zIIPs and zAAPs: Everything New and Old

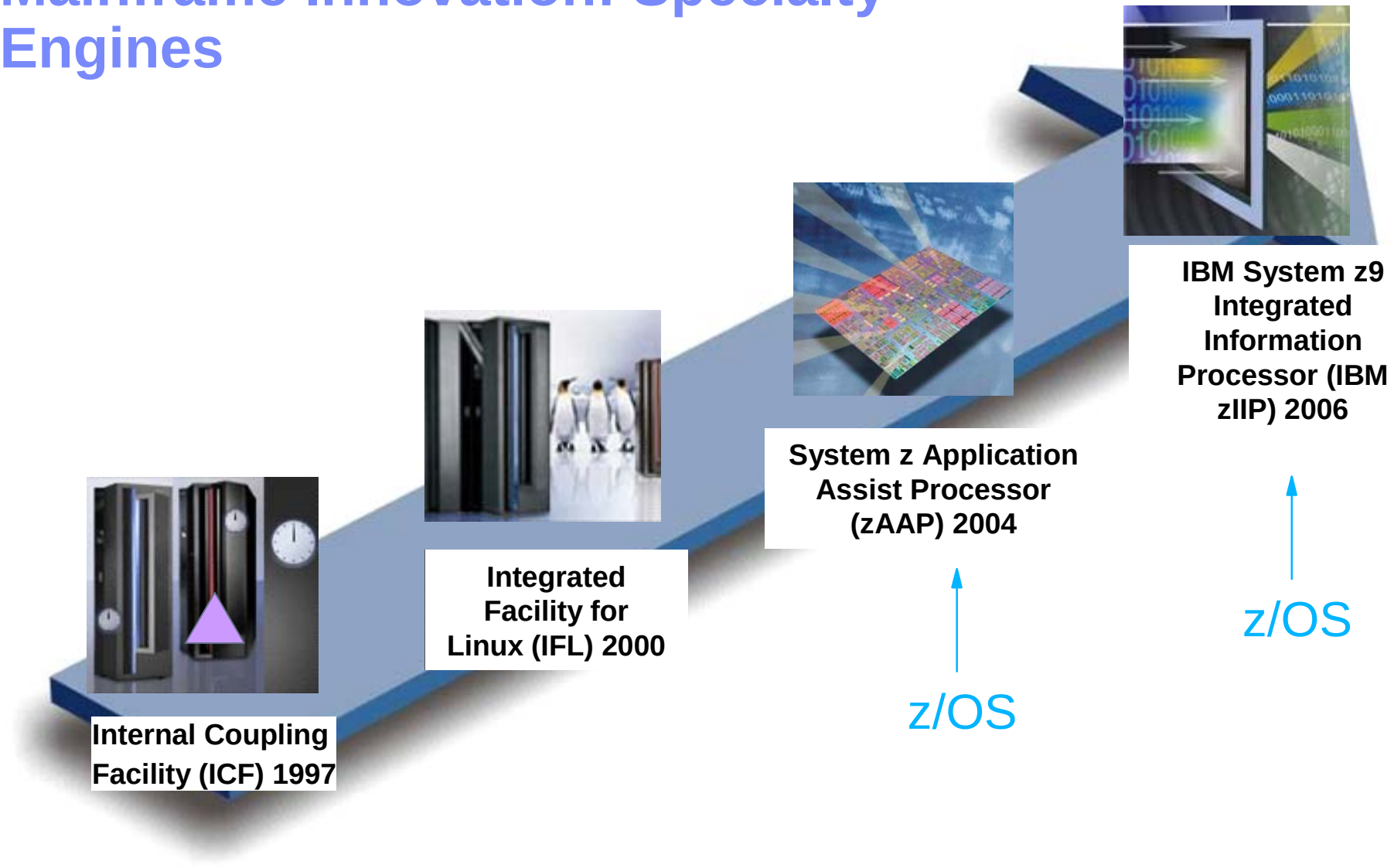
SHARE - Winter 2009
March1-6, 2009
Session: 2535

Kathy Walsh
IBM Corporation
IBM Distinguished Engineer
Washington Systems Center

Agenda

- Overview of zAAPs and zIIPs
- Processor Impacts
  - LPAR definitions and Controls
  - Capacity Planning
- z/OS Impacts
  - MP Effects
  - CPU times
  - Parmlib controls
  - RMF Reporting

Mainframe Innovation: Specialty Engines



Specialty CP Overview

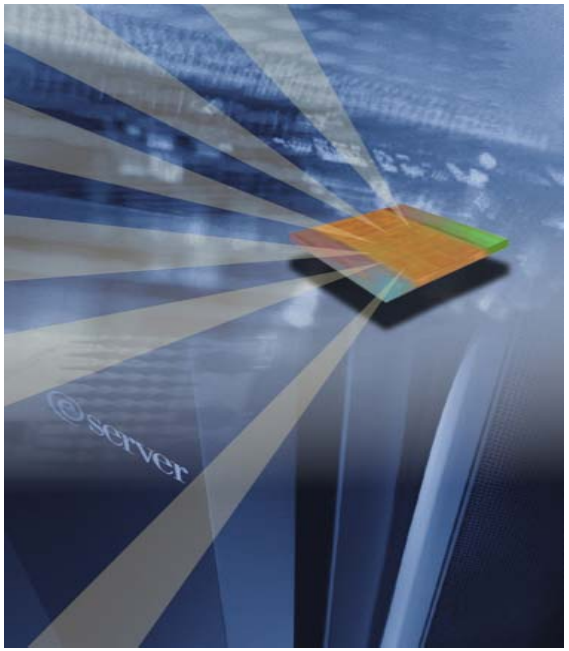
- Specialty engines for the mainframe designed to help:
  - Customers integrate data across the enterprise
  - Improve resource optimization and lower the cost of ownership for eligible workloads
  - Attractively priced, much lower than standard CPs
  - Significantly lower maintenance costs than standard CPs
  - Traditional IBM System z software charges unaffected
  - Sub-capacity eligible IBM software charges can be reduce

- z/OS manages and directs work between the general purpose processor and the Specialty CPs
  - Number of zAAPs CPs not to exceed number of general CPs
  - Number of zIIP CPs not to exceed number of general CPs

IBM System z Application Assist Processor (zAAP)

New specialty assist processor dedicated exclusively to execution of Java workloads under z/OS® – e.g. WebSphere®, CICS, IMS, DB2

- Available on IBM System z10, IBM System z9 and @server zSeries z990 and z890
- Used by any workload with Java cycles, e.g. WebSphere, CICS, Batch, DB2®,
- Executes Java code with no changes to applications
- ***Objective: Enable integration of new Java based Web applications with core z/OS backend database environment for high performance, reliability, availability, security, and lower total cost of ownership***



IBM System z Integrated Information Processor (zIIP)

New specialty assist processor dedicated exclusively to execution of any workloads under z/OS®r

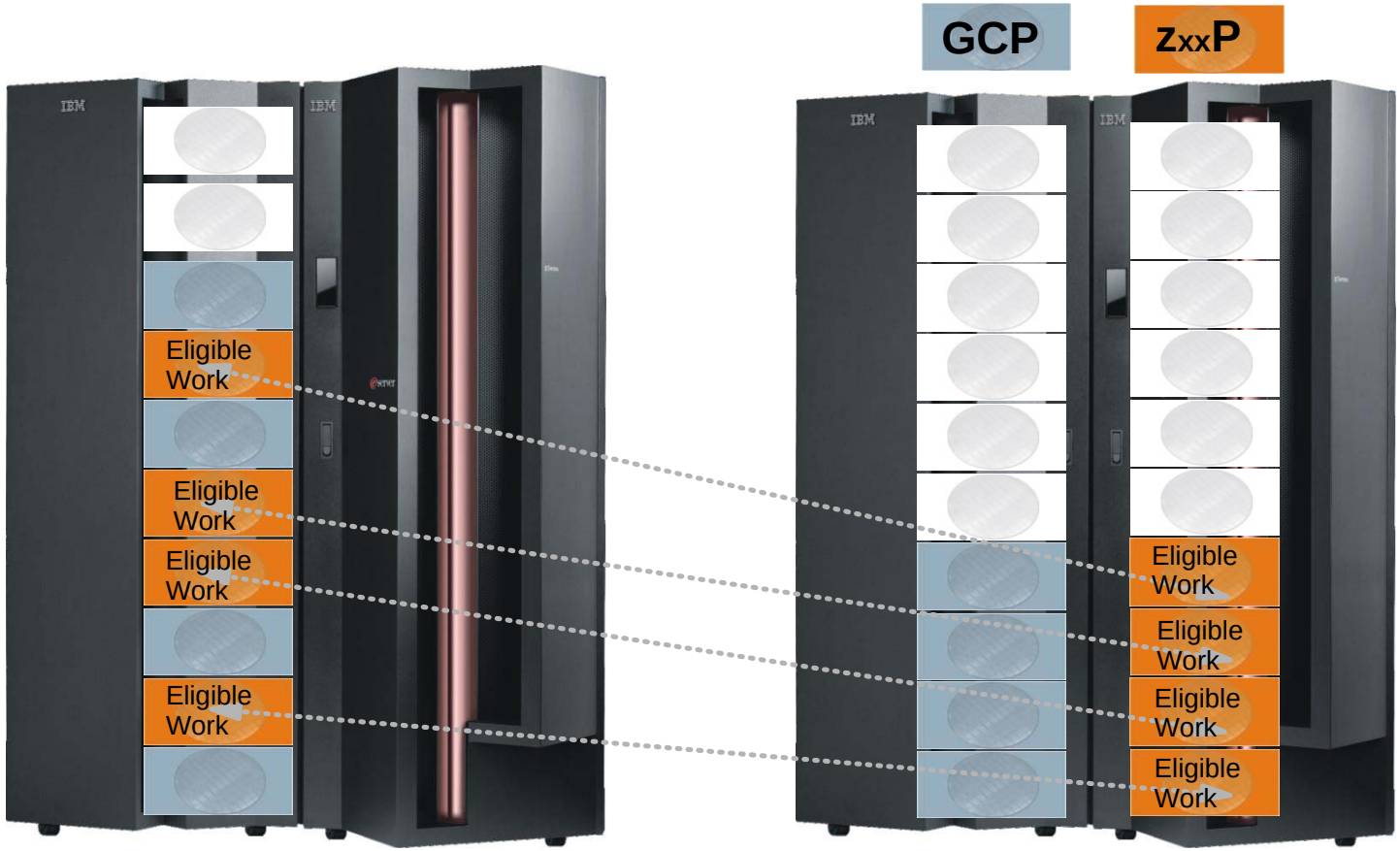
- Specialty engine for the IBM System z9 and IBM System z10 mainframe
- Requires work to be run as an enclave SRB
- Exploiters work with z/OS to determine how much capacity should be redirected to a zIIP
- Exploiters must use licensed interface to enable exploitation

Current IBM Exploitation

Specialty CP	Eligible	Major Users
zAAP	Any Java Execution	Websphere CICS Native apps XMLSS
zIIP	Enclave SRBs	DRDA over TCPIP Parallel Query DB2 Utilities Load, Reorg, Rebuild DB2 V9 z/OS remote native SQL procedures TCPIP - IPSEC XMLSS zIIP Assisted HiperSockets Multiple Write

Specialty CP Operation

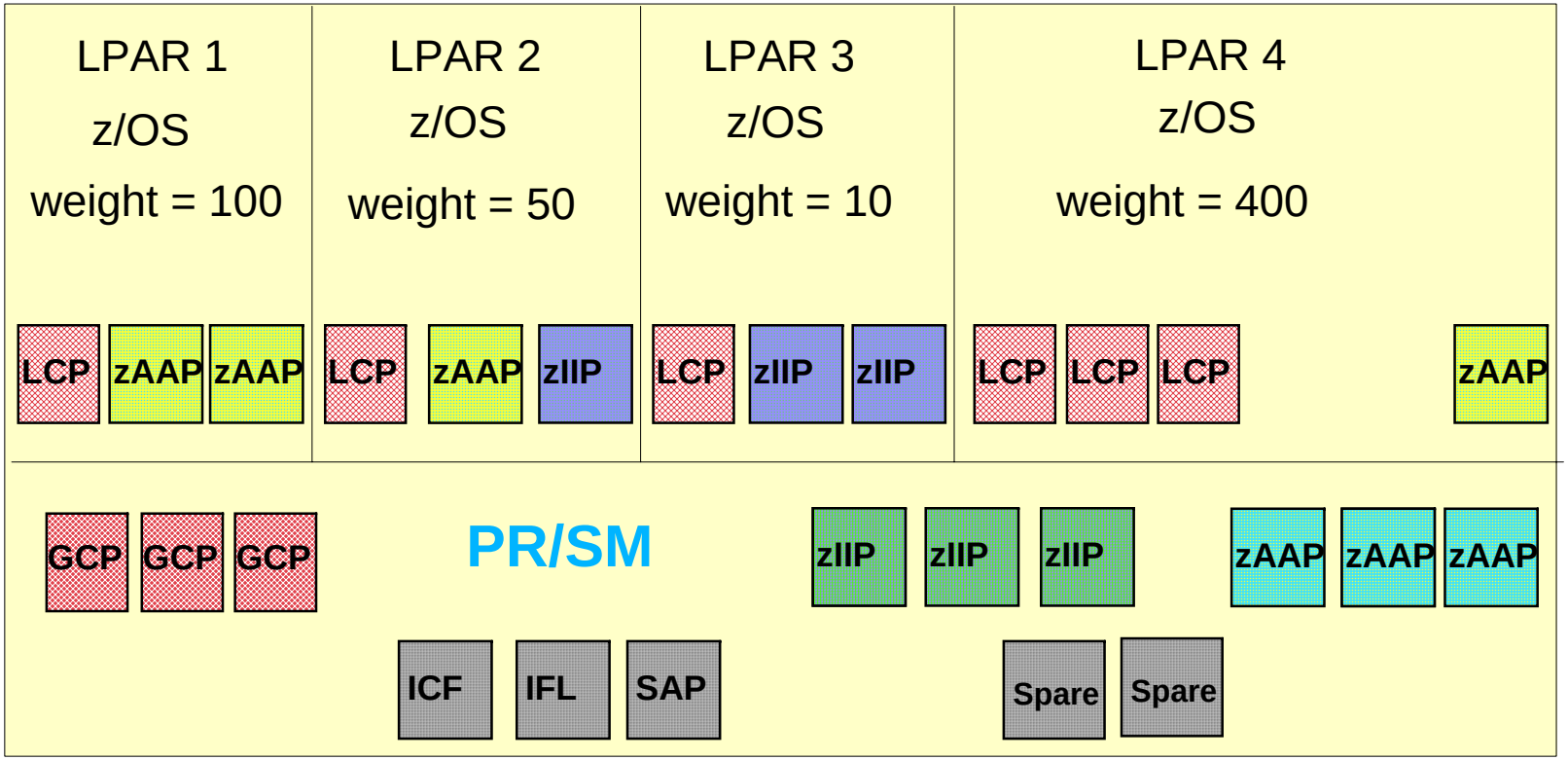
- Eligible work moves to the specialty CP under the control of z/OS
- Provide more capacity for eligible workloads without a change in MSU
- Provides more capacity for non-eligible workloads
- Reduces software stack charges for non-eligible products on CEC



* For illustrative purposes only

Hardware View

- Number of specialty processors can't exceed the number of GCPs
  - This relationship does not extend to the LPAR view
- PR/SM dispatches specialty CPs in the same manner as GCP
  - Managed to weight and number of logical specialty CPs
  - CPs can be shared or dedicated



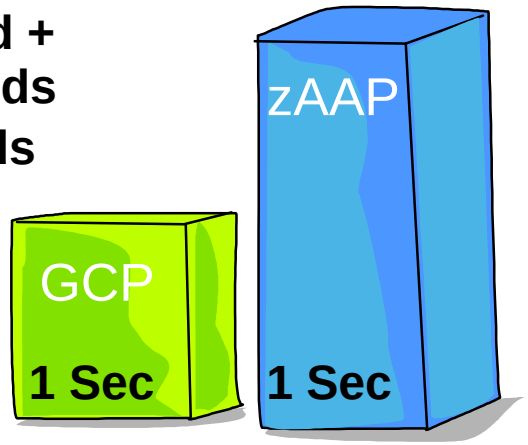
Subcapacity GCPs and Specialty CPs

- Specialty CPs always run at full speed of processor model
- Sub-capacity GCP are available on the z890, the z9 BC and EC processors, and the z10
- Same z/OS image has CPs running at different speeds
  - Requires CPU seconds to be normalized

Example: zAAP is 8 times the speed of the GCP

Execution Time = GCP seconds + (zAAP seconds * normalization factor)

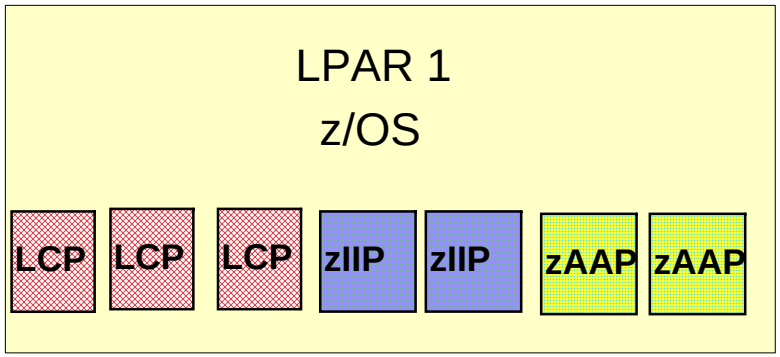
Execution Time = 1 second + (1 second * (8)) = 8 seconds
= 9 seconds



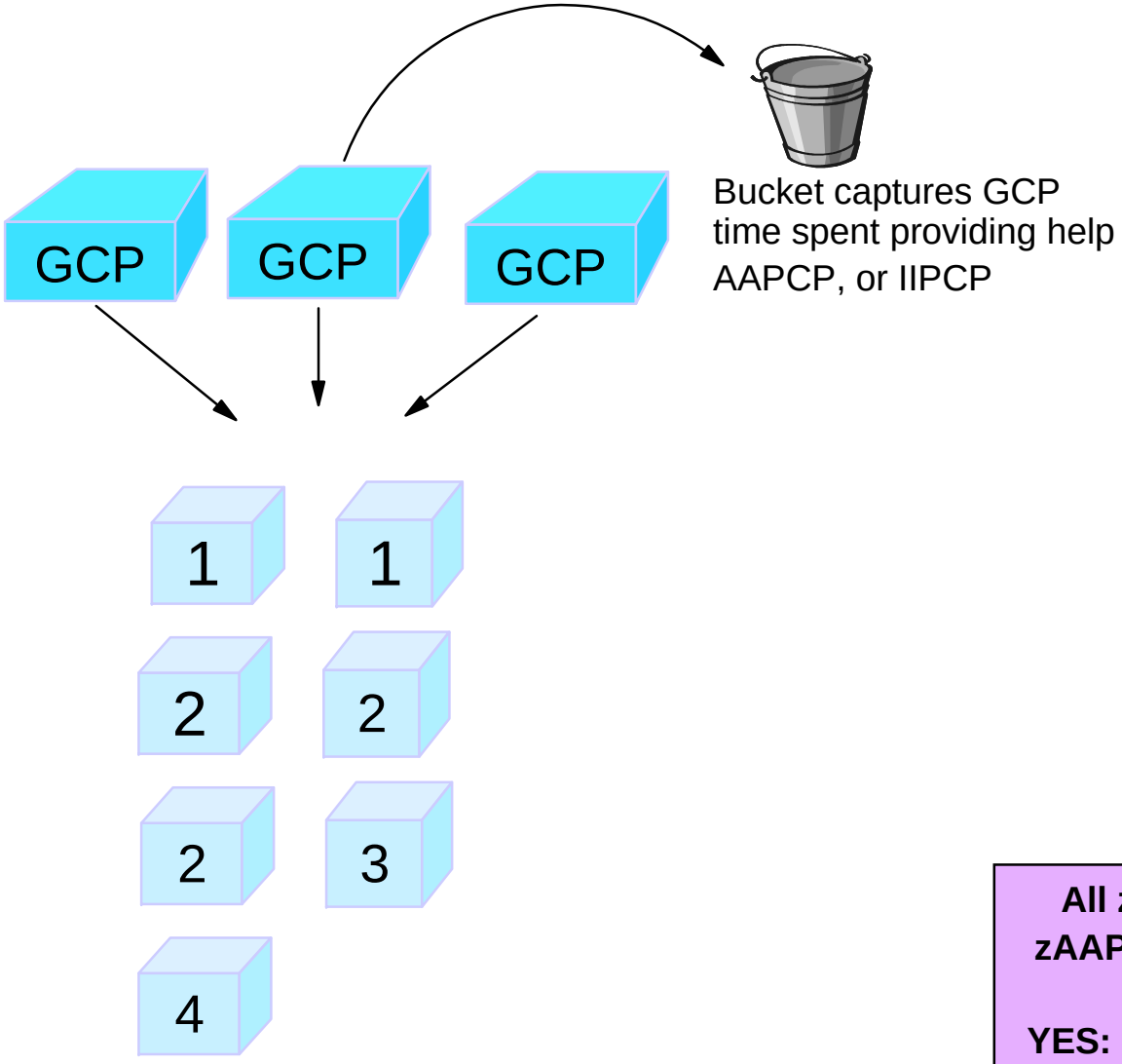
- **Normalization factor used is in RMF 72 subtype 3 record, R723NFFI (zAAP) or R791NFFS (zIIP)**
- **Normalization factor used is in the SMF 30 record, SMF30ZNF (zAAP) or SMF30SNF (zIIP)**
- **When zAAP/zIIP and GCP are the same speed the normalization factor resolves to 1**

z/OS View of Specialty CPs

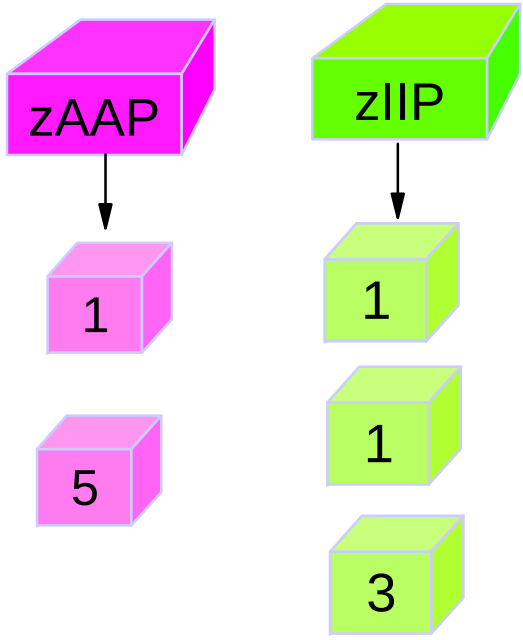
- Specialty CPs:
  - Are managed by z/OS and contribute to the upper limit of CPs managed by a copy of z/OS (64)
  - Do not impact the SU/SEC value nor the MSU rating
  - Can be configured on and off by CF CPU command
  - Are not managed by IRD
  - Are managed by hiperdispatch (zIIP support in z/OS 1.11)
  - Do not influence LPAR Group Capacity controls
  - Don't take I/O interrupts
  - No Clock Comparator interrupts
  - No affinity scheduling
  - Suspend Lock negotiation is always done on GCPs
  - Are not considered in WLM routing decisions until z/OS 1.9



Management of Work - Honorpriority=YES



GCPs will help in priority order



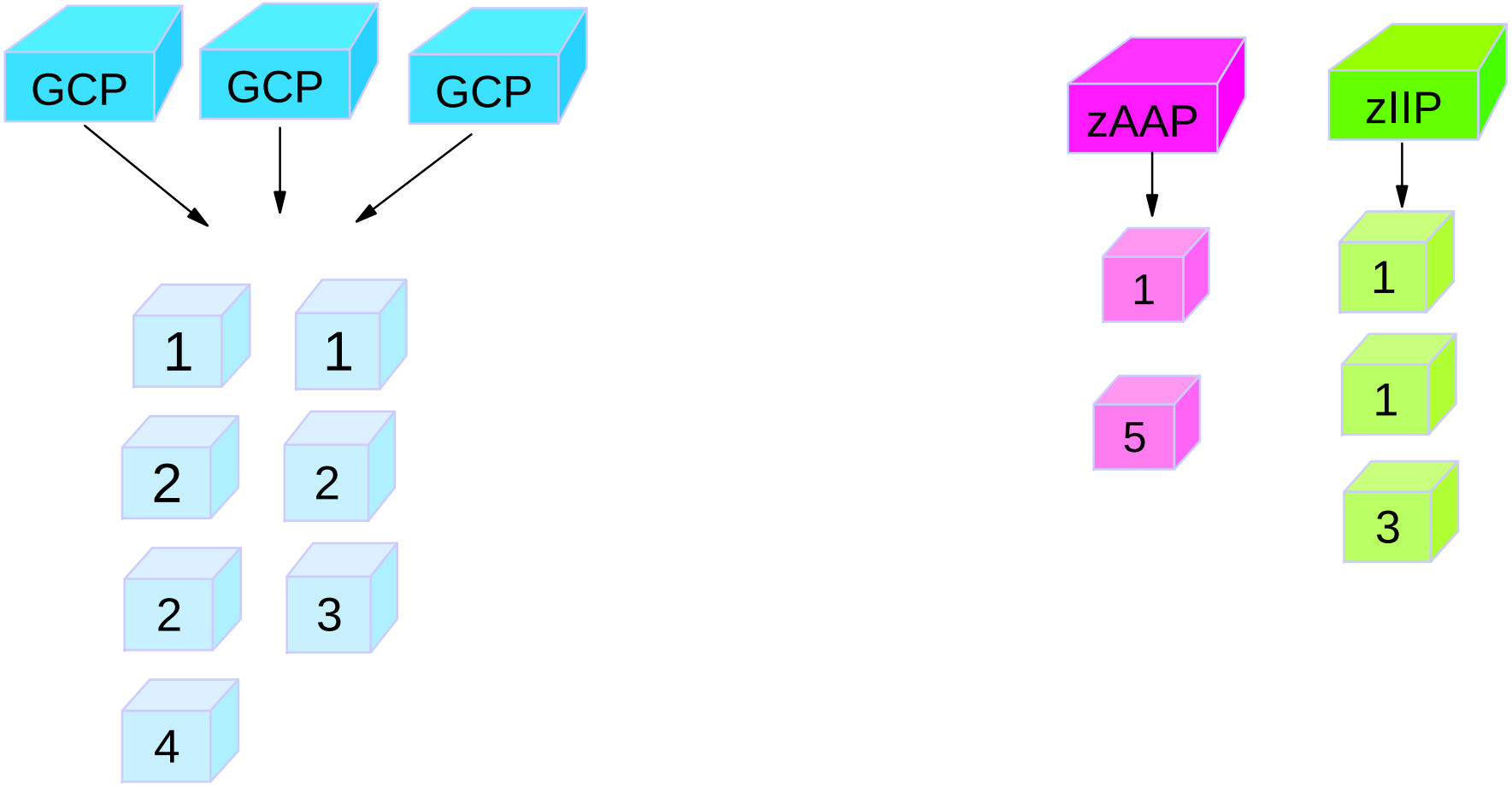
All zAAPs Busy?
zAAP work waiting?

YES: Signal GCP for help

All zIIPs Busy?
zIIP work waiting?

YES: Signal GCP for help

Management of Work - Honorpriority=NO



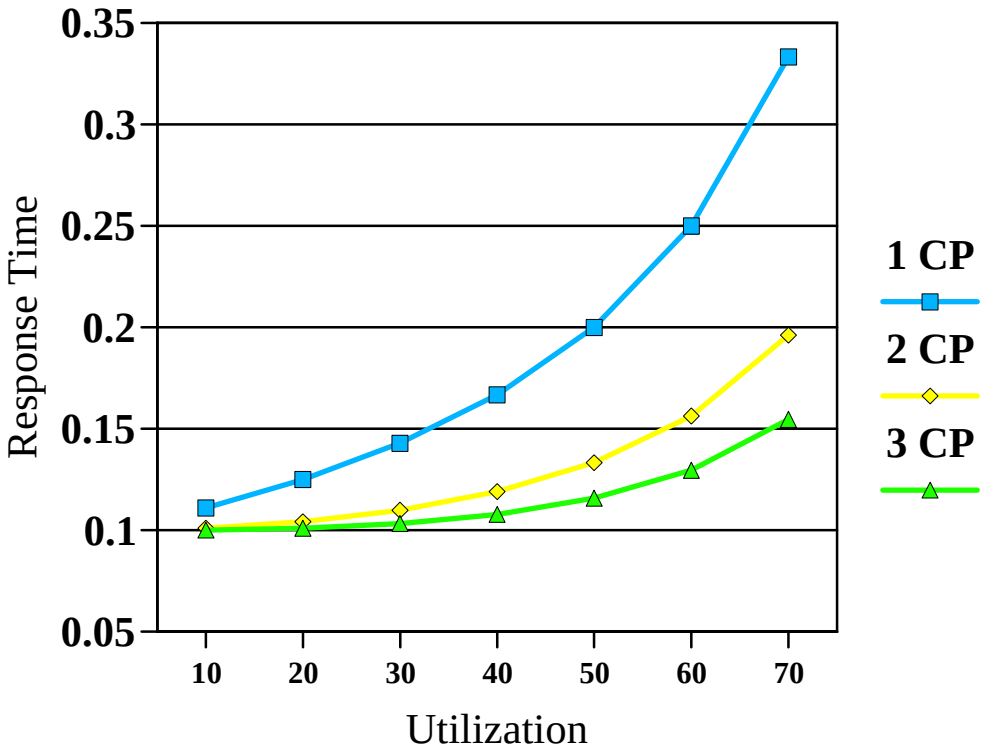
GCPs will never help

What is "Needs Help"

- Action taken by a GCP to process work queued to a zIIP or zAAP to protect response times
  - Configurations often have fewer specialty CPs than GCPs
  - Can be disabled by xxPHONORPRIORITY=NO
- Controlled by IEAOPTxx parameters: ZAAPAWT and ZIIPAWT
  - Defaut: 12000 - 12 ms **OR** Default: 3200 - 3.2 ms for Hiperdispatch=YES systems
  - Internally affects how frequently specialty CPs check the need for help
  - If help is required, the zxxP CP signals waiting zxxP to help
    - When all zxxP CPs are busy the zxxP ask for help from GCP
    - All available speciality engines (of the same type) must be busy before help is asked of the GCPs
  - IF the zxxPs needs help and all zxxPs are busy help is obtained from 1 GCP
    - If zxxPs continue to need help additional CPs may be asked to help
- Reducing the ZAAPAWT or ZIIPAWT value too low can cause the GCPs to run an excessive amount of zAAP or zIIP processor eligible workload
  - High AAPCP/IIPCP times
- Increasing the value specified for ZAAPAWT or ZIIPAWT causes the specialty engines to request help only after being busy for a longer period of time
  - Risks response times

Queuing Impacts of Server Busy

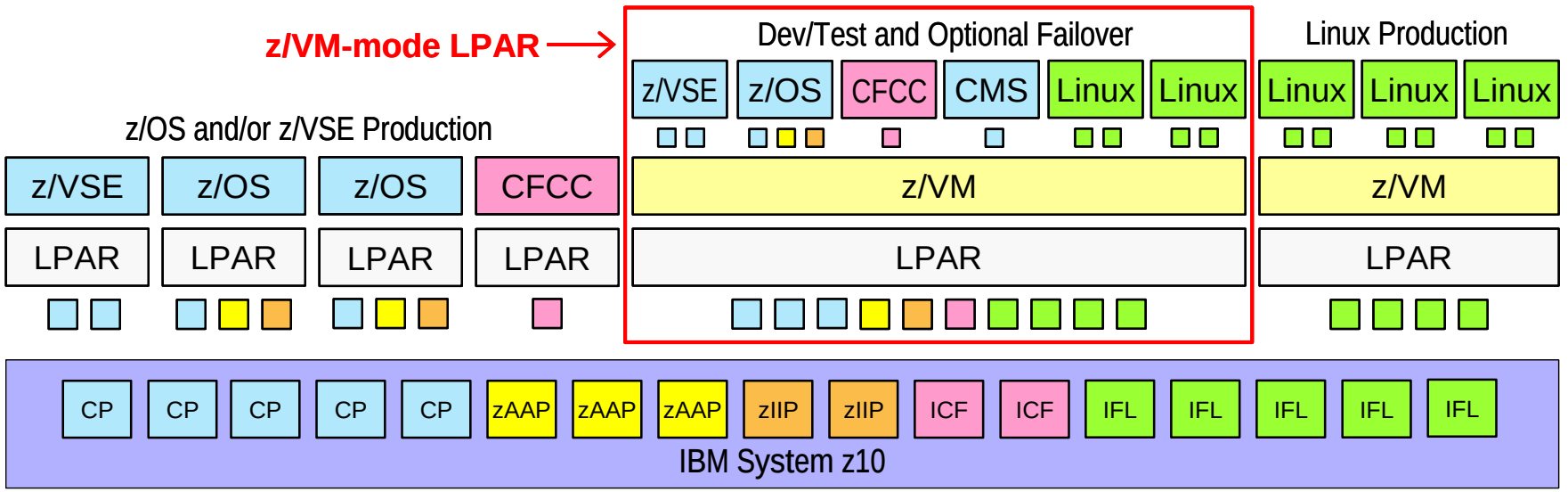
	1 CP	2 CP	3 CP	4 CP
1	0.1010	0.1000	0.1000	0.1000
10	0.1111	0.1010	0.1001	0.1000
20	0.1250	0.1042	0.1010	0.1003
30	0.1429	0.1099	0.1033	0.1013
40	0.1667	0.1190	0.1078	0.1038
50	0.2000	0.1333	0.1158	0.1087
60	0.2500	0.1563	0.1296	0.1179
70	0.3333	0.1961	0.1547	0.1357
80	0.5000	0.2778	0.2079	0.1746
90	1.0000	0.5263	0.3724	0.2969
99	10.0000	5.0251	3.3706	2.5448



- xxPHonorpriority=NO is a valid method of running Specialty CPs
- Arrival rates and Server busy will influence IIPCP/AAPCP time
- Can run Specialty CPs very busy IF there are multiple classes of work with different response time objectives

z/VM-Mode LPAR Support for IBM System z10

- New LPAR type for IBM System z10: z/VM-mode
  - Allows z/VM V5.4 users to configure all CPU types in a z10 LPAR
- Offers added flexibility for hosting mainframe workloads
  - Add IFLs to an existing standard-engine z/VM LPAR to host Linux workloads
  - Add CPs to an existing IFL z/VM LPAR to host z/OS, z/VSE, or traditional CMS workloads
  - Add zAAPs and zIIPs to host eligible z/OS specialty-engine processing
  - Test integrated Linux and z/OS and z/VSE solutions in the same LPAR
- No change to IBM software licensing
  - Software continues to be licensed according to CPU type



Performance / Capacity Planning

- Performance and Capacity data is fully integrated for zxxP
  - All RMF CPU reporting fields for GCPs are provided for zxxP
    - At the LPAR and application level
  - New zxxP fields created for accounting purposes
  - MSU rating of the processor does not change when zxxPs are added
  - For subcapacity pricing rolling 4 hr average does not include zxxP time

- Enhanced Capacity Planning support - AAPCP / IIPCP
  - Time spent on a GCP but zxxP eligible
    - Can be used to determine zxxP requirements without the hardware
    - Can also be used to determine when more capacity is needed if have the hardware installed

- Migration Actions:
  - Update the performance reporting systems to show zxxP usage
  - Update capacity planning to project zxxP usage
  - Update the accounting method to accommodate zxxP usage

CPU Activity Report for Specialty CPs

```
                                                C P U  A C T I V I T Y

            z/OS V1R7                  SYSTEM ID SYSD             DATE 08/18/2006            INTERVAL
00.59.999
                                       RPT VERSION V1R7 RMF       TIME 16.23.00              CYCLE 1.000

CPU  2094   MODEL  750   H/W MODEL  S54
---CPU---  ONLINE TIME   LPAR BUSY       MVS BUSY      CPU SERIAL  I/O TOTAL          % I/O INTERRUPTS
NUM  TYPE  PERCENTAGE    TIME PERC       TIME PERC     NUMBER      INTERRUPT RATE     HANDLED VIA TPI
 0   CP    100.00         2.50            2.41         04B10E      27.15               0.49
 1   CP    100.00         1.99            1.92         04B10E      24.37               0.14
 2   CP    100.00         1.85            1.77         04B10E      38.32               0.04
 3   CP    100.00         1.68            1.58         04B10E      28.18               0.12
CP   TOTAL/AVERAGE        2.01            1.92                     118.0               0.18
 6   AAP   100.00        40.19           40.96         04B10E
 7   AAP   100.00         0.31            0.29         04B10E
AAP  AVERAGE             20.25           20.63
 8   IIP   100.00         0.29            0.17         04B10E
 9   IIP   100.00         0.00            0.00         04B10E
IIP  AVERAGE              0.15            0.08
```

Partition Data Report with Specialty CPs

```
--------- PARTITION DATA --------- LOGICAL PARTITION PROCESSOR DATA --   -- AVE PROC UTILIZATION PERCENT
              ----MSU----  PROCESSOR-  ----DISPATCH TIME DATA----   LOGICAL PROCESSORS  PHYSICAL PROC
NAME     S   WGT  DEF    ACT  NUM   TYPE   EFFECTIVE       TOTAL      EFFECTIVE TOTAL  LPAR MGMT  EFFECTIVE
SOSP4    A   500    0      4    4   CP    00.00.04.349  00.00.04.814   1.81     2.01      0.02       0.14
....
SOSP15   A    10    0      2    2   CP    00.00.01.593  00.00.02.541   1.33     2.12      0.03       0.05
SOSP16   A    10    0      4    2   CP    00.00.04.305  00.00.05.208   3.59     4.34      0.03       0.14
SOSP17   A    10    0      3    2   CP    00.00.02.671  00.00.03.614   2.23     3.01      0.03       0.09
*PHYSICAL*                                 00.00.25.703                                    0.86
                                           ------------  ------------                    ------     ------
  TOTAL                                    00.20.33.086  00.21.04.443                    1.05       41.10

SOSP4    A   500                2   AAP   00.00.24.118  00.00.24.295  20.10    20.25      0.15       20.10
SOSP15   A    10                1   AAP   00.00.00.353  00.00.00.876   0.59     1.46      0.44       0.29
SOSP16   A    10                1   AAP   00.00.01.772  00.00.02.282   2.95     3.80      0.43       1.48
SOSP17   A    10                1   AAP   00.00.01.357  00.00.01.875   2.26     3.13      0.43       1.13
*PHYSICAL*                                 00.00.06.608                                    5.51
                                           ------------  ------------                    ------     ------
  TOTAL                                    00.00.27.601  00.00.35.939                    6.95       23.00

SOSP4    A   500                2   IIP   00.00.00.058  00.00.00.174   0.05     0.15      0.10       0.05
*PHYSICAL*                                 00.00.02.002                                    1.67
                                           ------------  ------------                    ------     ------
  TOTAL                                    00.00.00.058  00.00.02.176                    1.76       0.05
```

* RMF report has been modified to fit into the space

Specialty CP work running in a WLM Service Class

```
REPORT BY: POLICY=WLMPOL     WORKLOAD=BAT_WKL     SERVICE CLASS=BATSPEC     RESOURCE GROUP=BATMAXRG
TRANSACTIONS    TRANS-TIME HHH.MM.SS.TTT  --DASD I/O--  ---SERVICE----  SERVICE TIMES  ---APPL %---
AVG      0.98   ACTUAL              6.520  SSCHRT  11.5  IOC      8326   CPU     24.7   CP       0.97
MPL      0.98   EXECUTION           6.128  RESP     7.0  CPU    662386   SRB      0.0   AAPCP    0.01
ENDED      10   QUEUED                391  CONN     6.9  MSO         0   RCT      0.0   IIPCP    0.00
END/S    0.17   R/S AFFIN               0  DISC     0.0  SRB       965   IIT      0.0
#SWAPS      0   INELIGIBLE              0  Q+PEND   0.1  TOT    671677   HST      0.0   AAP     40.27
EXCTD       0   CONVERSION              0  IOSQ     0.0  /SEC    11195   AAP     24.2   IIP      0.00
AVG ENC  0.00   STD DEV                 0                               IIP      0.0

 GOAL: EXECUTION VELOCITY 35.0%     VELOCITY MIGRATION:   I/O MGMT  99.2%     INIT MGMT 92.2%

          RESPONSE TIME EX   PERF  AVG    ------ USING% ----- ------------ EXECUTION DELAYS % --------
SYSTEM                   VEL% INDX ADRSP    CPU  AAP  IIP  I/O  TOT  CPU

SYSD          --N/A--    99.2  0.4   1.0    0.8 45.9  0.0  3.9  0.4  0.4

SYSTEMS
  ---ID--- OPT SU/SEC  --TIME-- INTERVAL
  SYSD      00 26845.6 16.24.00 00.01.00

RESOURCE GROUPS
  --NAME--    ----------DESCRIPTION-----------     -SERVICE-      ACTUAL        ---CAPACITY---
                                                     CLASS       CONSUMED        MIN      MAX
  BATMAXRG                                                           11K                 6000
                                                  BATSPEC            11K
```

Capacity Planning Factors in a zxxP Environment

- Software
  - MP effects on z/OS
  - Switch request rate
  - Utilization of the zxxP
- Hardware
  - Number of books
  - Cost of additional partitions
- Use zPCR tool to estimate the costs
  - Build customer specific configurations including zxxPs
- Download zPCR and training on the tool at:
  http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/PRS1381

Impact of Multi-Book on Capacity

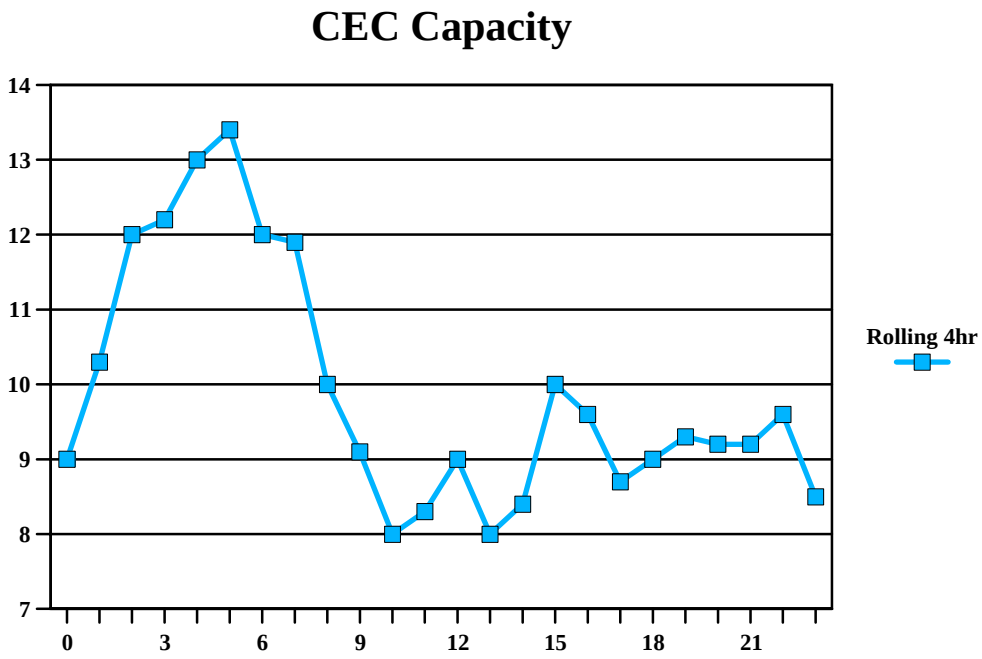
- There will be a slight impact on capacity when running multiple books and specialty CPs
  - Busy of the specialty CPs will govern the extent of the impact
  - zPCR assumes the specialty CPs are fully utilized (90%) and so gives a conservative view of capacity *
- Understand the impact of configurations set up for Disaster Recovery
  - More books to allow for CBU or CUOD

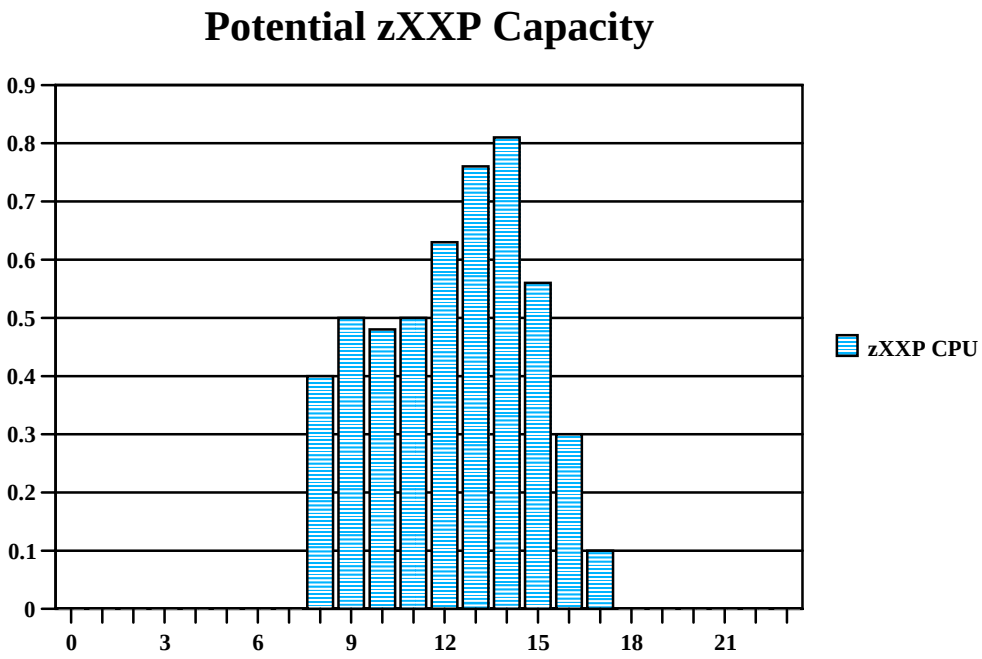
Environment	GCP MIPS	PER CP	zIIP MIPS	zAAP MIPS
2094 S08 with 4 GCPs	2292	573	N/A	N/A
2094 S08 with 4 GCPs, 2 zIIPs and 2 zAAPs	2204	551	1069	1069
2094 S08 with 8 GCPs	4291	536	N/A	N/A
2094 S38 with 4 GCPs	2242	561	N/A	N/A
2094 S38 with 4 GCPs, 2 zIIPs and 2 zAAPs	2106	527	998	998
2094 S38 with 8 GCPs	4198	525	N/A	N/A

* All MIPS values are created using zPCR with the reference CPU set to a 2084-301 rated at 450 MIPS and using the LoIO workload

Common Capacity Planning Mistakes

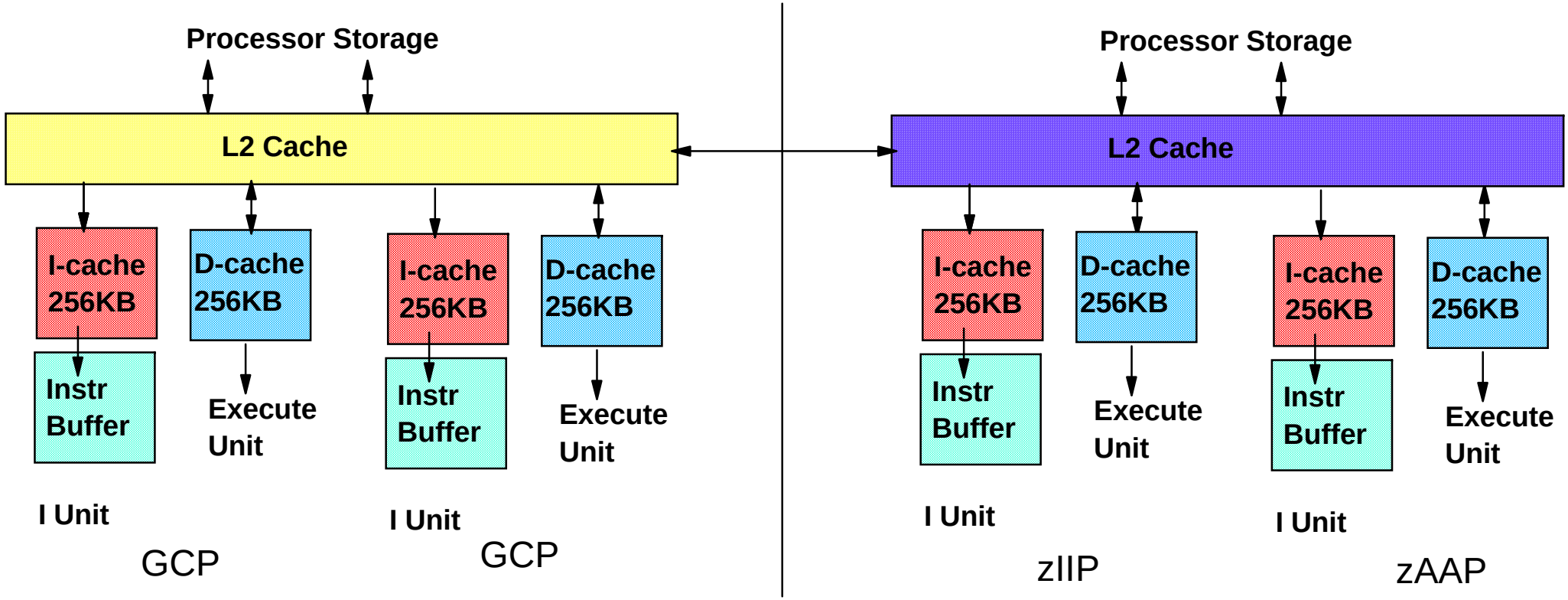
- Need to look across time for zXXP eligible work, and understand interval which drives need for processor capacity
- Make sure the peak which drives capacity can use the zXXP CPs to reduce costs





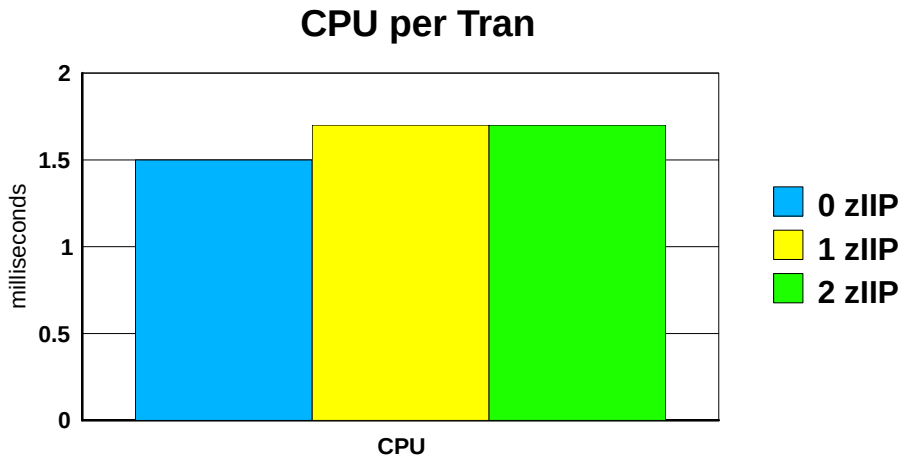
CPU per Transaction with Specialty CPs

- May see an increase in CPU time as work switches between processor types
- Recorded as increased CPU per transaction
  - Impacts may vary based on workloads, arrival rates, switch rate, cache hit rates, and book configuration
- Increased amounts of L2 Cache may provide performance benefits

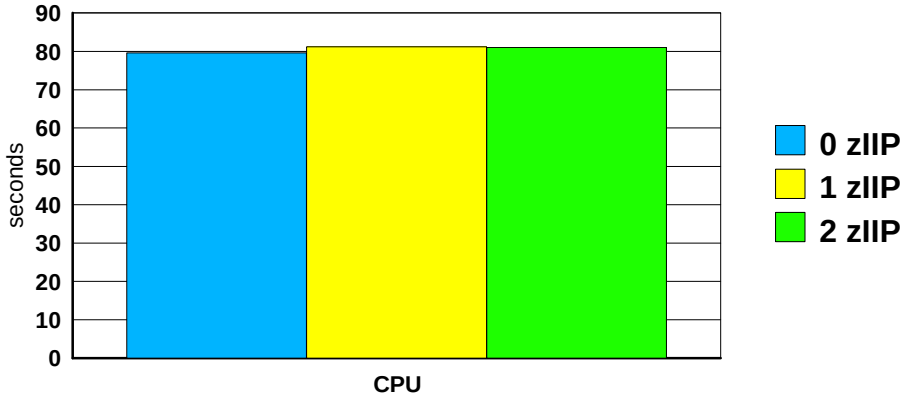


CPU per Transaction Examples

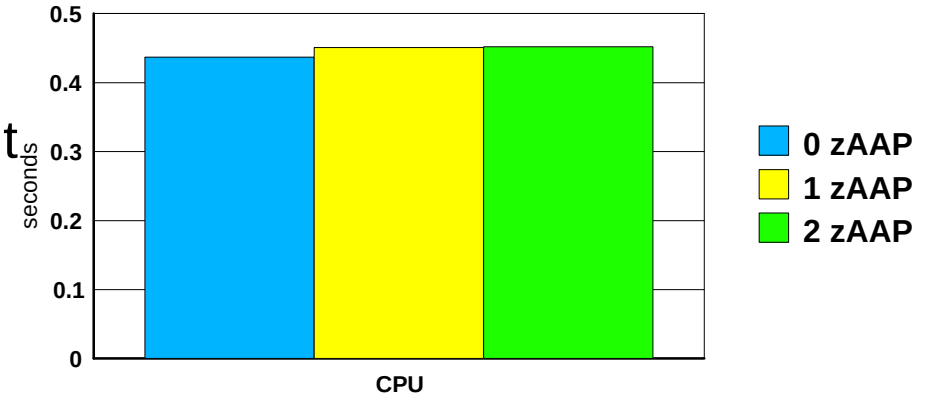
- DRDA work on a zIIP - Trivial CPU content



- DB2 Utility on a zIIP - Complex CPU content

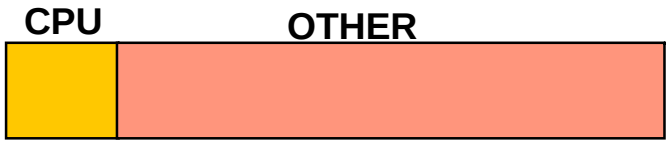


- CICS/Java on a zAAP - Average CPU content

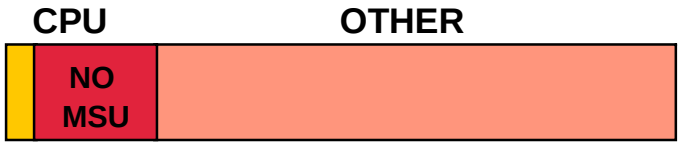


CPU per Tran: Need for a new view

- CPU portion of a transaction is usually such a small part an increase in CPU time doesn't generally impact response times
  - CPU per tran is a capacity metric
  - Higher CPU per tran requires more capacity for the same transaction rate
- Capacity costs with Specialty CPs needs to be managed differently
  - Software costs (Not incurred with Specialty CPs)
  - Hardware and maintenance costs (Incurred with Specialty CPs but at a lower cost)



Traditional View - No Specialty CPs
CPU equal MSU cost

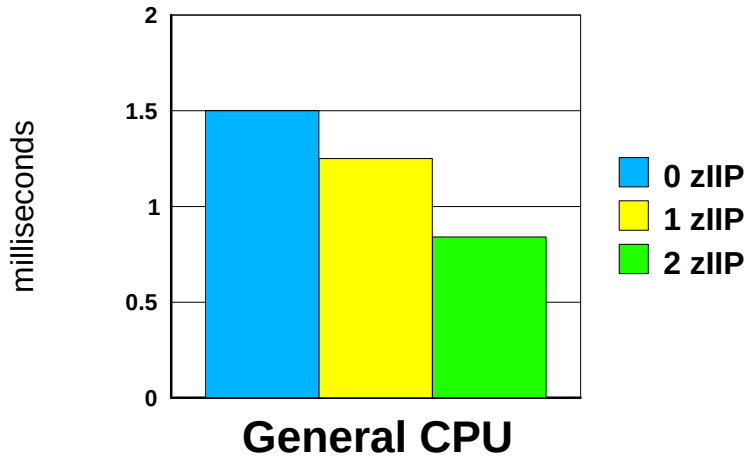


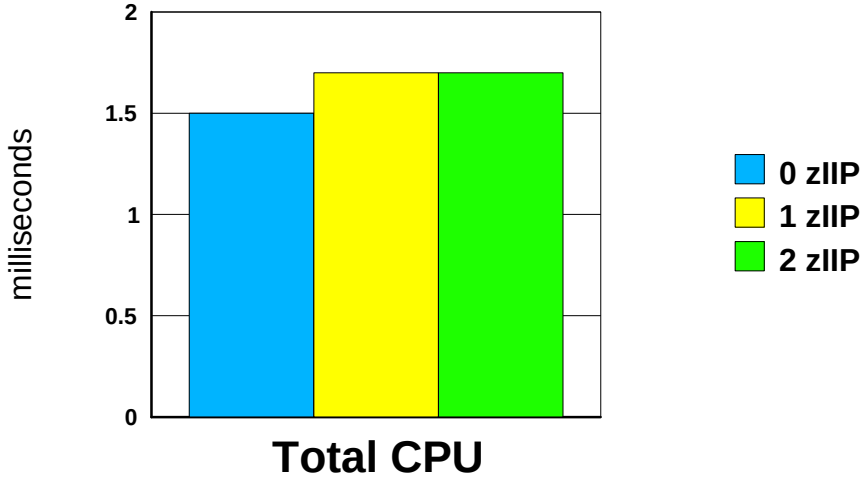
New View - Specialty CPs
CPU not equal MSU cost
CPU per Transaction may go up slightly

Unique View - Full Speed Specialty CPs and Sub-Capacity GCP
CPU not equal MSU cost
CPU per Transaction may go down

CPU per Tran: Need for a new view

- Need to profile work to understand CPU by general GCP time and Specialty CP time
  - Acquire processor capacity with different MSU relationships
  - Comparing capacity across platforms, a MIPS view is mixing cost structures

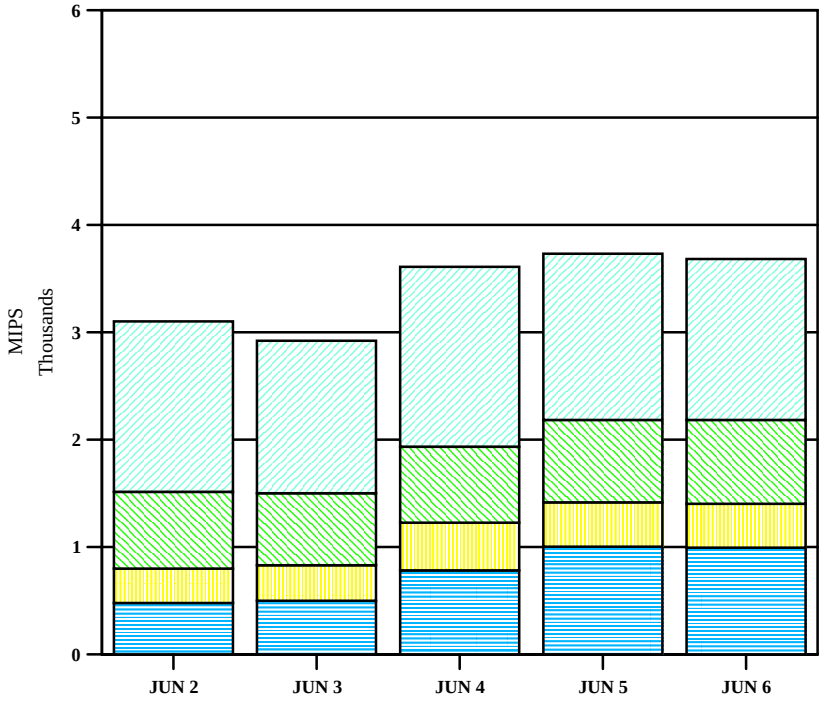




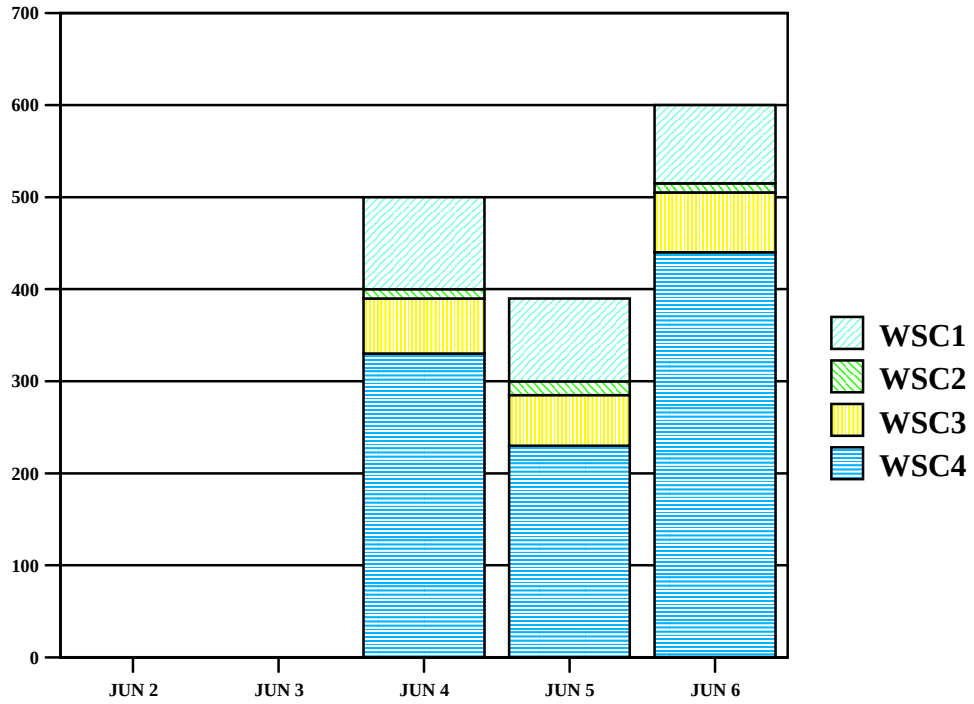
Performance reporting - How Busy is the CEC?

- Need to be aware of the amount of IIPCP and AAPCP offload percentages when building reports
- This capacity has moved from Specialty CP to GCPs
  - Not a workload increase, but rather a shift in CP Type

General CPU Busy by LPAR

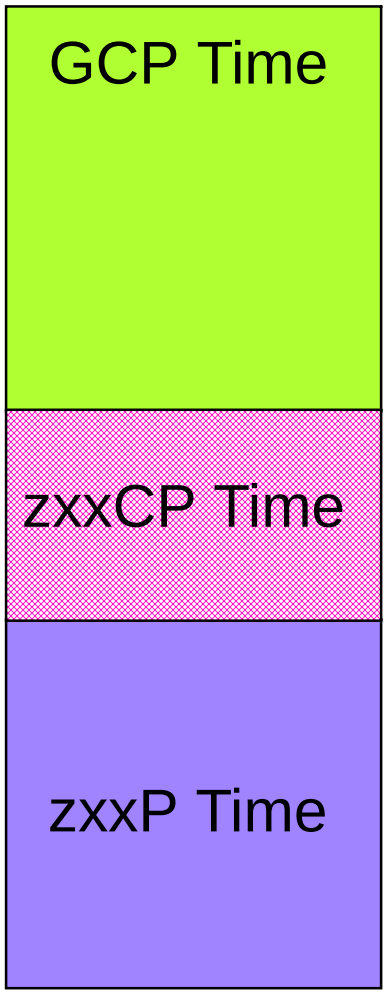


IIPCP CPU Busy by LPAR



Accounting Methods with Specialty CPs

- The issues:
  - How do you recover the zxxP costs
  - How do you handle impacts of zxxP "need help" on billing
    - Not something the application does, it's something the system does to the application
    - Highly variable, hence not repeatable
- Existing CPU fields do not include zxxP time, all new CPU fields for zxxPs
  - Done to ensure proper billing for zxxP Specialty CPs
    - Specialty CPs have different cost structures, (lower cost, don't carry IBM software charges, and have lower maintenance costs)
    - If zxxP time was in SMF30CPT you would mix CPU seconds with different cost structures
  - By not updating SMF30CPT current billing programs do not have to be changed and end users exploiting specialty CPs get the benefits



Summary

- Specialty CPs offer dramatic new capabilities to z/OS workloads
  - Assist Processors seamlessly managed by the operating system and PR/SM
  - Reduced Cost of Ownership
  - Integrated performance and capacity reporting
- Planning for System z workloads requires the use of zPCR tool to get the best view of capacity
  - N-way impacts
  - MP effects
  - Book configurations
- Update Capacity Planning methods to differentiate the CPU profile of work using Specialty CPs