Lower Mainframe Costs by up to 20% with Intelligent Capping

Balance cost with performance using MainView Intelligent Capping
# Table of Contents

1  EXECUTIVE SUMMARY  

2  INTRODUCTION  

   SOFTWARE COSTS AND THE MAINFRAME  

   THE FEAR, UNCERTAINTY, AND DOUBT  

3  WHY WORKLOAD CAPPING CAN BE SAFER AND EASIER THAN YOU THINK  

   THE SOLUTION: INTELLIGENT CAPPING  

4  HOW WILL YOU USE INTELLIGENT CAPPING?  

5  SUMMARY
Executive Summary

Optimizing your mainframe software environment, including costs, is a long-term proposition for which most companies running mainframes have been highly successful. The challenge is optimizing monthly license charge (MLC) costs that result from increasing transactions and activity. To complicate the matter, IBM increases the cost of MLC software each year by 4 to 7 percent.

If you run MainView systems management solutions, you have a unique advantage that enables your company to implement easy and effective workload capping to reduce peak MSU usage and lower MLC costs.

This white paper explains the benefits of leveraging MainView capping capabilities in an intelligent, dynamic manner using the Intelligent Capping solution.
INTRODUCTION

Does your mainframe organization currently use workload-defined capacity capping as a key strategy for lowering your mainframe software costs? If not, your company could be spending more than necessary. Workload capping can be the single biggest way to lower your IBM monthly license charges (MLC).

The purpose of this white paper is to help remove some of the roadblocks to your workload capping efforts. The authors and contributors to this paper have conducted extensive research into capping best practices from which you can benefit.

SOFTWARE COSTS AND THE MAINFRAME

The cost of the mainframe is a key focus for most large companies these days. Mainframe user surveys, such as the BMC Annual Mainframe Research Survey, indicate that cost optimization is the number one priority each year. IBM MLC costs are rising at a high rate, as much as 4% to 7 percent per year on average. And, 58 percent of respondents say they spend more than 30 percent of their company’s mainframe IT budget on MLC software costs. IT teams have implemented many cost reduction strategies to impact the 30 percent, including workload capping. Some have been successful while others have not. When it comes to capping, there’s a fine line between saving costs and risking workload interruption. You must strike the balance between cost and performance.

THE FEAR, UNCERTAINTY, AND DOUBT

Many organizations are facing a critical decision when it comes to workload capping: Do we invest in a comprehensive capping strategy? Or do we just do a minimum job of capping to avoid the impact of unusual spikes, while ensuring we don’t impact the business? When it comes to workload capping, there is a lot of fear, uncertainty, and doubt:

- Fear – Some people fear capping will take down their business if critical workloads are impacted.
- Uncertainty – Some people have attempted capping with poor results and are uncertain if they want to try it again.
- Doubt – Some people doubt the benefits and don’t have the time to further understand the true benefits of capping, much less the time required to design and implement the strategy.

What have been your experiences with workload capping? Ask yourself whether it’s been effective. Have you seen only marginal returns? Has it been effective but labor-intensive? Has it been catastrophic (hopefully not)?

Capping saves money. So why don’t all people cap? The implementation of capping is a difficult balancing act. A very fine line exists between performance of the workload and the cost of the resources. Workload capping can control the cost of the resources consumed, but it can also cause the performance of the workload to suffer. Questions must be answered, such as:

- What should I cap?
- When should I cap?
- How will the capping affect the performance of my workload? Or my MLC charges?

So how do we address these concerns? You want to react instantly to workload volume and volatility changes. You need some sort of automation to do it for you to ensure MLC is lowered while looking across all logical partitions (LPARs) and groups of LPARs so that your critical workloads are not impacted. While it is recognized that capping can cause the performance of the workload to suffer, an intelligent capping solution can mitigate this risk.
WHY WORKLOAD CAPPING CAN BE SAFER AND EASIER THAN YOU THINK

Workload capping can be safer and more effective than you think if you are utilizing an intelligent, automatic, and dynamic solution. You may not be aware that you have whitespace (unused capacity) available in your environment. You can take advantage of this unused capacity to avoid capping high importance workloads. There is no risk to the SLAs in your environment, as this capacity was not being used, and the increase in capacity where it is needed is offset by the decrease in capacity where it is not. In this case, you have prevented capping of critical work by utilizing unused capacity while also minimizing your MLC costs.

You might have low importance work contributing to your peak and driving up your costs. What if you can intelligently cap and delay that low importance work and give capacity to another LPAR that is struggling for the resources it needs to run its high importance work, automatically and without increasing your costs?

To maximize the safety in your capping strategy, you’ll want to use a mechanism that caps only low importance workloads, while ensuring high importance workloads, especially business-critical workloads, are never capped.

As a best-case capping scenario, you should be able to use capping to reduce your MLC costs without impacting your SLAs and impacting any workload, even low priority workloads. If capping is necessary, only low importance work should be delayed.

As part of IBM® z/OS®, IBM provides the ability to manage groups of LPARs with Workload Manager (WLM) Capacity Groups, which allows LPARs to share MSU, but doesn’t make dynamic changes based on the importance of the work running on the LPARs.

Ideally, you would use a solution that is constantly workload-aware and makes adjustments, thus alleviating the worry and manual fine tuning, and is built and supported by a company you already trust to manage your mainframe environment.

THE SOLUTION: INTELLIGENT CAPPING

Intelligent Capping (iCap) is a policy-based solution. You can define the policy directly in the product, and it requires only a few parameters to get started, including the LPARs to manage and how many total MSUs you want to use for those LPARs. iCap does the rest. It monitors the LPARs and makes intelligent decisions about which LPARs need MSUs to get their important work done. It’s all about ensuring the MSUs are in the right place at the right time.

iCap makes decisions based on:

- Current usage
- MSU limits you configure
- Workload importance and relative LPAR importance
- Optionally, you can factor in differences in MLC cost per MSU on LPARs to provide cost-aware capping

iCap is all about maximizing MLC savings and reducing risk. It has built-in MainView technology to alert you if any capping is occurring, and will even let you know if it’s just low importance work. When high importance work is about to be capped, you can use the built-in automation to raise the overall MSU limit or take other action. This process results in lower MLC costs overall.

Today, if you are using MainView to monitor and manage your critical systems, you know that MainView is a robust and highly efficient solution that you can depend on when you need it most. Now, you can cap MSU peak resource usage safely and effectively, at your own pace, without risking your SLAs.

With (iCap) and MainView, you have:

- **MainView Explorer** – The modern user interface for both iCap and MainView removes the need to deploy end user interfaces, provides better graphics, provides containers to allow you to group information from iCap as well as MainView for z/OS, and is easier to use for operations that do not have an IBM 3270 background.
- **Alarm Manager** – iCap ships with sample alarms that can alert you when system resources are overused. You use MainView Alarm Manager to set these alarms or to create your own alarms. The alarms can also be part of a MainView Explorer container. The alarms can be used by automation to take action and/or for notification/escalation.
- **Action** – MainView gives you the ability to issue z/OS commands in response to cap alerts or alarms. If you also have MainView AutoOPERATOR, you can have even more extensive automated responses to capping issues. The goal is to react quickly to mitigate risk.
- **Single View and Control** – This gives you the ability to see and control all LPARs from one view or the same resources across LPARs in one view.
You can use Intelligent Capping to help you decide what and when to cap. You can use it to strike the balance between reducing MLC costs and ensuring that critical workloads are not delayed. After all, your most important priority is to avoid any deterioration in the performance of your highest importance workloads. Using the facilities that you’re already familiar with in MainView, Intelligent Capping can help you lower your peak MSU usage, in a staged approach and at your own pace.

**FIGURE 1:** Typical capping just sets a standard limit, which is usually set too high to effectively save any MSU cost.

**FIGURE 2:** MainView Intelligent Capping provides dynamic adjustments to your MSU limits, helping you save money without adding risk.

**HOW WILL YOU USE INTELLIGENT CAPPING?**

You can use MainView Intelligent Capping to achieve several goals, depending on your organization’s priorities. Here are some ways to make excellent use of the solution:

- Save money on a monthly basis (lower MLC costs)
- Ensure business-critical performance (never delay level-1 importance workload)
- Improve manual processes (react to workload changes automatically and dynamically)
- Better support the volatility of digital engagement by your customers
- Improve capacity utilization (maybe even delay capacity upgrades as a result)
- Gain more control over monitoring and management of workloads with MainView alerts and automation
SUMMARY

This paper has provided you with some insight into what dynamic, intelligent capping can do for your organization. By fully exploiting workload capping, your company can save as much as 5 to 30 percent¹ more than companies that aren’t actively managing their MLC costs. Make the most of iCap for MainView and deliver substantial cost savings to benefit your organization.

¹ David Wilson, SZS Consulting - “10 Steps to Reducing MLC Costs.”

FOR MORE INFORMATION

To learn more about MainView Intelligent Capping, please visit bmc.com/it-solutions/mainview

BMC is a global leader in innovative software solutions that enable businesses to transform into digital enterprises for the ultimate competitive advantage. Our Digital Enterprise Management solutions are designed to fast track digital business from mainframe to mobile to cloud and beyond.

BMC digital IT transforms 82 percent of the Fortune 500.