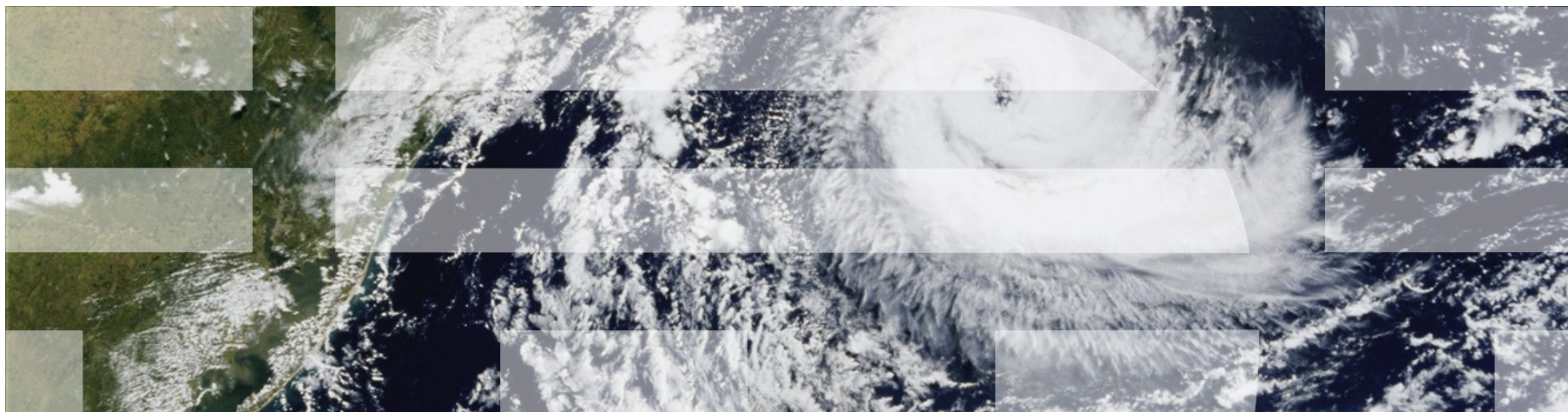


z/OS Capacity Provisioning Overview



Michael Groetzner, *z/OS Capacity Management and Support*

IBM Germany Research & Development

Email: Michael.Groetzner@de.ibm.com

Trademarks

The following are trademarks of the International Business Machines Corporation in the United States and/or other countries.

APPN*	HiperSockets	OS/390*	VM/ESA*
CICS*	HyperSwap	Parallel Sysplex*	VSE/ESA
DB2*	IBM*	PR/SM	VTAM*
DB2 Connect	IBM eServer	Processor Resource/Systems Manager	WebSphere*
DirMaint	IBM e(logo)server*	RACF*	z/Architecture
e-business logo*	IBM logo*	Resource Link	z/OS*
ECKD	IMS	RMF	z/VM*
Enterprise Storage Server*		S/390*	z/VSE
ESCON*	Language Environment*	Sysplex Timer*	zSeries*
FICON*	MQSeries*	System z	z10
GDPS*	Multiprise*	System z9	Z9
Geographically Dispersed Parallel Sysplex	NetView*	System z10	zEnterprise
* Registered trademarks of IBM Corporation	On demand business logo	TotalStorage*	

The following are trademarks or registered trademarks of other companies.

Java and all Java-related trademarks and logos are trademarks of Sun Microsystems, Inc., in the United States and other countries

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Microsoft, Windows and Windows XP are registered trademarks of Microsoft Corporation.

Red Hat, the Red Hat "Shadow Man" logo, and all Red Hat-based trademarks and logos are trademarks or registered trademarks of Red Hat, Inc., in the United States and other countries.

SET and Secure Electronic Transaction are trademarks owned by SET Secure Electronic Transaction LLC.

* All other products may be trademarks or registered trademarks of their respective companies.

Notes:

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput 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 improvements equivalent to the performance ratios stated here.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.

All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the 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.

Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.



**ENABLING BUSINESS.
A THROUGH Z.**

Agenda

- *Environment*
- *Capabilities*
- *Management Domain*
- *Policy*



Today's challenges to manage capacity

- Unexpected events and workload spikes can afford higher processing capacity
- Manual capacity management can be time-consuming and error prone
- Capacity provisioning decisions must be made without sound data



Failures,
Workload moves

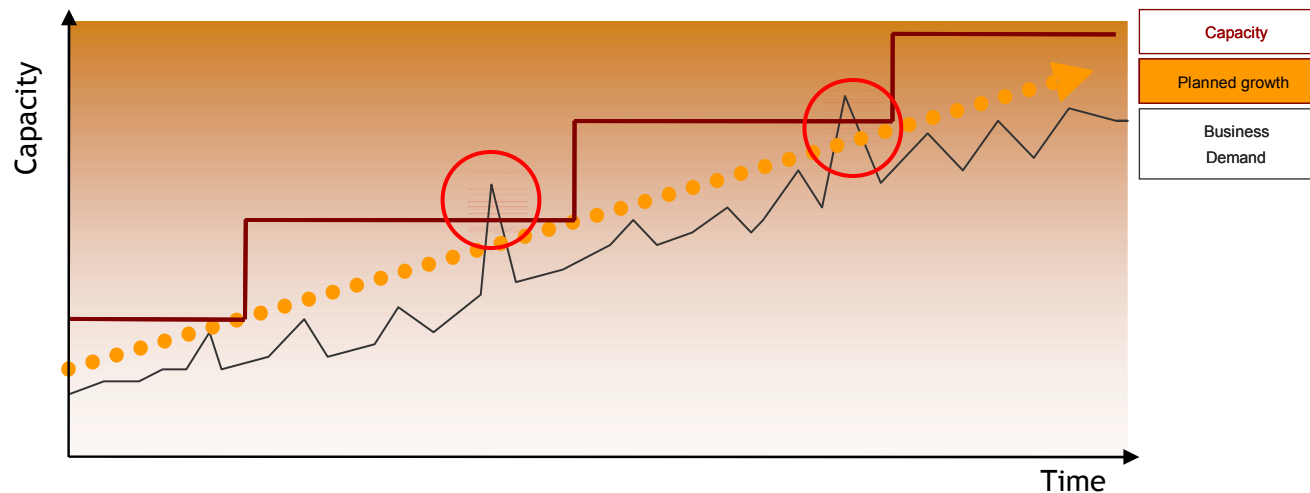
Unexpected
Workload
spikes



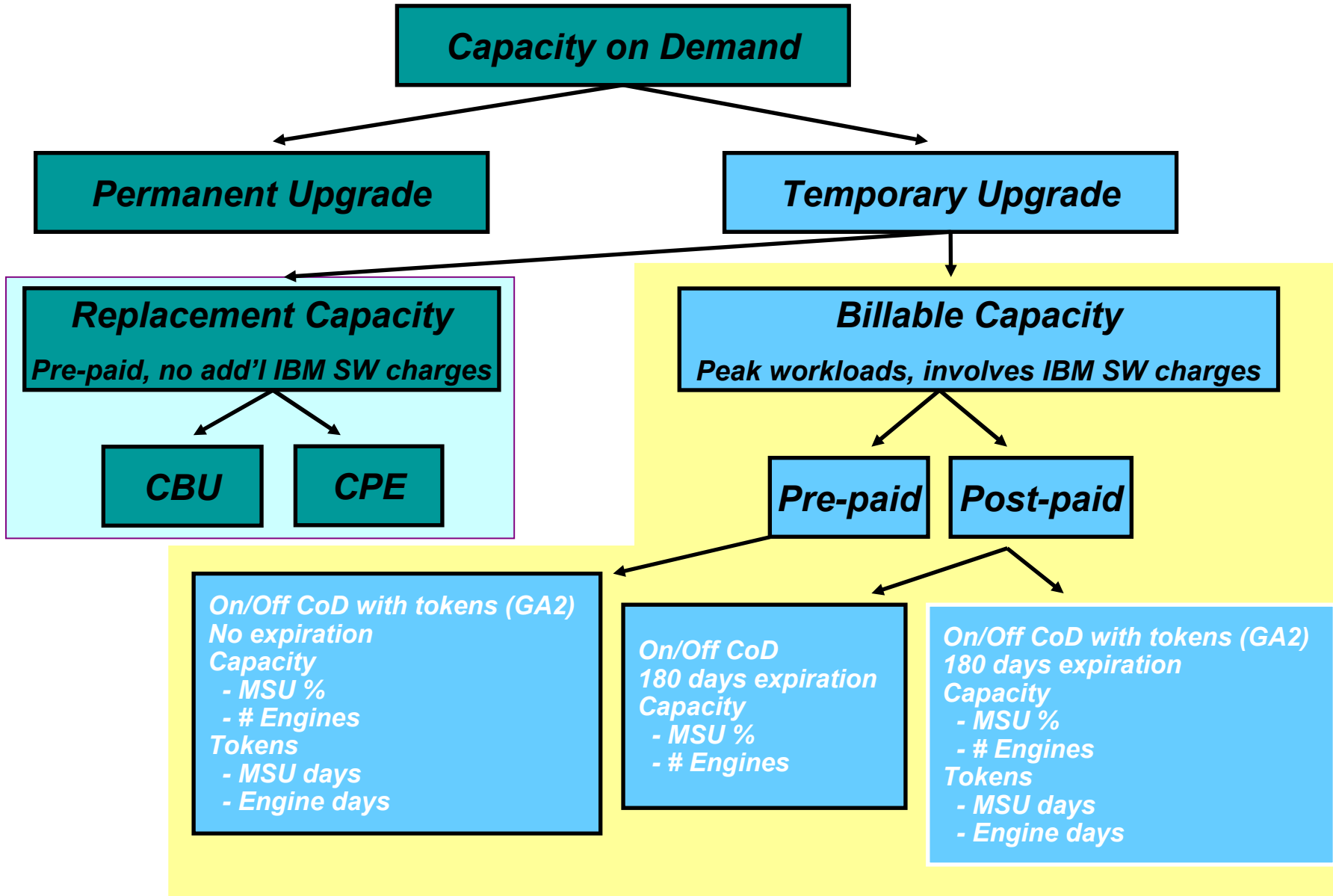
IBM z/OS Capacity Provisioning Basics



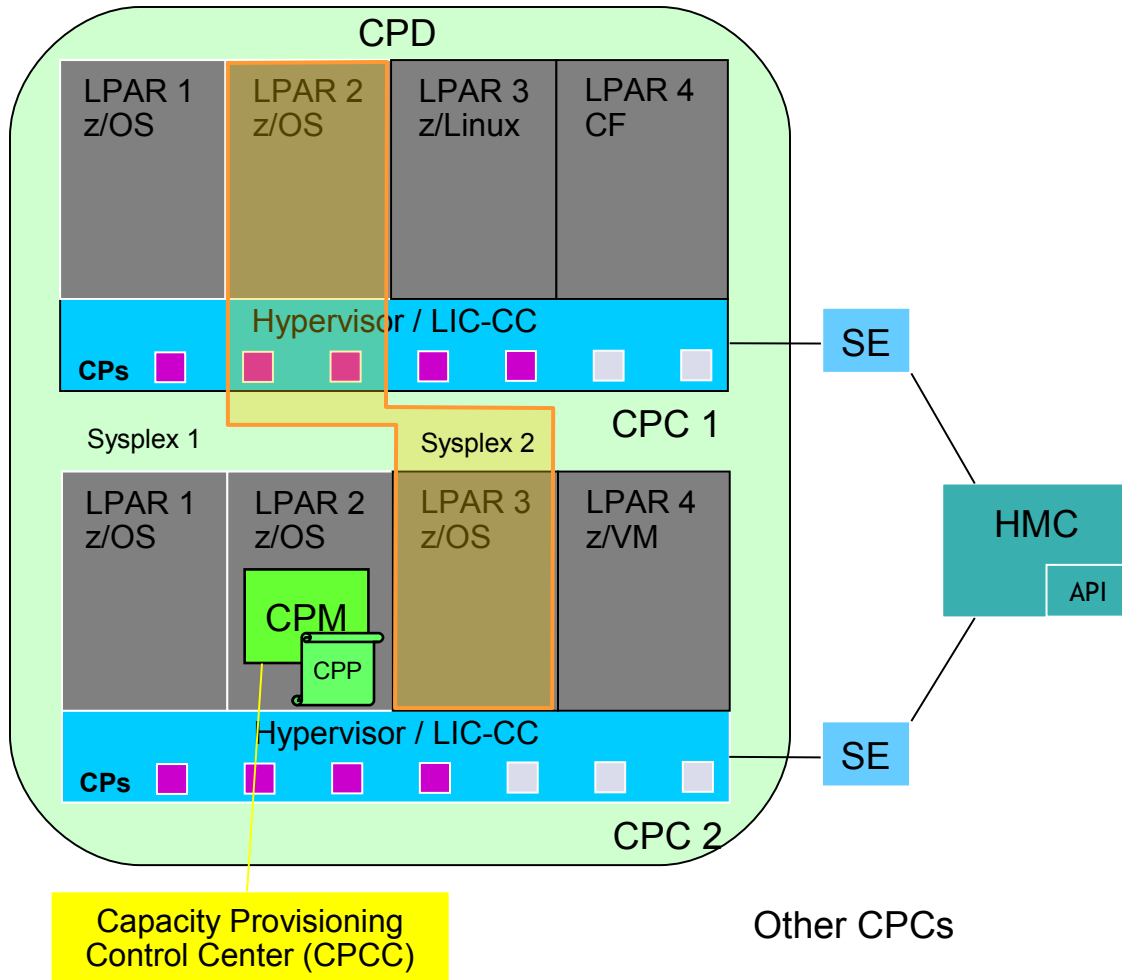
- Contained in z/OS base component free of charge
- Based on System z On/Off Capacity on Demand Feature (zEnterprise 196 or System z10)



z10 – Basics of CoD



Domain Configuration



- Domain configuration defines the CPCs and z/OS systems that are controlled by an instance of the CPM
- One or more CPCs, sysplexes and z/OS systems can be defined into a domain
- Sysplexes and CPCs do not have to be completely contained in a domain but must not belong to more than one Capacity Provisioning domain
- One active Capacity Provisioning policy per domain
- Multiple Sysplexes and hence multiple WLM service definitions may be involved

Manual capacity upgrades – How it could look like

1.	Workload increases	0 min
2.	Operator realizes bottleneck	5-10 min
3.	Operator informs system programmers and manager	2 min
4.	Discussion	10 min
5.	Logon to HMC, activate record	5 min

... meanwhile, so much workload may have queued up that one additional processor would be insufficient to decrease the queued workload

→ Two processors have to be added

CPM can react faster and reduce cost

Capacity Provisioning Capabilities Overview

- The Capacity Provisioning Manager (CPM) can control temporary processor resources on IBM z196 or z10
 - Number of zAAPs or zIIPs
 - General purpose capacity
 - Considers different capacity levels (i.e. effective processor speeds) for subcapacity processors (general purpose capacity)
 - Can advise on logical processors
 - Can control one or more IBM zEnterprise or System z10 servers
 - Including multiple Sysplexes

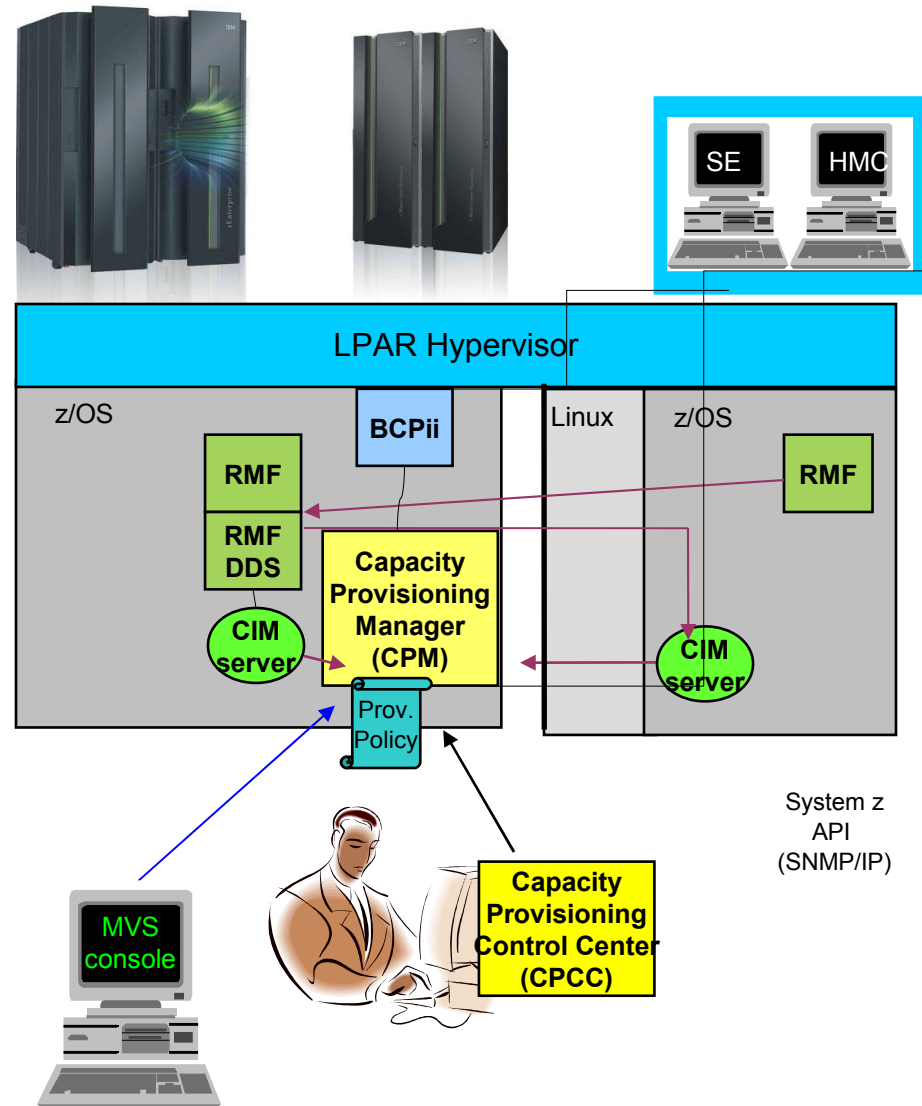


CPM differentiates between different types of provisioning requests:

- Manually at the z/OS console through Capacity Provisioning Manager commands
- Via user defined policy at specified schedules
- Via user defined policy by observing workload performance on z/OS

Capacity Provisioning – Infrastructure in a Nutshell

- z/OS WLM manages workloads to goals and business importance
- WLM indicators available through monitoring component
 - E.g. z/OS Resource Measurement Facility (RMF)
 - One RMF gatherer per z/OS system
 - RMF Distributed Data Server (DDS) per Sysplex
- Capacity Provisioning Manager (CPM) retrieves critical metrics through CIM
- CPM communicates to support elements or HMC, via
 - System z API (SNMP via IP)
 - BCPii
- Capacity Provisioning Control Center is front end to administer Capacity Provisioning policies



Policy Approach

The Capacity Provisioning policy defines the circumstances under which additional capacity may be provisioned:

- Three “dimensions” of criteria considered:
 - **When** is provisioning allowed
 - **Which** work qualifies for provisioning
 - **How much** additional capacity may be activated
- These criteria are specified as “rules” in the policy:

```
If
{ in the specified time interval
  the specified work “suffers”
}
Then up to
{ - the defined additional capacity
  may be activated
}
```

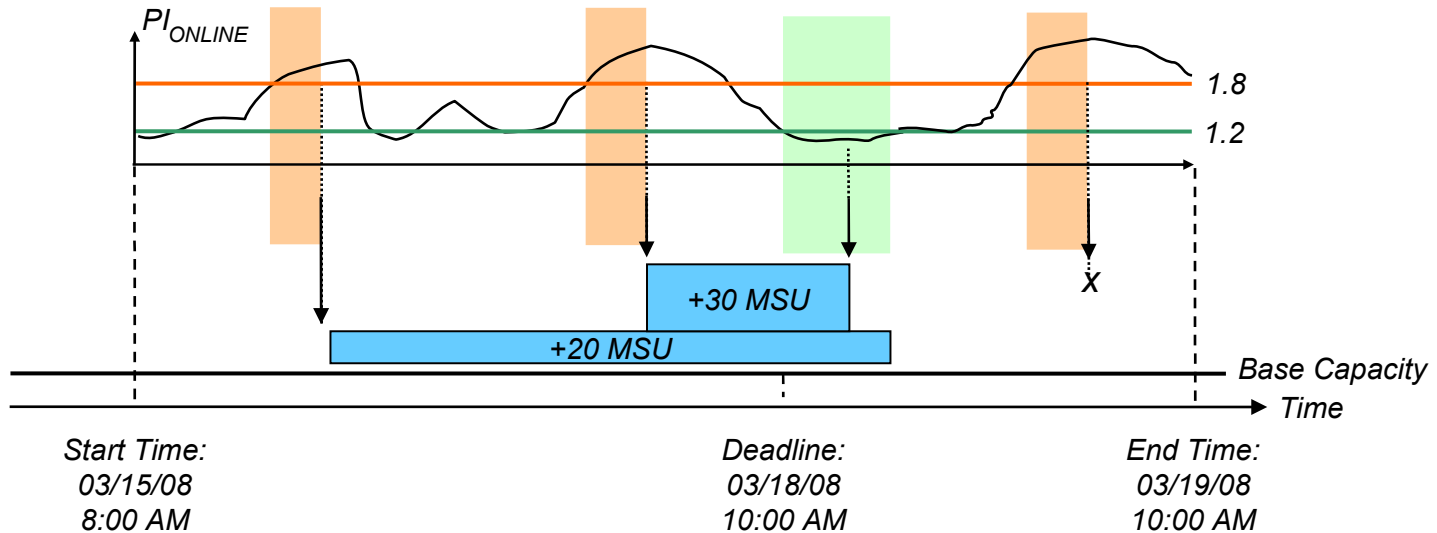
- The specified rules and conditions are named and may be activated or deactivated selectively by operator commands

Sample Workload Condition

Sample definition:

Name: PT1
Sysplex: PLEX1
System: SYSA
Included Service Class Periods:
 ONLINE in WLMSSD with $PI \geq 1.8$ for 10 min until $PI \leq 1.2$ for 10 min
Excluded Service Class Periods:
 BACKUP in WLMSSD

Monitor Service Class PI's:



Capacity Analysis

- If workload goal is not achieved
 - What's the reason (Processor, other)
 - Is there an overall bottleneck
 - What kind of resources will help
 - Can the workload benefit from additional resources

Processing Modes

Capacity Provisioning Manager can operate in one of four modes that allow for different degrees of automation

- **Manual mode**
 - Server capacities can be controlled via CPM commands
 - Command driven mode where no CPM policy is active
- **Analysis mode**
 - CPM processes capacity provisioning policy and informs the operator when a provisioning / deprovisioning action would be due according to criteria specified in the policy.
 - It is up to the operator either to ignore that information or to perform the up-/downgrade manually (using the HMC/SE or the available CPM commands)
- **Confirmation mode**
 - CPM processes the policy and interrogates the On/Off CoD record to be used for capacity provisioning.
Every provisioning action needs to be authorized (confirmed) by the operator.
- **Autonomic mode**
 - Similar to the confirmation mode, except that no human (operator) intervention is required.

Various reports are available with information about workload and provisioning status, and the rationale for provisioning recommendations

Questions?

