Zürich, Switzerland

techforge llc

MSU / MIPS savings whitepaper

Author: Christian Wehrli
Version: 1.0
Date: 07.11.2023

Document Summary:

This document is intended for mainframe customers who are looking to gain an overview on how to create strategies to achieve mainframe capacity reductions.

Table of Contents

Overview

Introduction to the Series
This whitepaper is the first in a series aimed at guiding IT professionals on how to minimize costly mainframe workloads, conserve valuable compute capacity, and understand mainframe pricing models. It introduces essential concepts, strategies, and practical insights for managing and optimizing mainframe operations efficiently.

Understanding Capacity and Pricing Models
The first and second parts delve into the nuances of the mainframe pricing model, particularly the 4-hour rolling average (4HRA) and the new tailored fit pricing models, which include "consumption-based" and "capacity-based" options. These sections explain how costs are calculated based on peak consumption and highlight how understanding these models can lead to significant cost reductions.

Analyzing and Managing Workload Footprints
The series discusses the identification of heavy workload footprints and offers a comprehensive overview of mainframe capacity measurement. It emphasizes the importance of extracting and interpreting data from z/OS systems to make it actionable and user-friendly.

Strategies for Reducing Mainframe Workload
In part three, the focus shifts to effective strategies for reducing the identified workload footprint. It explores various approaches such as redesigning, refactoring, rehosting, migrating to different products, or rescheduling applications. This part is designed to help IT professionals choose the best strategy for their specific needs to optimize performance and reduce costs.

Practical Implementation and Conclusion
Each part of the series provides practical strategies for analyzing extracted data, identifying optimization opportunities, and implementing changes. By applying these techniques, businesses can not only reduce costs but also improve the efficiency of their mainframe environments.

Analyzing Workload Footprints

Where to start

To effectively reduce MIPS/MSU consumption, it's crucial to align reduction efforts with both the mainframe's pricing model and its capacity limits. The first step is to identify the areas where consumption is highest—understanding where and how resources are being used allows for targeted interventions. From there, the approach involves a combination of optimizing existing processes and potentially restructuring or replacing resource-intensive applications.

MIPS / MSU

MIPS stands for "Million Instructions Per Second," a benchmark used to measure the processing speed of a computer, specifically on z/OS mainframes. Each type of mainframe processor has a different maximum MIPS capacity, which varies depending on the machine and the number of processors activated. Meanwhile, MSUs, or "Million Service Units," are another metric used in the industry. Sometimes humorously referred to as "mystery service units," the exact calculation for MSUs can seem opaque, though it is understood that the ratio between MIPS and MSUs is approximately 1:8, varying with the machine type.

SMF and RMF Reporting

Introduction to SMF and RMF Reporting

Obtaining accurate consumption metrics is critical for optimizing mainframe operations. System Management Facilities (SMF) and Resource Measurement Facility (RMF) are essential tools in this regard. SMF collects detailed records, such as CICS 115, 116, AS 30, DB2 100, 101, 102, for each address space, SQL statement, DLI call, and CICS/IMS transaction, helping establish a comprehensive consumption footprint. RMF supplements this by capturing most of the workload activity across a mainframe's Logical Partition (LPAR).

Utilizing SMF Records

SMF is equipped with utilities that facilitate the dumping of records. However, making these records readable often requires additional formatting tools such as SAS, MICS, or MXG. The process of formatting and the specifics of making SMF records printable are well-documented, although it can be complex to interpret these formats directly. An alternative approach could involve using RMF's capacity reporting features, which might already provide the necessary data in a more accessible format.

You can find more information in IBM’s online documentation on how to deal with SMF records:

Dumping SMF Records

Key Metrics to Monitor

For effective management and optimization, it is crucial to identify:

High-Consumption Programs and Utilities: Determining which batch programs and utilities consume the most resources.
Impactful Transactions and Statements: Identifying which transactions and SQL/DLI statements lead to the highest consumption.
Timing and Frequency of Consumption: Understanding when the highest consumption occurs and how frequently these peaks happen.

Identifying Factors Contributing to Workload Peaks

Pinpointing High-Consumption Elements
To formulate an effective MIPS/MSU reduction strategy, it's crucial to understand which components contribute most significantly to your workload peak. This involves analyzing specific address spaces, task control blocks, transactions, and SQL/DLI statements that drive high consumption rates.

Common Causes of High MIPS Usage
Several factors can lead to increased MIPS usage, which, in turn, drives up costs:

- Sub-Optimal Programming: Inefficient code can consume excessive CPU time, significantly increasing MIPS usage. Regular code reviews and optimization can help mitigate this issue.

- Sub-Optimal Data and Database Usage: Poor database management and inefficient data handling also contribute to high MIPS consumption. Optimizing data access patterns and refining database indexing can lead to more efficient resource use.

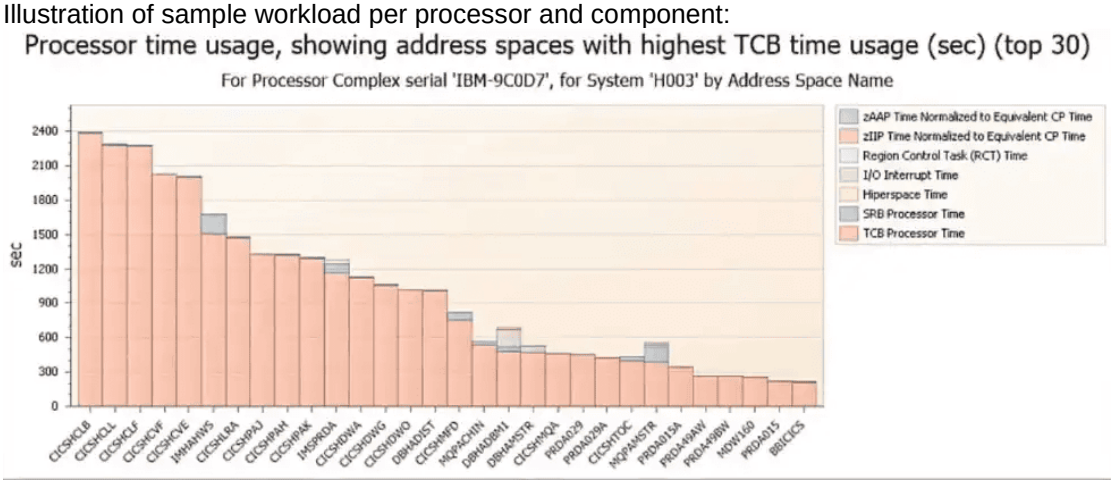
- Not Shuffling Workload into Off-Peak Hours: Running intensive tasks during peak hours can cause spikes in MIPS usage. Scheduling heavy jobs for off-peak times can help distribute load more evenly and reduce costs.

- Not Leveraging Cheaper Workload Processors or Runtimes: Utilizing specialized processors, such as zIIPs (z Integrated Information Processors), for eligible workloads can offload tasks from the main processor, thus reducing MIPS usage and associated costs.

Implementing MIPS/MSU Reduction Strategies

Understanding these factors allows IT managers to implement targeted strategies to reduce MIPS/MSU consumption effectively. Addressing each of these areas not only helps in reducing the immediate financial impact of high MIPS/MSU usage but also improves overall system performance and efficiency. This holistic approach ensures that mainframe resources are managed optimally, leading to sustained operational excellence.

Illustration of sample workload per processor and component:



Introduction to Mainframe Pricing Models

In this installment of the series, we delve into the mainframe pricing model and explore how a deep understanding of these models can aid in reducing costs. The typical cost calculation for mainframe usage is based on the four-hour rolling average (4HRA), where the highest peak consumption within any given four-hour period determines the monthly charges.

Detailed Examination of 4HRA and Tailored Fit Pricing

The traditional 4HRA model continuously tracks the MSU consumption for each Logical Partition (LPAR) over the course of a month. Charges are then calculated based on the peak usage during the most intensive four-hour period, as recorded by specific System Management Facility (SMF) records.

In response to the limitations of the 4HRA model, IBM introduced the "Tailored Fit Pricing" model, which includes two options: Enterprise Consumption and Enterprise Capacity. These models offer a more flexible approach:

- Enterprise Capacity: This model charges based on the maximum capacity of the machine, simplifying billing to a consistent monthly fee regardless of fluctuating usage.
- Enterprise Consumption: This model sets a baseline from the previous year's usage, allowing organizations to pay primarily for what they consume. Any underutilization against the baseline allows for MSUs to be rolled over to future periods, and any growth in consumption above the baseline is charged at a reduced rate.

Importance of Software and Consumption Metrics

Most of IBM’s system software, such as z/OS, Db2, IMS, and others, are charged under a monthly license charge (MLC). The 4HRA for each product is calculated using instantaneous MSU usage data collected every five minutes, ensuring that billing is reflective of actual usage.

Reporting and Usage Tracking

Usage metrics are meticulously collected and reported through the Sub Capacity Reporting Tool (SCRT), which utilizes SMF 70-1 and SMF 89 records. This data collection is vital as it underpins the billing process, ensuring organizations are billed fairly based on actual usage.

Strategic Implications for Cost Reduction

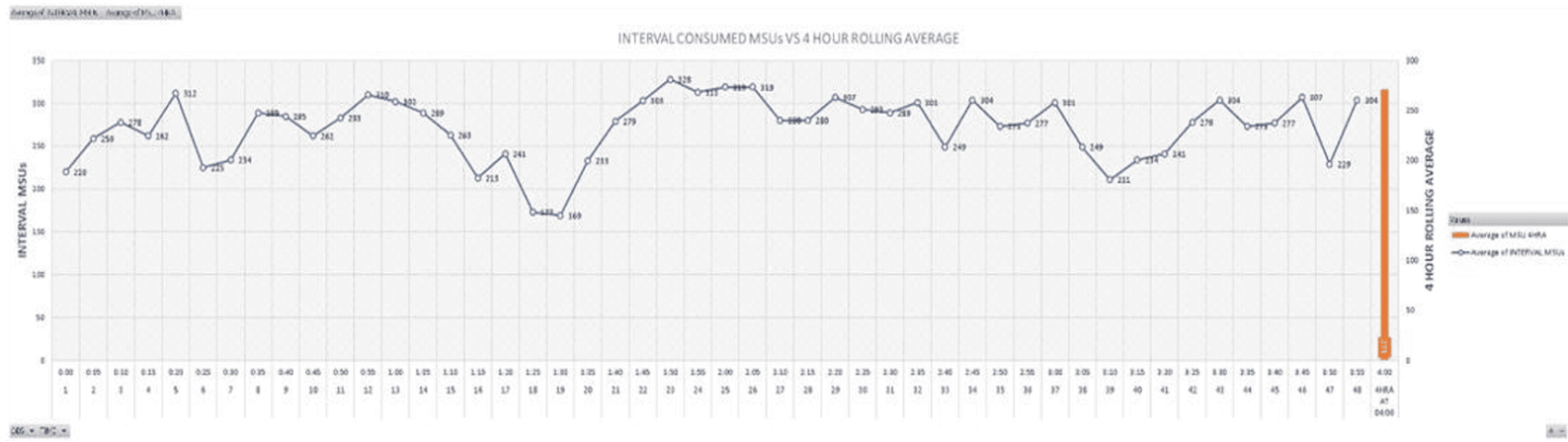
By understanding and managing the times when high consumption occurs, organizations can significantly reduce their peak usage and, consequently, their costs. Strategies such as shifting heavy workloads to off-peak hours or optimizing software and data usage can have a profound impact on reducing the 4HRA.

A detailed description of the pricing models can be found on those two links:

4HRA Model
Tailored Fit Pricing

Managing mainframe costs effectively requires not only understanding the intricate details of these pricing models but also actively managing and optimizing workloads. Reducing peak periods and smoothing out consumption can lead to substantial cost savings. This series aims to equip IT professionals with the knowledge to make informed decisions about their mainframe operations to achieve optimal financial and operational results.

Illustration of an example interval curve for calculating the 4HRA



Workload reduction strategies

In the concluding part of this series, we focus on practical strategies to reduce the MSU/MIPS footprint. After discussing the intricacies of pricing models and identifying high-consumption areas in previous parts, we now explore actionable steps to mitigate these workloads effectively.

Strategic Options for Workload Management

Rescheduling Workload:

One effective strategy is to shift non-critical workloads to off-peak hours. This can be achieved through transaction and batch scheduling. Additionally, prioritizing critical workloads during peak periods and configuring Workload Management (WLM) settings to be more restrictive can prevent non-essential processes from inflating peak usage.

Code Optimization:

Analyzing transactions, batch jobs, and database queries that consume significant capacity allows for targeted code improvements. Options include:

- Leveraging more efficient coding functions in languages like C370, PL/I, and COBOL.
- Recompiling code with updated compilers, such as COBOL v6.4, to take advantage of advanced, efficient instructions.
- Eliminating redundant or inefficient code, particularly costly instructions identified within your codebase.

Database Optimization:

Heavy database activities often contribute substantially to workload peaks. Utilizing tools like "DB2 explain" can provide insights into the cost of SQL statements and guide optimizations. Revisiting and potentially normalizing the database design can mitigate expensive queries, particularly those run frequently or those critical for business intelligence and data warehousing.

Leveraging Cheaper Processors:

IBM’s zIIP and zAAP processors offer cost-effective alternatives for specific types of workloads. These specialized processors are typically less expensive and, in some cases, free to use, though they come with certain limitations. Workloads suitable for these processors include container workloads and programs running in SRB mode, such as Java applications.

On this link you can find out nore how to make your applications zIIP eligable

Rehosting and

Independent Software Vendor (ISV) Replacements:

Migrating certain workloads to more cost-efficient platforms can also yield significant savings. Platforms like MicroFocus Enterprise Server or LzLabs’ Software Defined Mainframe (SDM) allow for the decentralization of processing, reducing reliance on traditional mainframe resources. Additionally, relocating non-essential services like message queuing or job scheduling to decentralized platforms can further optimize costs.

Conclusion

Hopefully, this document has equipped you with a robust understanding of MSU/MIPS pricing models and detailed strategies for reducing your mainframe's workload footprint. By implementing these targeted strategies, you can significantly reduce costs while maintaining or enhancing system performance and reliability.

Headquarters:
techforge llc
Juchstrase 9
CH-8173 Neerach
Switzerland
+41 76 437 54 53
web: https://techforge.llc

Email Us at:
info@techforge.llc