

IBM z15 System Recovery Boost (SRB)

Heinz Tschumi
htsi@ch.ibm.com
+41 79 323 0515



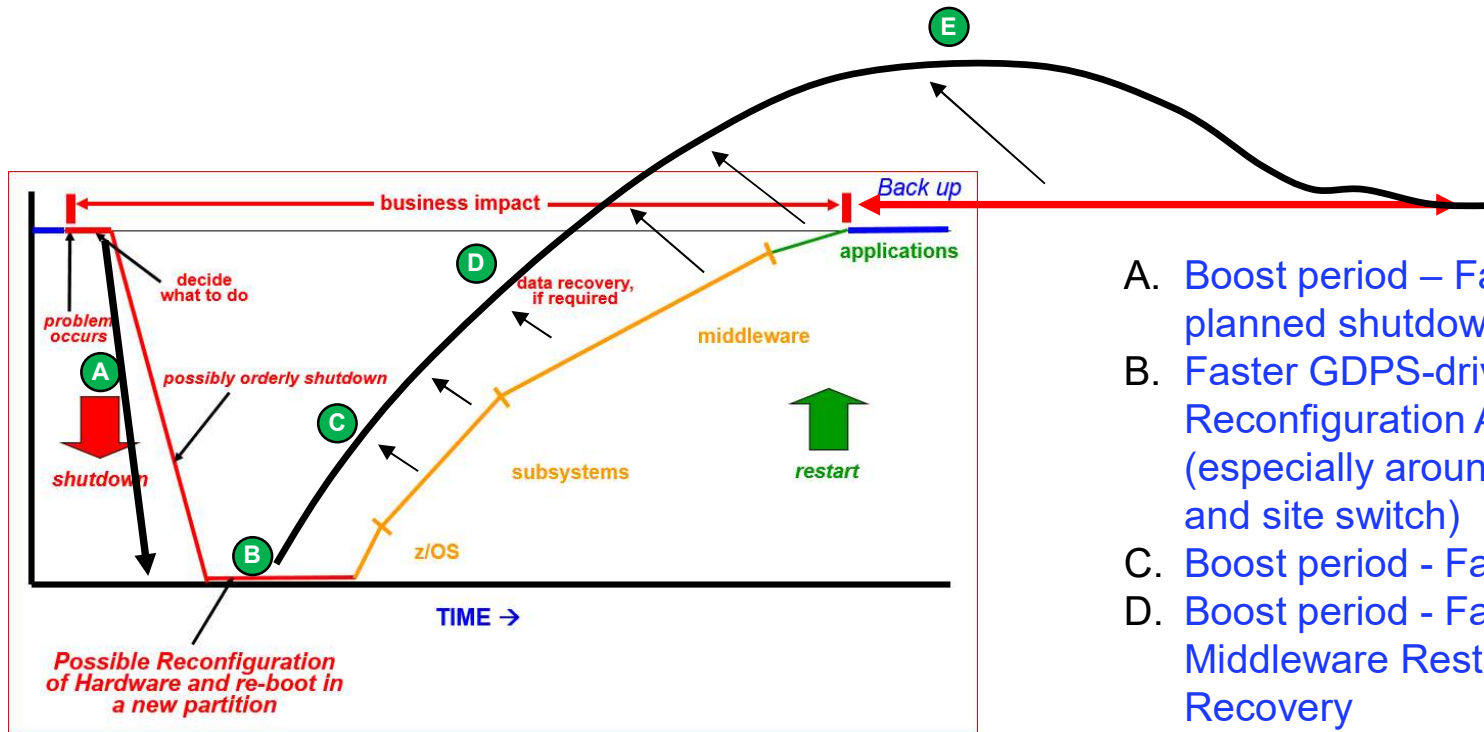
Agenda

- **Overview**
- **zIIP Capacity boost**
- **Speed boost (sub-capacity engines)**
- **GDPS enhancements**
- **System Recovery Boost – Stage 2**

Overview

- Provides **additional capacity** for **shutting down images**, to prepare for planned shutdown activities, substantially faster than any prior Z machine
- Drives faster and more efficient **GDPS** automation actions to reconfigure and recover your environment more quickly
- Provides **additional capacity** for **restarting and recovering images, middleware environments**, and **client workloads** substantially faster than any prior Z machine, to get your systems back to the “**steady state**”
- **Delivers higher processor capacity for a limited time following an IPL**, during a “**Boost period**,” so client workloads can catch up and work through a backlog after downtime
- **Deliver Boost for Recovery Events** (Stage 2)

Shutdown/Startup/Recovery Boost



- A. Boost period – Faster planned shutdown
- B. Faster GDPS-driven HW Reconfiguration Activities (especially around DR and site switch)
- C. Boost period - Faster IPL
- D. Boost period - Faster Middleware Restart and Recovery
- E. Boost period - Boosted Capacity to do work following IPL

No Increase in Client's IBM Software Costs!

Overview

- **System recovery boost** offers the customer **additional CP capacity** during particular system recovery operations so as to **“speed up” the recovery**
 - **Shutdown** - when an LPAR is being shutdown
 - **IPL** - when an LPAR is being IPLs
- **Boosts are for 30 minutes for shutdown and 60 minutes for load**, and are initiated by the operating system automatically
- There are two forms of System Recovery Boost
 - **zIIP capacity boost** - Allows CP capacity to be run on all currently active zIIP engines for the boost period
 - **Speed boost** - Boosts sub-capacity CP engines to full capacity for the boost period
- There are no additional HW, SW, or maintenance charges for the boost period
- Requirements:
 - **z15**
 - **z/OS V2.5, V2.4 or z/OS V2.3 with PTFs**
 - **GDPS V4.2**

zIIP Capacity Boost

- During the zIIP capacity boost, **all active zIIPs associated with the LPAR** are used to **extend CP capacity**
- zIIP boost is only supported on **shared processor pools**
- zIIP processor boost applies only to **Hiperdispatch** mode partitions
- A **new temporary record**, the **“Boost” record** is available that allows the customer to **activate additional temporary zIIPs for a limited number of hours**
 - Customers must **own at least 1 zIIP** on the machine in order to purchase a zIIP boost record
 - The number of zIIPs provided by the record is **1 to 20 zIIPs**
 - The boost record will activate the zIIPs for **up to 6 hours**
 - The Boost record has an expiration date of one year
 - The number of boost records the customer may order is one per system
- Customers should **activate the boost record prior to the “boost” event**. They should plan on deactivating the record when the boost event is done
 - The record will self deactivate when it has used all of it’s hours, but customer deactivation should be encouraged

zIIP Capacity Boost

- Clients will configure a z/OS system-level parameter (parmlib) to control whether or not a particular z/OS image should opt-in for the zIIP processor Boost
 - IEASYSxx **BOOST=SYSTEM | ZIIP | SPEED | NONE** parameter
- z/OS images configured to exploit this capability will automatically opt-in at IPL time and opt-out at the end of their Boost period (**30' for shutdown, 60' for IPL**)
 - Start and End of Boost period is signaled via console message, ENF signal (84), and cutting an SMF record
 - Start and End of Boost period starts new SMF interval
 - A system command or PROC (name IEABE) will be provided to allow for early opt-out
- HMC/SE will show images that are currently opted-in for Boost
 - No z/OS display command provided (in stage 1)

zIIP Capacity Boost

- While in Boost period, each z/OS image sets up its dispatching so that general-purpose workload is eligible to be dispatched on zIIPs, **effectively merge all available GP and zIIP capacity together!**
 - “Entitled” **purchased GPs**
 - “Entitled” **purchased zIIPs**
 - Any **additional zIIPs** provided by the Recovery Boost record (temporary), if activated
- If **reserved logical zIIPs** are available and backed by physical zIIP capacity (e.g. additional zIIP capacity added using Boost eBod temporary capacity record), **images will bring additional logical zIIP processors online** to make use of the available physical zIIP capacity during the Boost period
 - **These additional logical processors are also automatically taken offline at the end of the Boost period**
- While in boost period, **boosted zIIP capacity is ignored for all forms of image and group capping**
 - Given that the boosted zIIP capacity is ignored, it will not contribute to the image being capped, **though capping might still happen based on other capacity**

zIIP Capacity Boost

- While in boost period, **capacity-related metrics show the boosted zIIP capacity** for the CPC and all opted-in images
 - **But this is zIIP capacity, not GP capacity or MSUs**
- While in boost period, **SMF records** used for pricing purposes (e.g. 30, 70.1, 72.3, 89) will **contain information about the boosted zIIP capacity**
 - A bit will be added to the SMF records to indicate whether or not the data pertains to an image that was in a processor Boost period; **SCRT will report on the zIIP Boost**
 - SCRT does not report on zIIP usage or consumption, for pricing purposes
- **After the Boost period ends**, z/OS dispatching of work on GPs vs zIIPs goes back to normal, and the activation of the **extra physical zIIP capacity** via the temporary capacity record, if any, **is deactivated**

System Recovery Boost Record (Temporary Capacity eBod)

- Temporary zIIP Processors for Turbo feature of System Recovery Boost
 - **Priced and Prepaid** – based on an Annual Subscription model for “recharging” the record indefinitely during that period
 - **Model z15 T01 only**
 - **zIIP processor type only**
 - Number of activations – **fixed number of activations**, but this can be “recharged” dynamically based on Annual Subscription model
 - Each activation has a **fixed 6-hour time period** - the record supports planned change activity windows of up to 6 hours during which **one or more system shutdown/startup** actions are being performed on the machine
 - Each activation has a **Number of zIIPs** supported in the activation: up to 20 zIIPs, and it could permit the violation of the 2:1 ratio rule between zIIPs and GPs
 - **Auto-deactivation at the end of the time period (6 hours)** the record is configured for (it could be manually deactivated sooner if desired)

zIIP capacity boost - Summary

- zIIP processor boost applies only to **Hiperdispatch** mode partitions
- zIIP processor boost applies only to **shared processor partitions**
- Definition and use of **SMT mode** by an image **is not affected** by zIIP processor boost; SMT will either be used for all zIIPs, or not, as currently defined
- Activation/deactivation of additional zIIP engines as zIIPs via the eBod record will trigger LPAR to perform **Dynamic Processor Reassignment** actions, as needed, to **remap physical to logical processor** definitions
- Note that even if additional zIIPs are not provided by the Boost temporary capacity record, there is still significant value provided by z/OS blurring work across GPs and any entitled, purchased zIIPs that the image may have

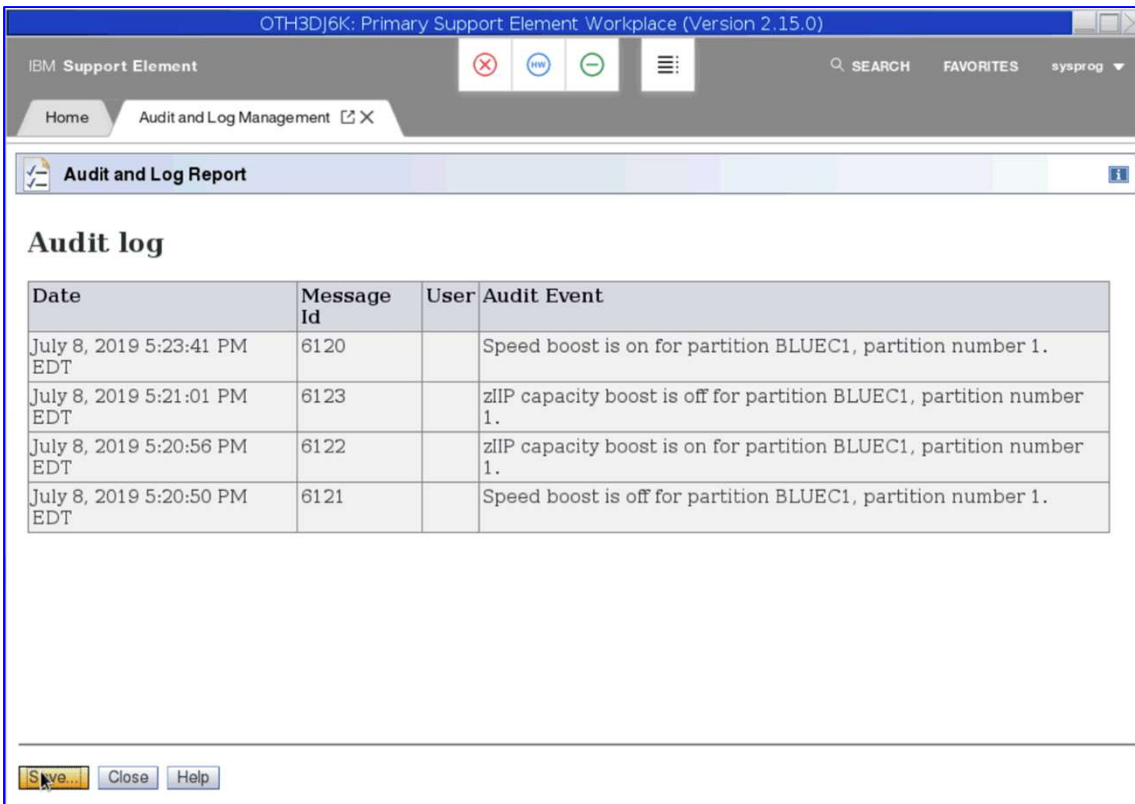
Subcapacity to full-capacity Speed Boost

- **Summary: When running on a subcapacity machine, temporarily run images at full-capacity speed to provide a short-term GP processor capacity and thread speed Boost, for the Boosting image(s) only**
- Applicable **ONLY to images running on subcapacity** machine models
 - T01: Models 4xx, 5xx, 6xx
 - T02: not Models z0x
 - z/OS
 - Clients will configure a z/OS system-level parameter (parmlib) to control whether or not a particular z/OS image should opt-in for the Speed Boost
 - IEASYSxx BOOST=SYSTEM | ZIIP (SPEED) | NONE parameter
 - OS images configured to exploit this capability will opt-in at IPL time and opt-out at the end of their Boost period (60 min period)
 - A system command or PROC (name IEABE) will be provided to allow for early opt-out if desired
 - Start and End of Boost period is signaled via console message, ENF signal (84), and cutting an SMF record
 - Start and End of Boost period starts new SMF interval

Subcapacity to full-capacity Speed Boost

- **HMC/SE** will show images that are opted-in for Boost
 - No z/OS display command provided
- While in Boost period, LPAR and CEC millicode will work together to dispatch CP processors for opted-in images in such a way that they will run at **full-cap speed, for the boosted image(s) only**
 - **While GP processors for other un-boosted images will continue to run at subcapacity speed**
- **While in Boost period, boosted GP capacity is ignored for all forms of LPAR and OS image and group capping; for capping purposes, the GPs appear as if they are *unboosted***
- Applicability notes:
 - **Speed boost** applies to both **Hiperdispatch** and **non-Hiperdispatch** mode partitions
 - **Speed boost** applies to both **shared processor** and **dedicated processor partitions**
 - Definition and use of **SMT** mode by an image **is not affected by GP Speed boost**, as SMT does not apply to GPs
 - Opting in for Speed boost on a CEC model that is **not subcapacity** to begin with has **no effect**

Audit log entries showing partition boosting



The screenshot displays the IBM Support Element Workplace interface. The title bar reads "OTH3DJ6K: Primary Support Element Workplace (Version 2.15.0)". The main window shows the "Audit and Log Management" section with a sub-tab "Audit and Log Report". Below this, the "Audit log" section contains a table with the following data:

Date	Message Id	User	Audit Event
July 8, 2019 5:23:41 PM EDT	6120		Speed boost is on for partition BLUEC1, partition number 1.
July 8, 2019 5:21:01 PM EDT	6123		zIIP capacity boost is off for partition BLUEC1, partition number 1.
July 8, 2019 5:20:56 PM EDT	6122		zIIP capacity boost is on for partition BLUEC1, partition number 1.
July 8, 2019 5:20:50 PM EDT	6121		Speed boost is off for partition BLUEC1, partition number 1.

At the bottom of the window, there are buttons for "Save...", "Close", and "Help".

- Audit entries are created for system recovery boost activation and deactivation
- Both zIIP capacity boost, and speed boost events are recorded
- Partition being boosted is identified

SE partition image details: boost is active

The screenshot shows the 'LP01 Details - LP01' dialog box in the IBM Support Element Workplace. The dialog box has two tabs: 'Instance Information' (selected) and 'Status'. The 'Instance Information' tab displays the following details:

Group:	Images
Image mode:	Not set
Activation profile:	LP01
Last used profile:	LP01
Sysplex name:	
Processor boost:	On
Secure execution:	Off
Operating system name:	
Operating system type:	
Operating system level:	
Group capacity name:	
CP management cluster name:	
Simultaneous Multi-Threading (SMT):	Inactive
Include CP's in Standby state:	<input type="radio"/> Yes <input checked="" type="radio"/> No
Lockout disruptive tasks:	<input type="radio"/> Yes <input checked="" type="radio"/> No

At the bottom of the dialog box, there are buttons for 'OK', 'Apply', 'Change Options...', 'Cancel', and 'Help'. The 'Processor boost' field is circled in red.

- Boost activity is shown on the SE partition image details
- Processor boost is shown as On or Off

HMC – Partition image details – Also shows Boost Status

The screenshot displays the IBM Hardware Management Console (HMC) interface, version 2.15.0. The main window shows the details for a partition image named 'IA7UNO2I:BLUEC1'. The 'Instance Information' tab is active, and the 'System recovery boost' status is highlighted with a red circle, indicating it is 'On'. Other details include the group 'CPC Images', activation profile 'BLUEC1', last used profile 'BLUEC1', operating system 'INSTALL', CPU LPAR cluster name, operating system type 'SSC', and operating system level '1.0.0'. The 'Task Information' section is also visible, showing 'Task name' and 'Task status' fields, and a 'Lock out disruptive tasks' option set to 'No'.

Instance Information	
Instance Information	
Group:	CPC Images
Activation profile:	BLUEC1
Last used profile:	BLUEC1
Sysplex name:	
System recovery boost:	On
Operating system:	INSTALL
CPU LPAR cluster name:	
Operating system type:	SSC
Operating system level:	1.0.0

Task Information	
Task name:	
Task status:	
Lock out disruptive tasks:	<input type="checkbox"/> Yes <input checked="" type="radio"/> No

Buttons: OK, Apply, Change Options..., Cancel, Help

HMC Monitors Dashboard – Processor boost status

HMCDAILY03: Hardware Management Console Workplace (Version 2.15.0) - Mozilla Firefox
 https://9.60.14.37/hmc/connects/mainuiFrameset.jsp

IBM Hardware Management Console

Home Monitors Dashboard - SE... X

Monitors Dashboard

Last refresh time: 03:15:15 PM Date: 04/08/19 Time zone: UTC-04:00 [Pause Refresh](#)

Overview

Select	Name	Status	Type	Machine Type - Model	Processor Usage(%)	IO Usage(%)	Power Consumption (kW) (Btu/hr)	Ambient Temperature (°C) (°F)
<input type="checkbox"/>	SETR87	Operating	CPC	8562 - T02	0	2		

Page 1 of 1 Max Page Size: 100 Total: 1 Filtered: 1 Displayed: 1 Selected: 0

Details

SETR87

Processors

Select	Name	Processor Usage(%)	SMT Usage(%)	Thread 0 Usage(%)	Thread 1 Usage(%)
<input type="checkbox"/>	CP000	0	0	0	0
<input type="checkbox"/>	CP004	0	0	0	0
<input type="checkbox"/>	CP008	0	0	0	0
<input type="checkbox"/>	CP00C	0	0	0	0
<input type="checkbox"/>	IFP010	0	0	0	0

Page 1 of 1 Max Page Size: 100 Total: 5 Filtered: 5 Displayed: 5 Selected: 0

Aggregated Processors

Select	Type	All Processor Usage(%)	Shared Processor Usage(%)
<input type="checkbox"/>	All Types	0	0
<input type="checkbox"/>	CP	0	0
<input type="checkbox"/>	IFP	0	0

Page 1 of 1 Max Page Size: 100 Total: 3 Filtered: 3 Displayed: 3 Selected: 0

Logical Partitions

Select	Name	All Processor Usage(%)	CBP Processor Usage(%)	CP Processor Usage(%)	ICF Processor Usage(%)	IFL Processor Usage(%)	ZIP Process Usage(%)	Processor Boost
<input type="checkbox"/>	LPO2	0	0	0	0	0	0	Off

Page 1 of 1 Max Page Size: 100 Total: 1 Filtered: 1 Displayed: 1 Selected: 0

Channels

Adapters

GDPS HW API actions performance and parallelism

- GDPS drives BCPii HW APIs for orchestrating CBU capacity activations, image activations, resets, and IPLs, for multiple images, in many planned and unplanned DR site-switch scenarios
 - There is value in improving both the **performance** of, and the usage **parallelism** of, these HW services, in those scenarios
- Exploitation/usage changes in GDPS software will take greater advantage of the available parallelism in the underlying HW services
 - Implement additional multitasking to drive HW actions in parallel, taking advantage of **available cross-CEC parallelism**, and avoid redundant per-action or per-system activities in the automation engine
 - *GDPS scripting changes are required to take full advantage of these enhancements!*
- **z15 Firmware changes** will be made to **support greater parallelism** and **performance** improvements in the HW API services themselves

GDPS Scripting Enhancements

- **Address multiple systems in same script command:**

SYSPLEX BCPII_Command System_Criterias Command_Options

- **BCPII_Command** can be:

- ACTIVATE, DEACTIVATE, RESET, LOAD, STOP, PSWRESTART

- **System_Criterias** can be:

- List of systems with or without generic names:
System(TSYS1,TSYS2,PS*,VM*)
- Logical Group of systems:
GROUP(SITE1) , GROUP(ALL), GROUP(ZOS), etc...

- **Customized group of systems** using system type , site parms:

GROUP(Type=YV,Site=*) or GROUP(Type=*,Site=*) or GROUP(Type=N,Site=1), etc...

- **Examples:**

```
SYSPLEX ACTIVATE GROUP(SITE1)
```

```
SYSPLEX LOAD GROUP(Type=YN,Site=*)
```

```
SYSPLEX RESET SYSTEM(G0C1,G0P*,SVM*)
```

Pricing

Boost should NOT INCREASE customers' IBM software costs, regardless of whether client is using 4HRA Pricing, Solution Pricing, or Consumption based Pricing

- **System Recovery Boost Base Function**
 - **Base capabilities are included for no extra charge:**
 - Speed Boost for GPs
 - zIIP Boost using client's entitled zIIPs
 - GDPS scripting and firmware enhancements
 - Terms and conditions define how/when these functions may be used in an unpriced, unlimited way for system startup and shutdown purposes
 - No need for any capacity planning as will be included in base offering
- **System Recovery Boost With Boost Temporary Capacity Record**
 - **Priced activation of additional temporary physical zIIP capacity for Boost usage via the Boost temporary capacity record**
 - Annual subscription model for entitlement to unrestricted system startup and shutdown using zIIP Boost with activation of zIIP engines via temporary record
 - We intend that annual subscription model pricing eliminates the need for capacity planning around the use of the additional capacity Boosts

Capping Summary

- Some forms of capping are managed in LPAR, some in z/OS; some forms of capping apply to GPs, some to zIIPs, some to both
- Some forms of capping (Resource Group capping) apply to apportioning usage of capacity *within a z/OS image*, by different tenants, to avoid monopolization by particular tenants
- Our general approach to capping is:
 - **Speed boost:** during boost, count GPs as if they were still subcapacity speed, and cap normally based on that *understated* view of GP capacity
 - **zIIP processor boost:** during boost, do not count boosting partition's entitled zIIPs or boost zIIPs for capping purposes; do not suspend boosting partition based on zIIP capacity (essentially, ignore the cap)

SRB Stage 2 Recovery Events

- SRB Stage 2 provides boosted processor capacity and parallelism to accelerate the following recovery events:
 - **Sysplex Partitioning Recovery**
 - **Boost all surviving systems in the sysplex as they take on the additional workload of sysplex partitioning related recovery, after planned or unplanned removal of a system from the sysplex**
 - When a system in the sysplex is removed, the surviving systems have to do a large amount of recovery processing to clean up after the failed system, free up resources that were held on the failed system, etc.
 - **CF Structure Recovery**
 - **Boost all systems participating in CF structure recovery processing – CF structure rebuild, duplexing failover, re-duplexing**
 - Recovering failed CF structures and their data can be a laborious process that requires the participation of all systems that were using those CF structures, and can apply to many structures in cases like loss of a CF image
 - **CF Datasharing Member Recovery**
 - **Boost all systems participating in recovery from termination of a CF datasharing member**
 - When a datasharing member (e.g. a DB2 instance) fails, the other surviving members have to do a lot of recovery/cleanup processing to free up locks and other datasharing resources held by the failed member
 - **Hyperswap Recovery**
 - **Boost all systems participating in a Hyperswap recovery process**
 - Hyperswap processing is a coordinated, sysplex-wide recovery process that restores access to DASD devices following the failure of a storage controller. Its recovery time is sometimes limited by slow processing on one or more participating systems.

SRB Stage 2 Support Overview*

- Stage 2 provides support for **short-duration** “recovery process” boost periods, which are a separate and distinct new class of boosts from today’s IPL/startup and shutdown boost periods
- Each image can receive boosts as follows:
 - **Long-duration boost** for **image startup** (60 minutes), of each type (CP speed boost, zIIP boost), and **one** long-duration boost **for image shutdown** (30 minutes), of each type
 - **Short-duration recovery** process boosts, of each type, **each of less than 5 (five) minutes** duration, with a total usage of **no more than 30 minutes of recovery process boost time in any given consecutive 24-hour period**
- z/OS manages the recovery process boosts internally, with the **operating system initiating the boosts** *as these recovery events take place, and only on the images that are affected by these events*
 - If recovery process boosts happen to “overlap” – a second recovery process boost occurs before a first one has used its entire boost period – then the overlapping boosts are merged and the boost period may be extended to allow the full boost period time for the second recovery process
- New LPAR firmware in **z15 T01** and **T02** provides the support for the **short-duration recovery** process boosts
 - Concurrently-installable* LPAR MCL on z15 T01 and T02;
 - z/OS support will react dynamically to a concurrent installation of LPAR MCL to provide these capabilities
 - LPAR support also times the use of recovery process boosts and limits the total usage per-day for each image

SRB Stage 2 New Display Support

- In **SRB Stage 1** support, **z/OS** provided **no display command** support that showed the details regarding an active boost period for that image
- For **SRB Stage 2**, the **D IPLINFO,BOOST,STATE** command is enhanced to show an image's current boost state, both for Stage 1 startup/shutdown boosts, and the new recovery process boosts:

Example 1:

```
IEE257I Boost State
Boost class: IPL
zIIP boost: active with 5 transient zIIP cores
Speed boost: active
```

Example 2:

```
IEE257I Boost State
Boost class: Recovery Process
Requestor: Hyperswap
zIIP boost: active with 2 transient zIIP cores
Speed boost: active
```


SRB Stage 2 New Display Support

- For SRB Stage 2, **DISPLAY M=CPU** is also enhanced:
 - (Existing) “I” indicates “zIIP”
 - (New) “B” indicates (transient) “boost zIIP”. This CPU was config’d online at the start of the boost period, and will be config’d offline when the boost ends

```
SY1 IEE174I 09.58.10 DISPLAY M 328
PROCESSOR STATUS
ID CPU SERIAL
00 + 0449D74381
01 + 1449D74381
02 +I 2449D74381
03 +B 3449D74381
04 +I 4449D74381
```

RMF Monitor III

```
Samples: 316      System: TSTB  Date: 04/30/21  Time: 21.43.30  Range: 317  S
Partition:  TEST      8562 Model T04      Boost: All
CPC Capacity:  418    Weight % of Max: ****  4h Avg:  10
Image Capacity: 120    WLM Capping %:  0.0    4h Max:  67    Group: TEST
MT Mode IIP:   2      Prod % IIP:      77.4    AbsMSUCap: N   Limit:  120*
```

```
Samples: 300      System: TSTB  Date: 04/30/21  Time: 21.59.00  Range: 300  Sec
Partition:  TEST      8562 Model T04      Boost: N
CPC Capacity:  418    Weight % of Max: ****  4h Avg:  13
Image Capacity: 120    WLM Capping %:  0.0    4h Max: 103    Group: TEST
MT Mode IIP:   2      Prod % IIP:      49.2    AbsMSUCap: N   Limit:  120*
```

Boost

- ZIIP
- SPEED
- ALL

Questions?

End of Section

