*, IBM Global Technology Services® Thought Leadership white paper, October 2011

  http://resources.idgenterprise.com/original/AST-0059741_ibm_getmoreoutofcloud.pdf

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

The team who wrote this paper

This paper was produced by a team of specialists from around the world working at the IBM International Technical Support Organization, Poughkeepsie Center.

Deepti Naphade an Advisory Software Engineer at IBM Systems and Technology group; Enterprise Systems and Technology Development, Poughkeepsie, NY, US. Deepti's area of expertise is Algorithm Verification testing and, in her 8 years since joining IBM, she has demonstrated technical leadership in this area across multiple products. She holds a Bachelor degree in Instrumentation and Control Engineering from University of Pune, India.

Richard D. Prewitt is an IBM Distinguished Engineer and Chief Test Engineer in the Systems and Technology Group organization. He is a testing expert at IBM and has devoted his career of more than 28 years to the field of enterprise operating system testing. Rich has provided thought leadership for numerous testing innovations that are focused on operating system resiliency improvements. He is a subject matter expert in the field. Rich holds a Bachelor of Science in Computer Science from Pennsylvania State University.

Peter Yocom is a Senior Technical Staff Member at IBM in the US. He has 26 years of experience in the development organization for IBM System z. He holds an MS degree in computer science from Rensselaer Polytechnic Institute. His primary area of expertise is workload management for the System z environment.

Now you can become a published author, too!

Here's an opportunity to spotlight your skills, grow your career, and become a published author—all at the same time! Join an ITSO residency project and help write a book in your area of expertise, while honing your experience using leading-edge technologies. Your efforts will help to increase product acceptance and customer satisfaction, as you expand your network of technical contacts and relationships. Residencies run from two to six weeks in length. You can participate either in person or as a remote resident working from your home base. Find out more about the residency program and apply online at:

ibm.com/redbooks/residencies.html

Stay connected to IBM Redbooks

- Find us on Facebook:

  http://www.facebook.com/IBMRedbooks
- Follow us on Twitter:

  http://twitter.com/ibmredbooks
- Look for us on LinkedIn:

  http://www.linkedin.com/groups?home=&gid=2130806
- Explore new IBM Redbooks publications, residencies, and workshops with the IBM Redbooks weekly newsletter:

  https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm
- Stay current on recent Redbooks publications with RSS Feeds:

  http://www.redbooks.ibm.com/rss.html

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

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

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs.

This document REDP-4940-00 was created or updated on January 23, 2013.

Send us your comments in one of the following ways:
- Use the online **Contact us** review Redbooks form found at:
  ibm.com/redbooks
- Send your comments in an email to:
  redbooks@us.ibm.com
- Mail your comments to:
  IBM Corporation, International Technical Support Organization
  Dept. HYTD Mail Station P099
  2455 South Road
  Poughkeepsie, NY 12601-5400 U.S.A.

Trademarks