IBM System z Technical Conference

E73 - z/VSE Performance Update

Ingo Franzki – ifranzki@de.ibm.com

March 9, 2007

Trademarks

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

CICS*
DB2*
DB2 Connect
DB2 Universal Database
e-business logo*
Enterprise Storage Server
HiperSockets
IBM*
IBM logo*
IMS
Intelligent Miner
Multiprise*
MQSeries*
OS/390*
S/390*
SNAP/SHOT*
Virtual Image Facility
VM/ESA*
VSE/ESA
VisualAge*
VTAM*
WebSphere*
xSeries
z/Architecture
z/VM
z/VSE
zSeries

* Registered trademarks of IBM Corporation

The following are trademarks or registered trademarks of other companies.

LINUX is a registered trademark of Linus Torvalds

Tivoli is a trademark of Tivoli Systems Inc.

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

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

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

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

Intel is a registered trademark of Intel Corporation.

Disclaimer

The information contained in this document has not been submitted to any formal IBM test and is distributed on an "AS IS" basis without any warranty either express or implied. The use of this information or the implementation of any of these techniques is a customer responsibility and depends on the customer's ability to evaluate and integrate them into the operational environment. While each item may have been reviewed by IBM for accuracy in a specific situation, there is no guarantee that the same or similar results will be obtained elsewhere. Customers attempting to adapt these techniques to their own environments do so at their own risk.

In this document, any references made to an IBM licensed program are not intended to state or imply that only IBM's licensed program may be used; any functionally equivalent program may be used instead.

Any performance data contained in this document was determined in a controlled environment and, therefore, the results which may be obtained in other operating environments may vary significantly. Users of this document should verify the applicable data for their specific environments.

It is possible that this material may contain reference to, or information about, IBM products (machines and programs), programming, or services that are not announced in your country. Such references or information must not be construed to mean that IBM intends to announce such IBM products, programming or services in your country.

Agenda

- **Hardware support**
- **z/VSE V4 Considerations**
  - z/Architecture
  - 64-bit real addressing
  - Paging
- **Sizing a system for z/VSE**
- **Miscellaneous Considerations**

z/VSE 4.1 Hardware support

- **z/VSE 4.1 runs on the following machines**
  - IBM System z9 BC or z9 EC (formerly z9-109)
  - IBM zSeries: z800, z900, z990, z890
  - z/VM V5.2 (or later) is a prerequisite for running z/VSE V4.1 under VM.
- **z/VSE 3.1 and VSE/ESA 2.7 runs on the following machines**
  - IBM System z9 BC or z9 EC (z9-109)
  - IBM zSeries: z800, z900, z990, z890
  - 9672 Parallel Enterprise Server (G5/G6)
  - Multiprice 3000 (7060)
  - equivalent emulators (Flex-ES)

Supported VSE Releases

VSE Release	Available	End of Marketing	End of Service
z/VSE 4.1	03/16/2007		
z/VSE 3.1	03/04/2005		
VSE/ESA 2.7	03/14/2003	09/30/2005	02/28/2007 (out of service)
VSE/ESA 2.6	12/14/2001	03/14/2003	03/31/2006 (out of service)
VSE/ESA 2.5	09/29/2000	12/14/2001	12/31/2003 (out of service)
VSE/ESA 2.4	06/25/1999	09/29/2000	06/30/2002 (out of service)
VSE/ESA 2.3	07/12/1997	06/30/2000	12/31/2001 (out of service)

Running z/VSE V4 under z/VM

- **z/VM V5.2 (or later) is a prerequisite for running z/VSE V4.1 under z/VM**
  - If you IPL z/VSE V4.1 in a guest system of z/VM version 4 or z/VM 5.1, you may experience severe performance problems
  - Because of that the following message is issued during IPL:
    - `0J86I WARNING: VM RELEASE NOT SUPPORTED BY VSE 4.1 - Z/VM 5.2 OR LATER REQUIRED`
  - If you receive this message, you must urgently upgrade your VM system to z/VM 5.2 or a later release.

- **Note: It is not required to run z/VSE under z/VM, you can also run z/VSE in an LPAR**

VSE Server Support

IBM Server	z/VSE 4.1	z/VSE 3.1	VSE/ESA 2.7 (*)	VSE/ESA 2.6 (*)	VSE/ESA 2.5 (*)	VSE/ESA 2.4/2.3 (*)
IBM System z9 BC/EC (z9-109)	Yes	Yes	Yes	Yes (PTF required)	Yes (PTF required)	No
zSeries 890, 990	Yes	Yes	Yes	Yes (PTF required)	Yes (PTF required)	No
zSeries 800, 900	Yes	Yes	Yes	Yes	Yes	Yes
S/390 Parallel Enterprise Server G5/G6	No	Yes	Yes	Yes	Yes	Yes
S/390 Multiprise 3000	No	Yes	Yes	Yes	Yes	Yes
S/390 Parallel Enterprise Server G3/G4	No	No	No	Yes	Yes	Yes
S/390 Multiprise 2000	No	No	No	Yes	Yes	Yes
S/390 Integrated Server	No	No	No	Yes	Yes	Yes
S/390 Parallel Enterprise Server G2 / G1 (out of Service)	No	No	No	Yes	Yes	Yes
ES/9000 – 9221, 9121, 9021 (out of Service)	No	No	No	Yes	Yes	Yes
P/390 and R/390 (out of Service)	No	No	No	Yes	Yes	Yes

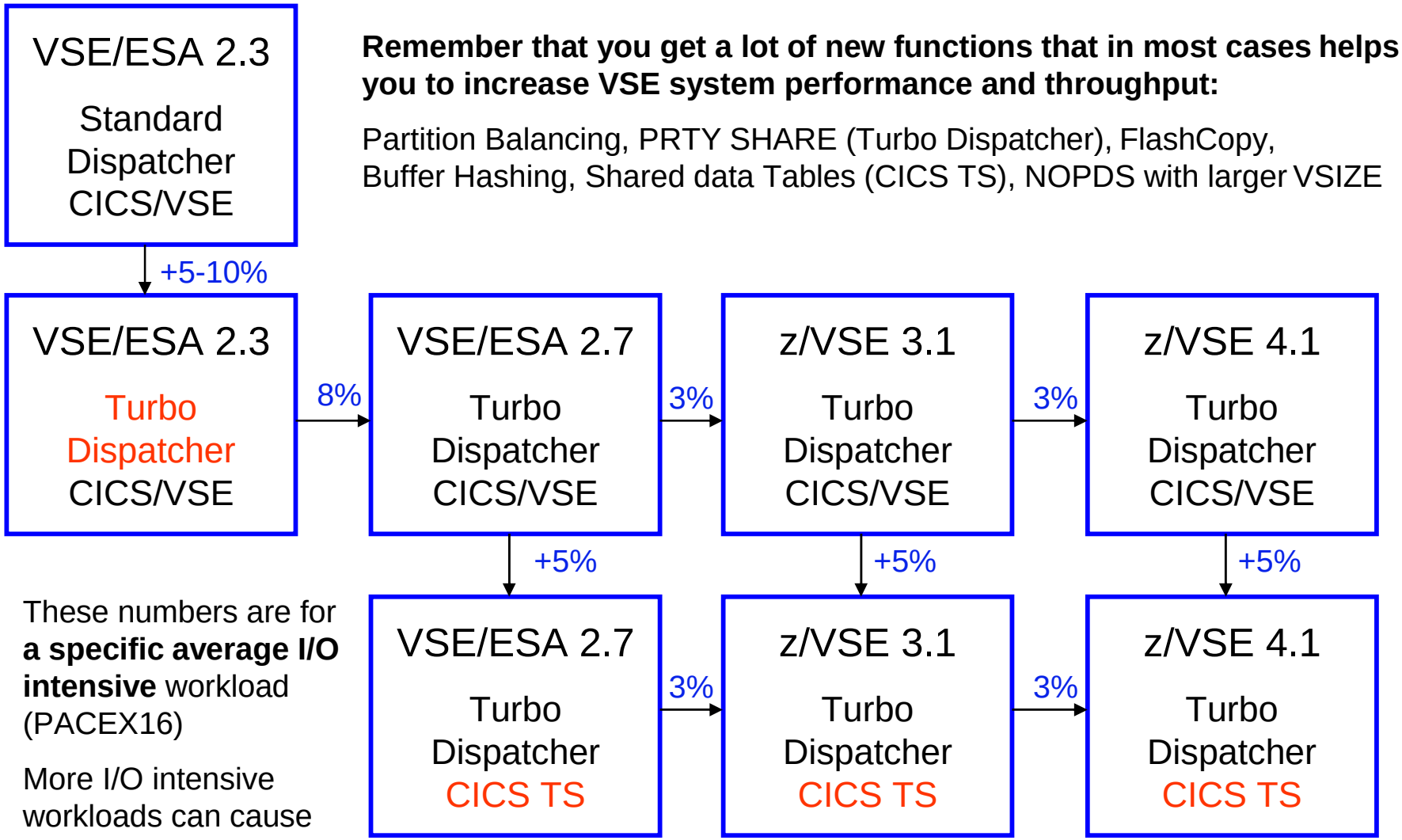
(*) Note: Although VSE/ESA 2.7 or earlier releases technically run on selected servers, these releases are Out-of-Service anyway.

VSE Hardware Support

VSE Release	HiperSockets	OSA Express (QDIO mode)	Hardware Crypto
z/VSE 4.1	Yes	Yes	Yes (PCICA, CEX2C, CEX2A, CPACF)
z/VSE 3.1	Yes	Yes	Yes (PCICA, CEX2C, CEX2A, CPACF)
VSE/ESA 2.7	Yes	Yes	Yes (PCICA, CPACF)
VSE/ESA 2.6	No	Yes	No
VSE/ESA 2.5	No	No	No
VSE/ESA 2.4	No	No	No
VSE/ESA 2.3	No	No	No

Crypto Card	z800	z900	z890	z990	z9 BC/EC
PCICA	No	Yes	Yes	Yes	No
CEX2C	No	No	Yes	Yes	Yes
CPACF	No	No	Yes	Yes	Yes
CEX2A	No	No	No	No	Yes

Overhead Deltas for VSE Releases



Remember that you get a lot of new functions that in most cases helps you to increase VSE system performance and throughput:

Partition Balancing, PRTY SHARE (Turbo Dispatcher), FlashCopy, Buffer Hashing, Shared data Tables (CICS TS), NOPDS with larger VSIZE

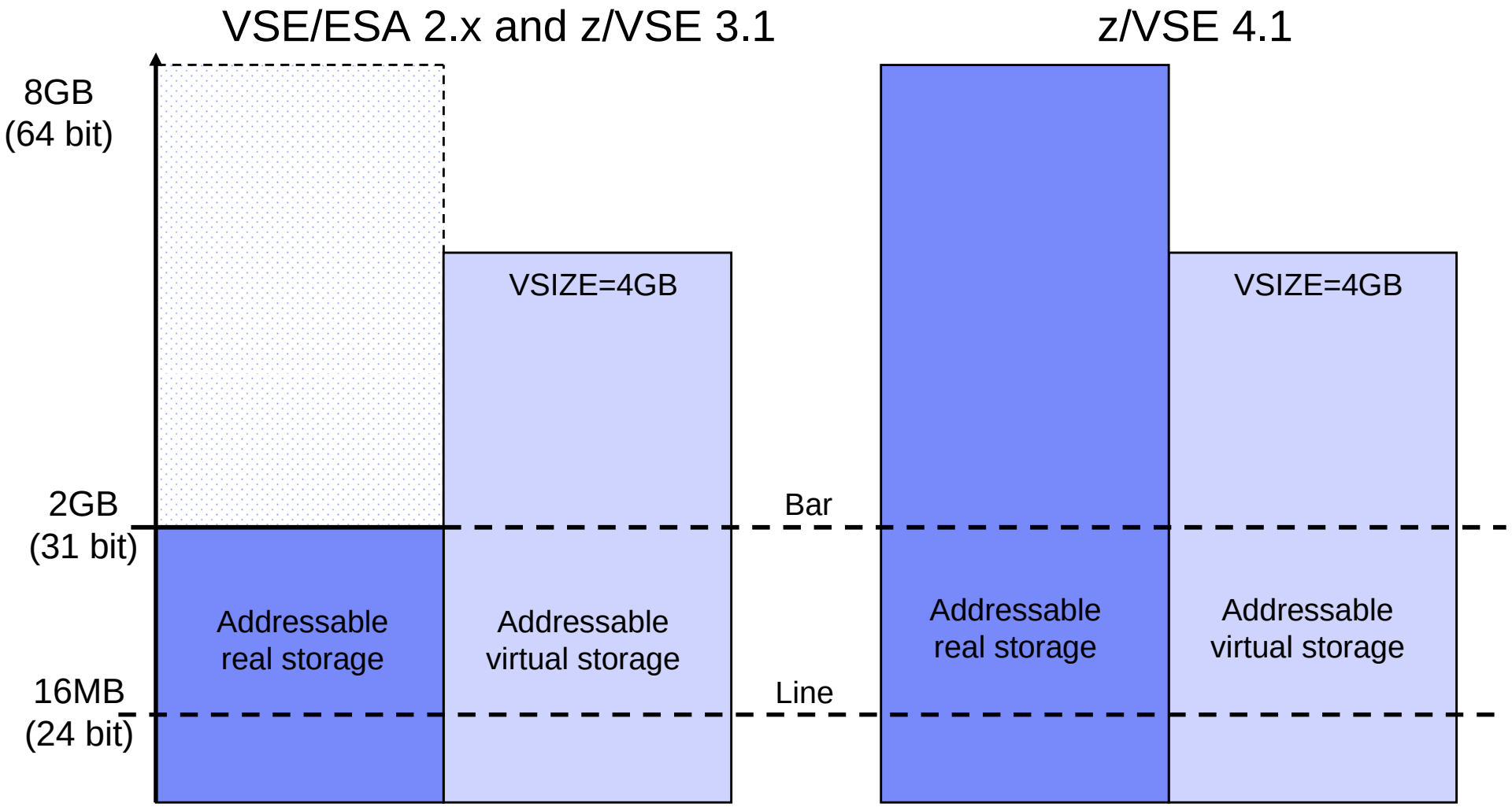
These numbers are for **a specific average I/O intensive** workload (PACEX16)

More I/O intensive workloads can cause more overhead

z/VSE 4.1 - z/Architecture mode

- z/VSE 4.1:
  - Supports z/Architecture-capable (64-bit) processors.
  - Executes in z/Architecture mode only.
  - Supports 64-bit real addressing for selected system functions.
  - Supports processor storage up to 8 GB.
  - The storage beyond 2 GB is managed exclusively by the z/VSE operating system.
- z/VSE 4.1 does not support:
  - 64-bit virtual addressing. The size of a virtual address or data space remains restricted to 2 GB.
  - For user applications, 64-bit addressing and operations that use 64-bit registers.
- 64-bit real addressing is transparent to your user applications providing you use IBM-supplied standard interfaces.
- Customers with especially large z/VSE environments might benefit from lower paging rates.
- Many z/VSE environments might be able to run without a page data set (using the NOPDS option).

What does 64 Bit ‘real addressing’ mean ?



What does 64 Bit ‘real addressing’ mean ?

- **VSE/ESA V2.x and z/VSE 3.1**
  - z/VSE 3.1 or below can address only 2 GB of processor storage
    - 31 bit (real) addressing
  - Page data set required if VSIZE (+VIO) > ~ 2GB
  - No page data set required as long 2 GB processor storage is sufficient
- **z/VSE 4.1**
  - z/VSE 4.1 can address up to 8 GB of processor storage
    - 64 bit (real) addressing
  - No page data set required
    - If processor storage >= VSIZE (+VIO)
  - Virtual address spaces and data spaces are still limited to 2 GB
    - 31 bit (virtual) addressing
    - No changes required to applications

z/VSE 4.1 – Example with 3 GB of real storage

```
map real
AR 0015 AREA        R-SIZE    R-ADDR    PFIX(BELOW)          PFIX(ABOVE)
AR 0015                                 ACTUAL    LIMIT      ACTUAL    LIMIT
AR 0015 SUP         52K       0
AR 0015 SYS-24                          592K      14136K
AR 0015 BG V                            0K        0K         0K        0K
AR 0015 F1 V                            148K      400K       0K        1400K
AR 0015 F2 V                            32K       144K       0K        0K
AR 0015 F3 V                            88K       424K       0K        0K
AR 0015 F4 V                            0K        0K         0K        0K
AR 0015 F5 V                            0K        0K         0K        0K
AR 0015 F6 V                            0K        0K         0K        0K
AR 0015 F7 V                            200K      400K       1044K     2100K
AR 0015 F8 V                            0K        0K         0K        0K
AR 0015 F9 V                            0K        0K         0K        0K
AR 0015 FA V                            0K        0K         0K        0K
AR 0015 FB V                            0K        0K         0K        0K
AR 0015 SYS-31                                                7404K     2052264K
AR 0015 DYN-PA                          0K        0K         0K        0K
AR 0015 AVAIL       64K
AR 0015 SYSTEM      25068K
AR 0015 TOTAL       3145728K                  <----'                 <----'
AR 0015
AR 0015 AVAILABLE FOR SETPFIX:                13544K              2044860K
AR 0015
AR 0015 1I40I READY
```

Exploiting 64 Bit real storage

- **Even on VSE/ESA or z/VSE 3.1 the VSIZE could 'theoretically' be up to 90G**
- **'Practically' you are limited by**
  - Page dataset size and number of extents
  - Page I/O rate
    - Too heavy page I/O rates makes VSE almost unusable
- **With z/VSE 4.1 you can run with VSIZE+VIO up to approximately 8GB without a page dataset (NOPDS)**
  - If enough processor storage is available
  - No time consuming page I/Os
- **You can have more large partitions in your system**

Paging considerations

- **'Paging' is another word for Page Manager activities**
  - Assigning real pages to virtual pages
  - Writing pages to the page data set (page-out I/O)
  - Reading pages from the page data set (page-in I/O)
- **Even with 8GB real storage and NOPDS paging happens**
  - Assigning real pages to virtual pages
  - Moving pages from <2G to >2GB (above the bar) and vice versa
- **'Paging' as such is not bad**
  - But page I/Os are 'bad' (dependent on the page I/O rate)

Paging considerations - PMRMON

- **New SIR command: SIR PMRMON**
  - Displays reports from the ‘Page Manager Monitor’
  - Shows number of page faults, page I/Os, page exchanges (31 <-> 64), ...
- **Usage:**
  - `SIR PMRMON=ON` (resets counters)
  - `// run your workload`
  - `SIR PMRMON`
  - `SIR MPRMON=OFF`
- **Output example, see next foil**
- **The output displays mainly internal counters that are for evaluation by IBM support persons**

Paging considerations - PMRMON

```
AR 0015                  PAGE MANAGER MONITORING REPORT
AR 0015               (BASED ON A 0000:01:04.801 INTERVAL)
AR 0015 IPFQ 31-BIT     =     517126   IPFQ 64-BIT     =     115185
AR 0015  PSQ 31-BIT     =       1500    PSQ 64-BIT     =        527
AR 0015 PF EXCH TOTAL   =       4673   PF EXCH 31->64  =          0
AR 0015 PF EXCH 64->31  =       4673   PGFLT TOTAL     =      25825
AR 0015 PGFLT PMGR      =          0   PGFLT USER      =      25825
AR 0015 PGFLT IMM PO 31 =          0   PGFLT IMM PO 64 =          0
AR 0015 SELCT ON PSQ 31 =          0   SELCT ON PSQ 64 =          0
AR 0015 SELC R=1 MAX 31 =          0   SELC R=1 MAX 64 =          0
AR 0015 RECLAIMS        =          0   NPSQ LOW        =          0
AR 0015 PGOUT I/O TOTAL =          0   PGIN I/O TOTAL  =          0
AR 0015 PGOUT I/O UNC.  =          0   PGOUT I/O PRE.  =          0
AR 0015 LRA PGM CHECK   =       3904   TFIX 64-BIT FR  =        768
AR 0015 1I40I  READY
```

Description of values see next foil

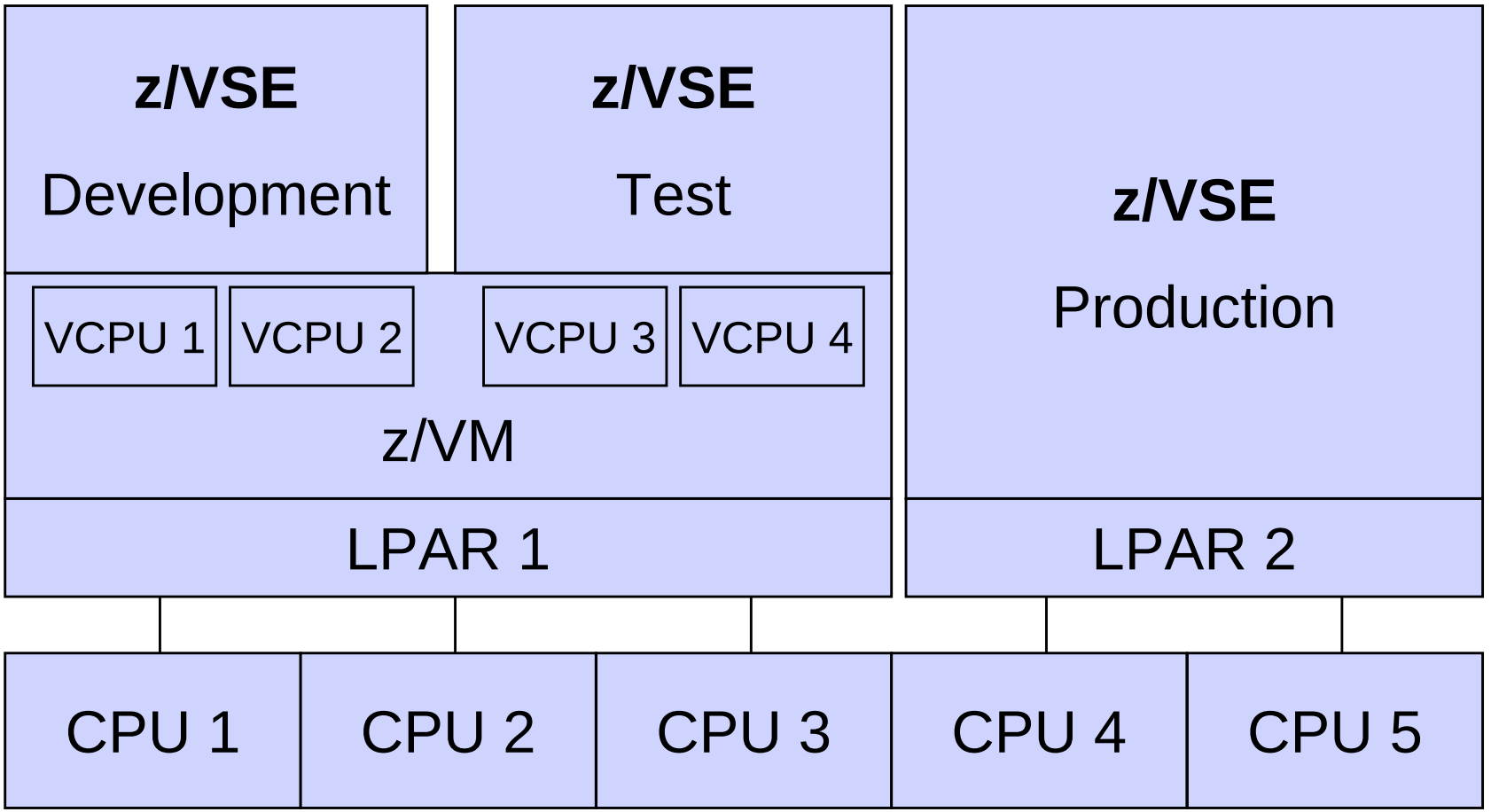
Paging considerations - PMRMON

IPFQ 31-BIT	Number of unused page frames below 2G
PSQ 31-BIT	Number of page frames below 2G in page selection queue
IPFQ 64-BIT	Number of unused page frames above 2G
PSQ 64-BIT	Number of page frames above 2G in page selection queue
PF EXCH TOTAL	Total number of times a page frame exchange is requested
PF EXCH 31->64	Number of times a page frame is exchanged from below 2G to above 2G (e.g. instead of page-out I/O)
PF EXCH 64->31	Number of times a page frame is exchanged from above 2G to below 2G (e.g. TFIX)
PGFLT TOTAL	Total number of page faults
PGFLT PMGR	Number of page faults handled by page manager task
PGFLT USER	Number of page faults handled immediately without activating page manager task
PGFLT IMM PO 31	Number of page faults forcing an immediate page-out of a page frame below 2G
PGFLT IMM PO 64	Number of page faults forcing an immediate page-out of a page frame above 2G
SELCT ON PSQ 31	Number of page selection events on page frames below 2G
SELCT ON PSQ 64	Number of page selection events on page frames above 2G
SELC R=1 MAX 31	Maximum number of cycles on a page selection event for page frames below 2G
SELC R=1 MAX 64	Maximum number of cycles on a page selection event for page frames above 2G
RECLAIMS	Number of times page frames are reclaimed from the page-out queue
NPSQ LOW	Number of times the available page frames are below a critical minimum
PGOUT I/O TOTAL	Total number of page-out I/O
PGIN I/O TOTAL	Total number of page-in I/O
PGOUT I/O UNC.	Total number of unconditional page-out I/O
PGOUT I/O PRE.	Total number of pre-page-out I/O
LRA PGM CHECK	Number of special operation program exceptions on LRA
TFIX 64-BIT FR	Number of TFIX requests with page frames above 2G involved

Sizing a system for z/VSE

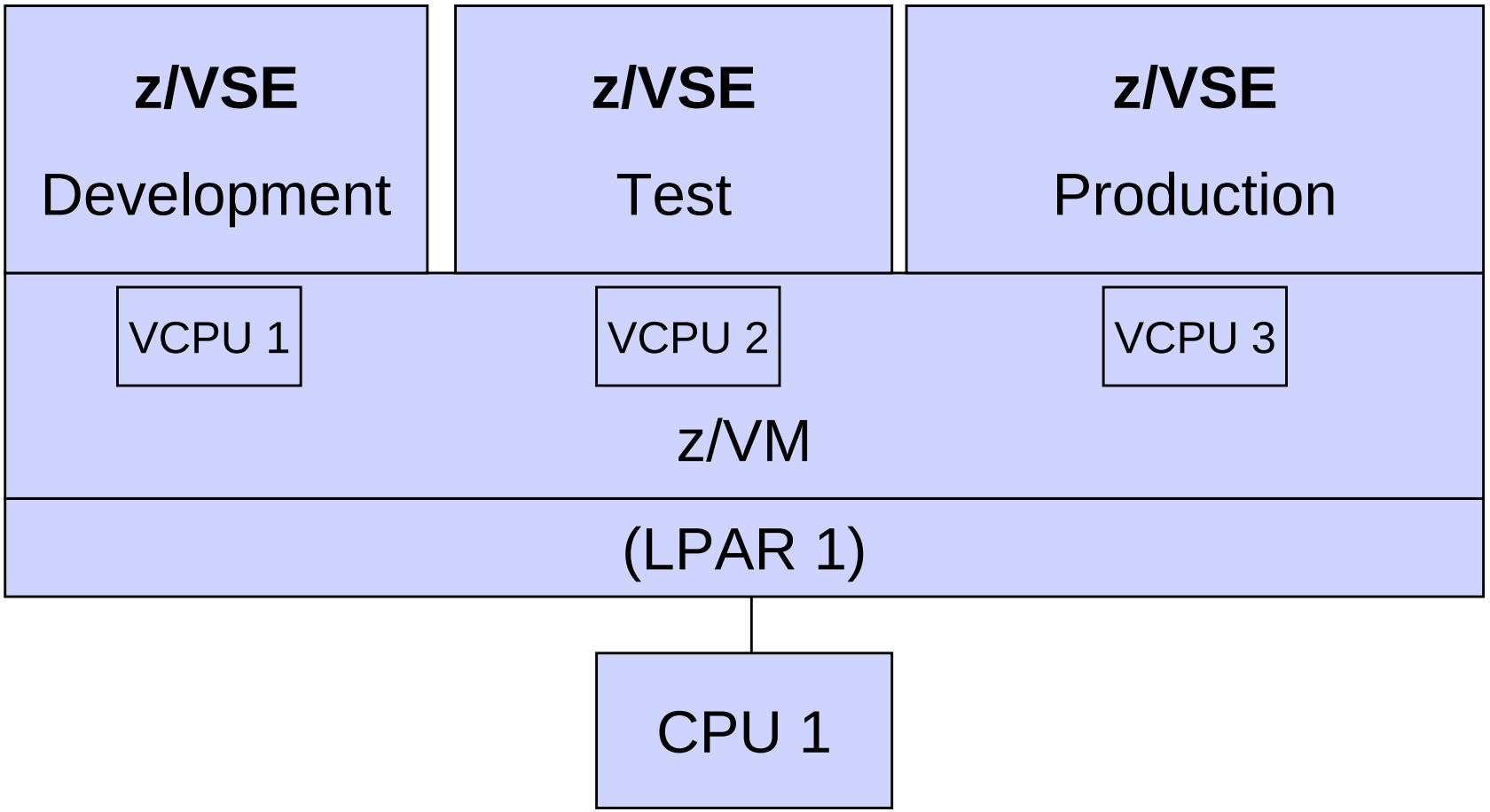
- **Sizing a system for z/VSE is different from sizing a system for z/OS**
  - Although z/VSE supports multiprocessing, z/VSE does not scale as good as z/OS does
    - Do not use more than 3 active processors per z/VSE LPAR or z/VM Guest
- **In general, a faster single CPU is better than multiple smaller CPUs**
  - One partition can only exploit the power of one CPU
    - The largest partition (e.g. CICS) must fit into one single CPU
  - Dependent on nonparallel share (NPS) value
- **Additional CPUs can be useful when multiple LPARs or z/VM Guests are used**
  - Define only up to 3 CPUs per LPAR or z/VM Guest, even if more than 3 CPUs are available on the CEC
- **Do not use MIPS tables for capacity planning purposes**
  - Use zPCR Tool (see page 24) instead with the CB-L workload

Sizing a system for z/VSE – Example 1



- **Although the CEC has 5 CPUs, each VSE only sees 2 CPUs**
- **Test & Development systems have no influence on production system**

Sizing a system for z/VSE – Example 2



- **z/VM does a very good job in sharing a single CPU across multiple guests systems**
- **Use SET SHARE (ABSOLUTE/RELATIVE) command in z/VM to prioritize production systems**

Sizing a system for z/VSE

The fastest uni-processor is (almost always *) the best processor

(*) from a single VSE-image point o view

IBM Processor Capacity Reference for zSeries (zPCR)

- **The zPCR tool was released for customer use on October 25, 2005**
  - **http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/PRS1381**
  - **'As is', no official support, e-mail to zpcr@us.ibm.com**
- **PC-based productivity tool under Windows**
- **It is designed to provide capacity planning insight for IBM System z9 and eServer zSeries processors running various workload environments**
- **Capacity results are based on IBM's LSPR data supporting all IBM System z9 and eServer zSeries processors**
  - Large System Performance Reference: http://www.ibm.com/servers/eserver/zseries/lspr/
- **For VSE use z/OS CB-L workload (similar to CBW2 = Commercial Batch Workload 2)**
  - CB-L fits best to most VSE workloads

New VSE CPU Monitor Tool

- **Intended to help customers to measure the CPU utilization of their VSE system over a period of time.**
- **When you plan for a processor upgrade it is very important to know the CPU utilization of your VSE system over a day or a week.**
  - Helps you to estimate the size of the new processor.
- **The VSE CPU Monitor Tool is not intended to replace any existing monitoring product provided by partners.**
- **It provides only very basic monitoring capabilities on an overall VSE system level.**
- **No details about CPU usage of certain applications are provided**
- **Download**
  - http://www.ibm.com/servers/eserver/zseries/zvse/downloads/tools.html
  - 'As is', no official support, e-mail to zvse@us.ibm.com

New VSE CPU Monitor Tool

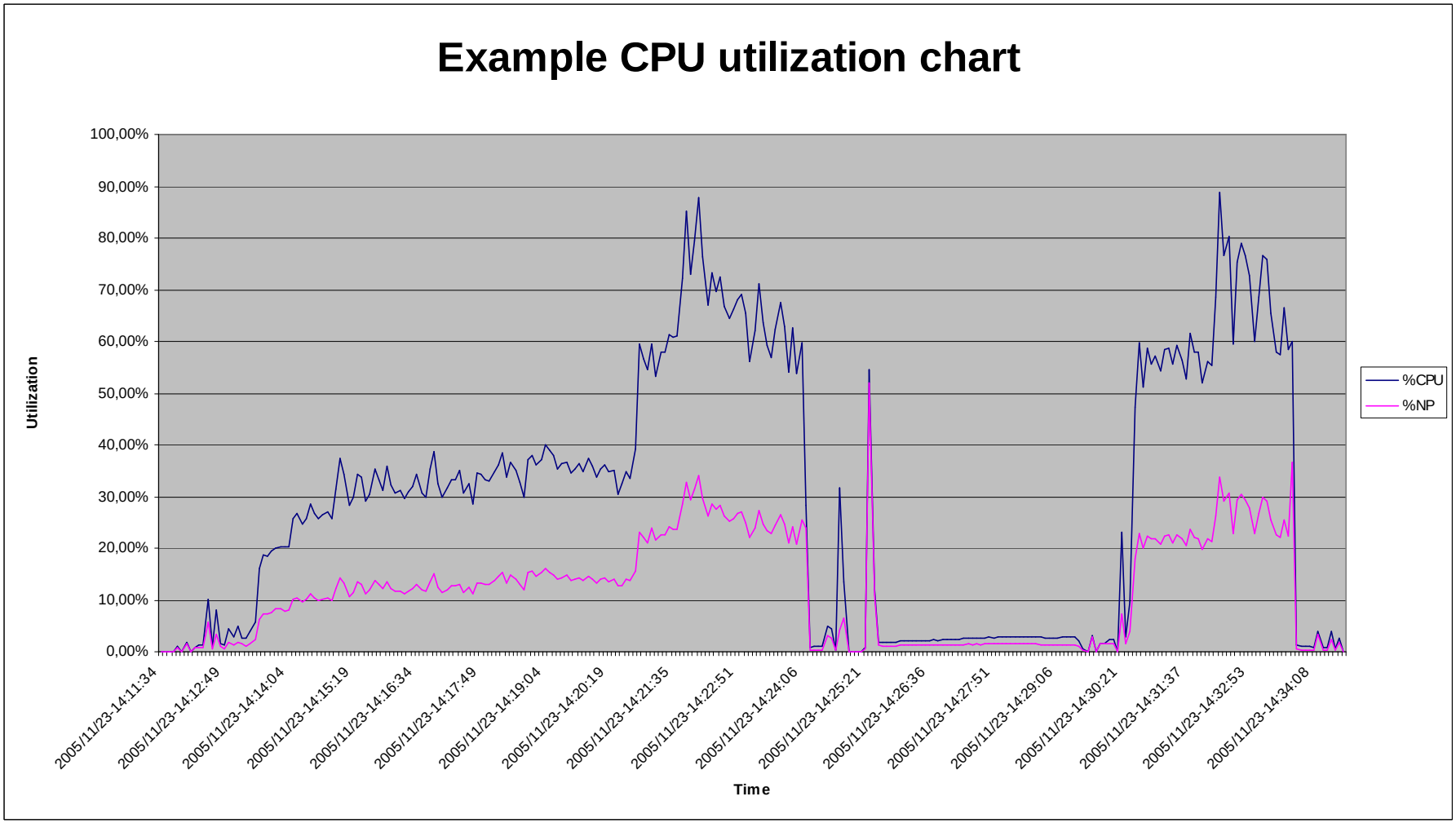
- **CPUMON periodically issues a TDSERV FUNC=TDINFO macro to get performance relevant data.**
- **The data provided by the macro is the same as command QUERY TD shows.**
- **The data from each measurement interval is printed to SYSLST in a comma separated format.**
- **Later on this data can be imported into a spreadsheet (EXCEL)**
- **CPUMON runs in a VSE partition (dynamic or static).**
- **CPUMON is started using:**

```
// EXEC DTRIATTN,PARM='SYSDEF TD,RESETCNT‘
/*
// EXEC CPUMON,PARM='nn‘  nn = interval in seconds
/*
```

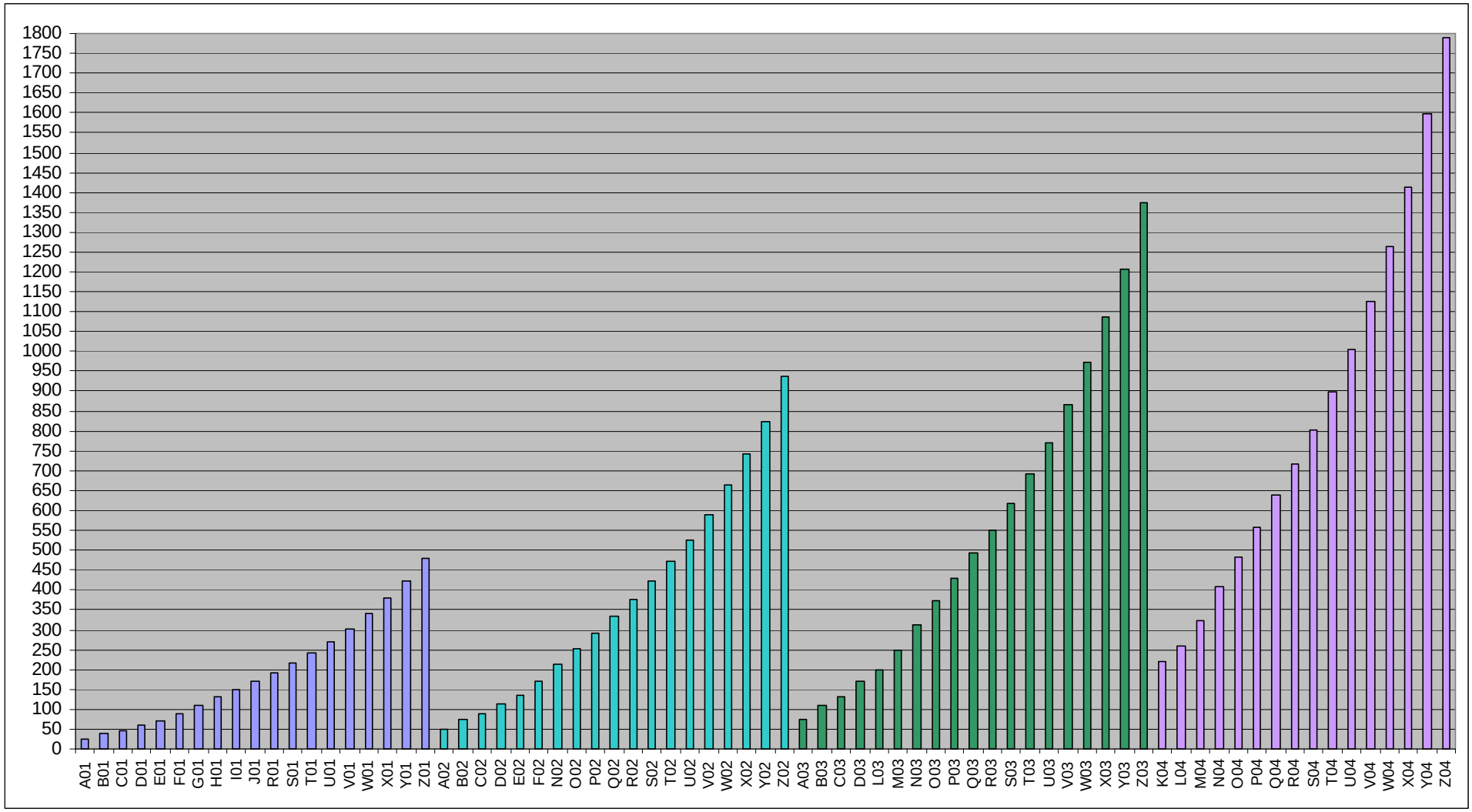
- **The tool can be stopped by entering the following command:**

```
MSG xx,DATA=EXIT          xx = partition id
```

New VSE CPU Monitor Tool

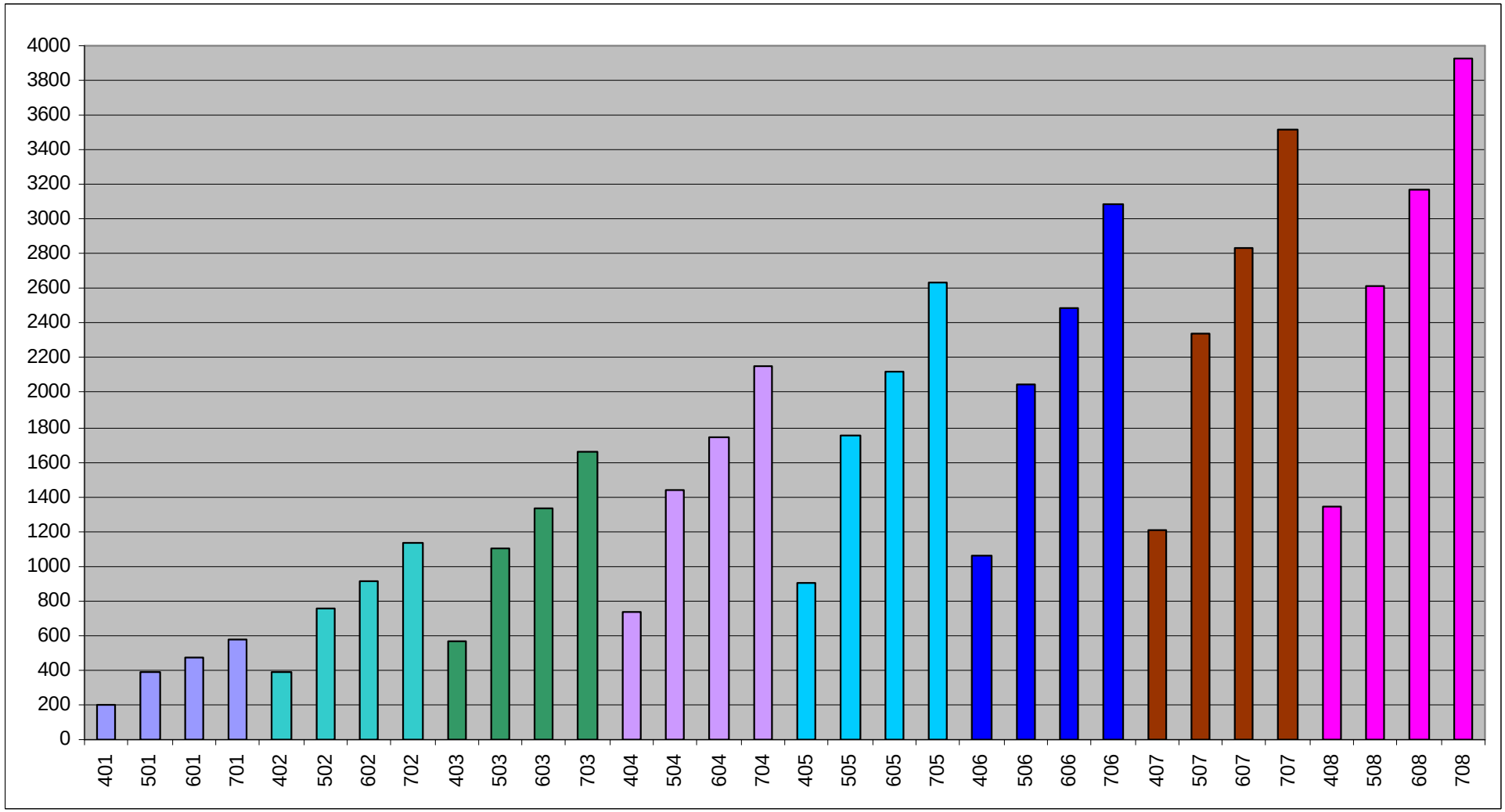


IBM System z9 BC (1-4 CPUs)



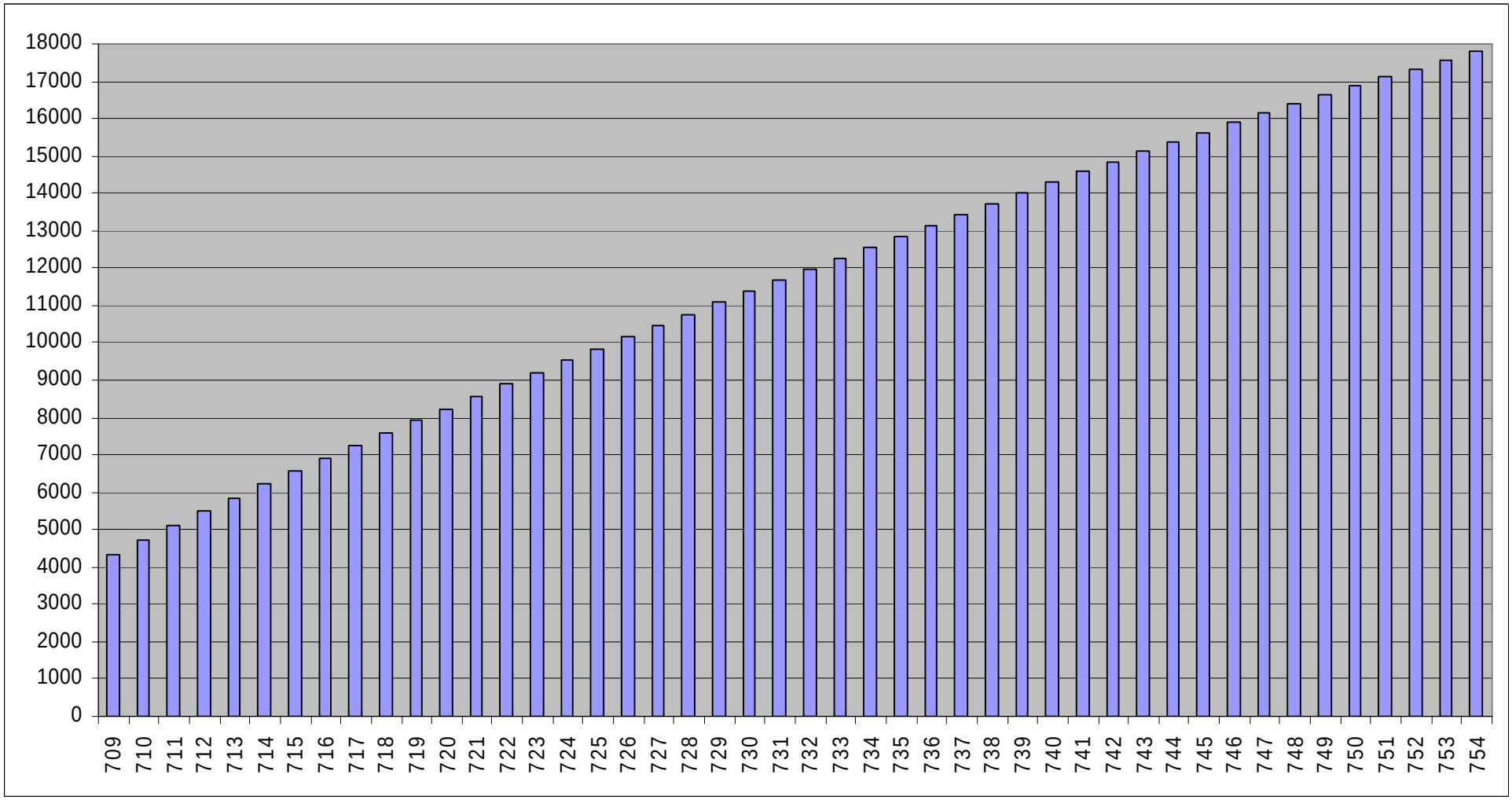
Note: Do not use MIPS to do any kind of capacity planning, use the zPCR tool instead !

IBM System z9 EC subcapacity models (1-8 CPUs)



Note: Do not use MIPS to do any kind of capacity planning, use the zPCR tool instead !

IBM System z9 EC (9-54 CPUs)



Note: Do not use MIPS to do any kind of capacity planning, use the zPCR tool instead !

Midrange Workload License Charge (MWLC)

- **MWLC is a new monthly license charge price metric on the IBM System z9 servers**
- **It applies to z/VSE V4 and 12 key VSE-related middleware programs**
  - such as CICS TS for VSE, ACF/VTAM for VSE, and DB2 Server for VSE.
- **MWLC is only available on z9 EC and z9 BC servers with z/VSE V4.**
- **It is NOT a performance topic**
  - Just for pricing
- **Capacity Measurement Tool**
  - Measures used MSUs (Millions of Service Units) per image (z/VM guest or LPAR)
    - Measurement interval = 30 minutes
    - Calculates 4 hour rolling average
  - Not to be used for performance tuning !
- **For more details see**
  - **IBM System z9 and eServer zSeries Software Pricing:** http://www.ibm.com/servers/eserver/zseries/swprice/
  - **IBM's MSU ratings for the z9 Servers:** http://www.ibm.com/servers/eserver/zseries/library/swpriceinfo/hardware.html

z890, z990 and z9 BC and z9 EC Considerations

- **The z890, z990 and z9 BC and z9 EC (formerly z9-109) are LPAR-only machines**
  - No basic mode any more
  - Even if you run just one VSE system, it now runs in an LPAR
  - Running z/VSE systems under z/VM means
    - running z/VSE in z/VM in an LPAR
  - No I/O Assist in LPARs
    - Only available if z/VM runs in basic mode, but no basic mode available on z890, z990, z9 BC and z9 EC

z/VM V5 considerations

- **z/VM V5 no longer supports V=R and V=F guests**
- **z/VM V5 no longer support I/O Assist**
  - If you currently run with preferred guests, you will need to estimate and plan for a likely increase in processor requirements as those preferred guests become V=V guests as part of the migration.
  - Refer to Preferred Guest Migration Considerations at http://www.vm.ibm.com/perf/tips/z890.html for assistance and background information
- **How to size the impact (on your current system)**
  - **Loss of I/O Assist:** Run your workload with CP SET IOASSIST OFF and measure the increase
  - **Loss of V=R/F:** Run your workload with V=V and use the CP Monitor to watch for increased CPU consumption
- **How to tune**
  - **Dedicated processors:** CP SET SHARE ABSOLUTE
  - **Dedicated memory:** CP SET RESERVED
  - **I/O Assist:** Use minidisks, turn minidisk caching on (MDC)
- **Note: z/VM V5.2 (or later) is a prerequisite for running z/VSE V4.1 under z/VM**

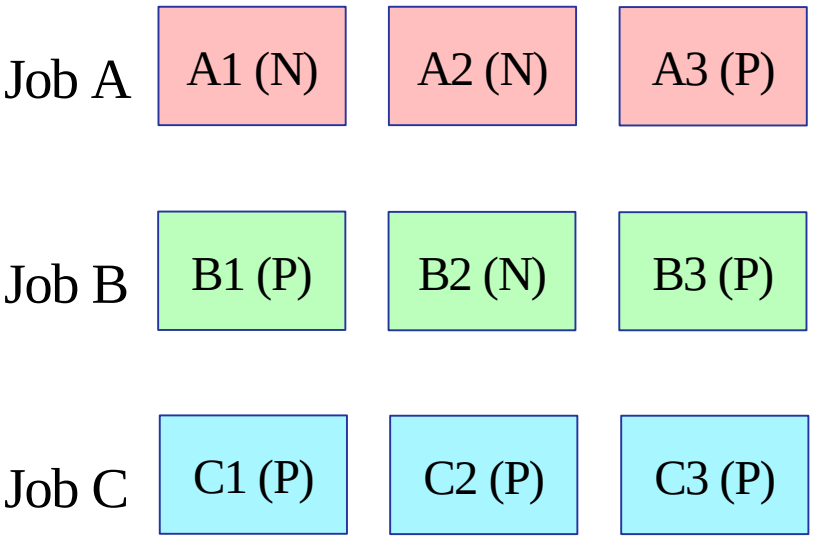
Turbo Dispatcher - Design

- **TD dynamically assigns partitions to CPUs**
  - Work unit = from assignment to one CPU until next interrupt/SVC
  - If one task (subtask) of a partition is active, no other task of the same partition will be selected
  - TD dispatches on partition-basis, not on task-basis
  - A job running in a partition is processed in several work units.

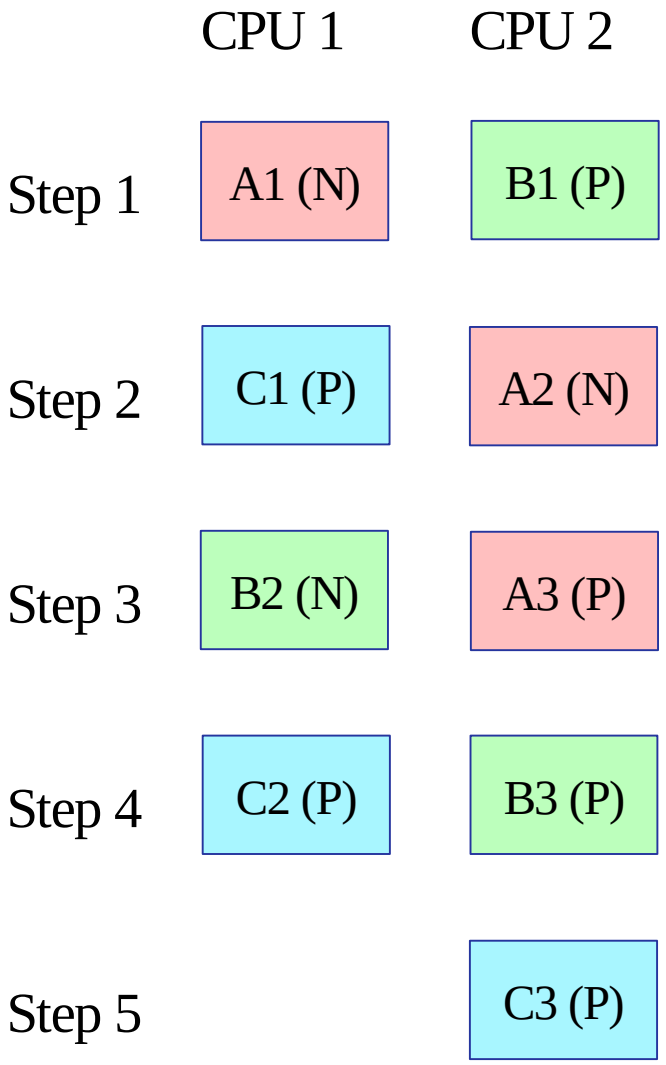
Turbo Dispatcher - Design (2)

- **parallel work units**
  - application code (CICS, Batch)
  - may run on any CPU concurrently with other parallel or non-parallel work units.
- **non-parallel work units**
  - system code (Services, VTAM, Vendor code)
  - As long as one non-parallel work unit is active on one CPU, no other non-parallel work unit can execute on any other CPU.

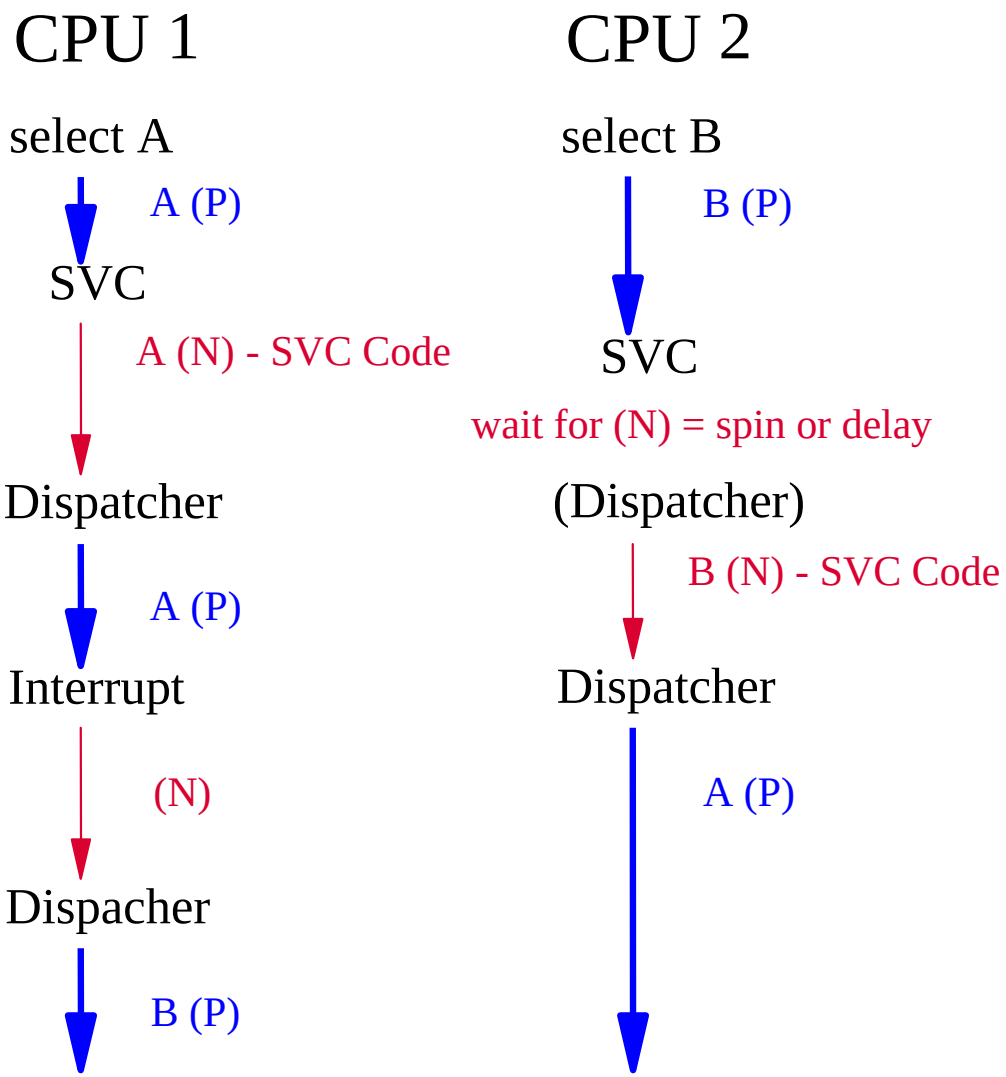
Turbo Dispatcher - Design - Example 1



Ax, Bx, Cx = workunits of job A, B, C
(N) = non-parallel work unit
(P) = parallel work unit



Turbo Dispatcher - Design - Example 2



Turbo Dispatcher - CPU time measurement

- **CPU time measurement (overall system)**
  - SYSDEF TD,RESETCNT
  - Workload (e.g. run a job)
  - QUERY TD (QUERY TD,INTERNAL)

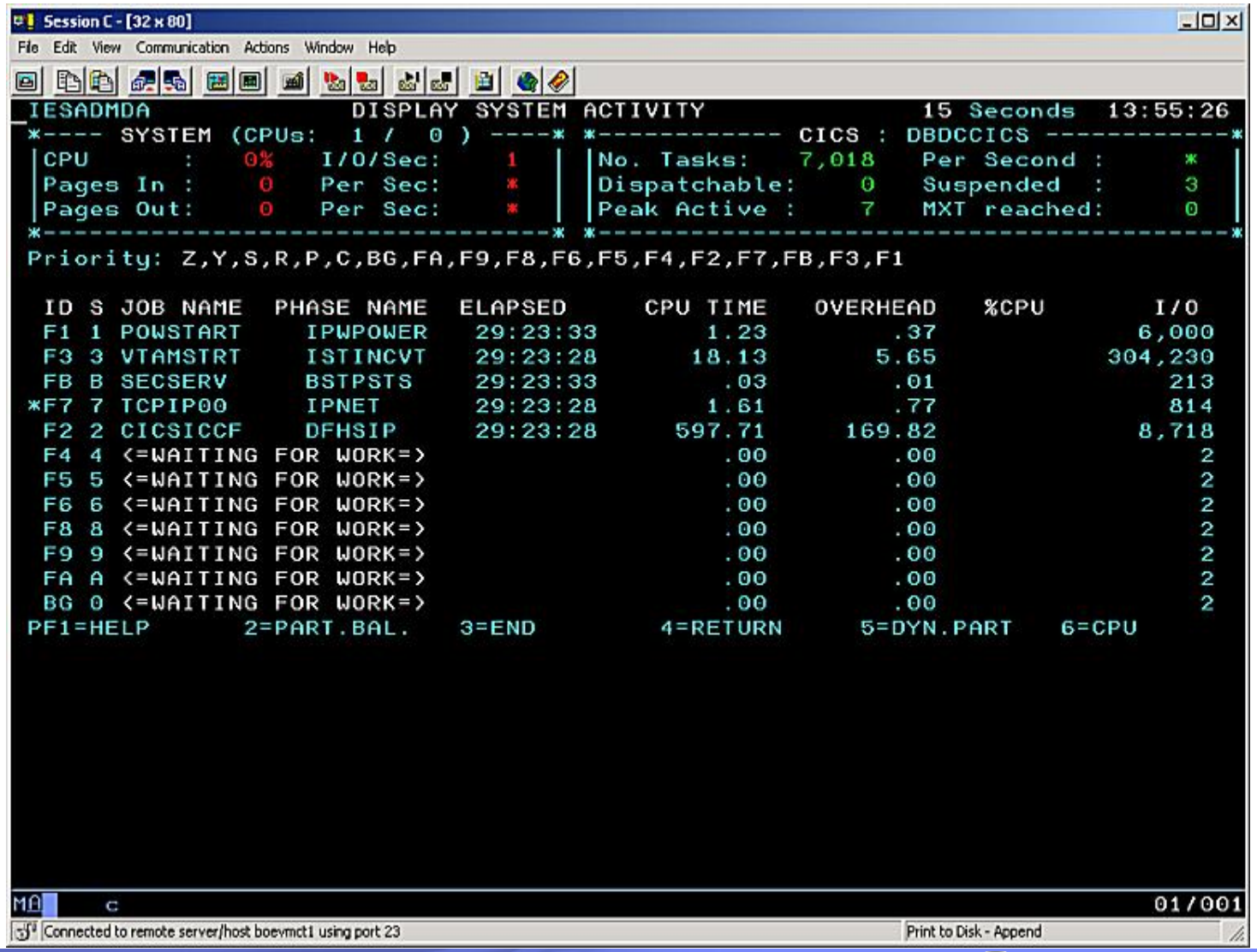
```
 CPU   STATUS     SPIN_TIME     NP_TIME TOTAL_TIME NP/TOT
  00   ACTIVE             0      237100     416698  0.568
  01   ACTIVE             0      157556     415229  0.379
  02   QUIESCED           0           0          0  *.***
  03   INACTIVE
                 ---------------------------------------
TOTAL                     0      394656     831927  0.474

               NP/TOT: 0.474      SPIN/(SPIN+TOT): 0.000
 OVERALL UTILIZATION: 179%        NP UTILIZATION:  85%

 ELAPSED TIME SINCE LAST RESET:       463433
```

NP/TOT = non-paralell share (NPS)
SPIN_TIME = CPU time waiting for NP

Display System Activity Dialog



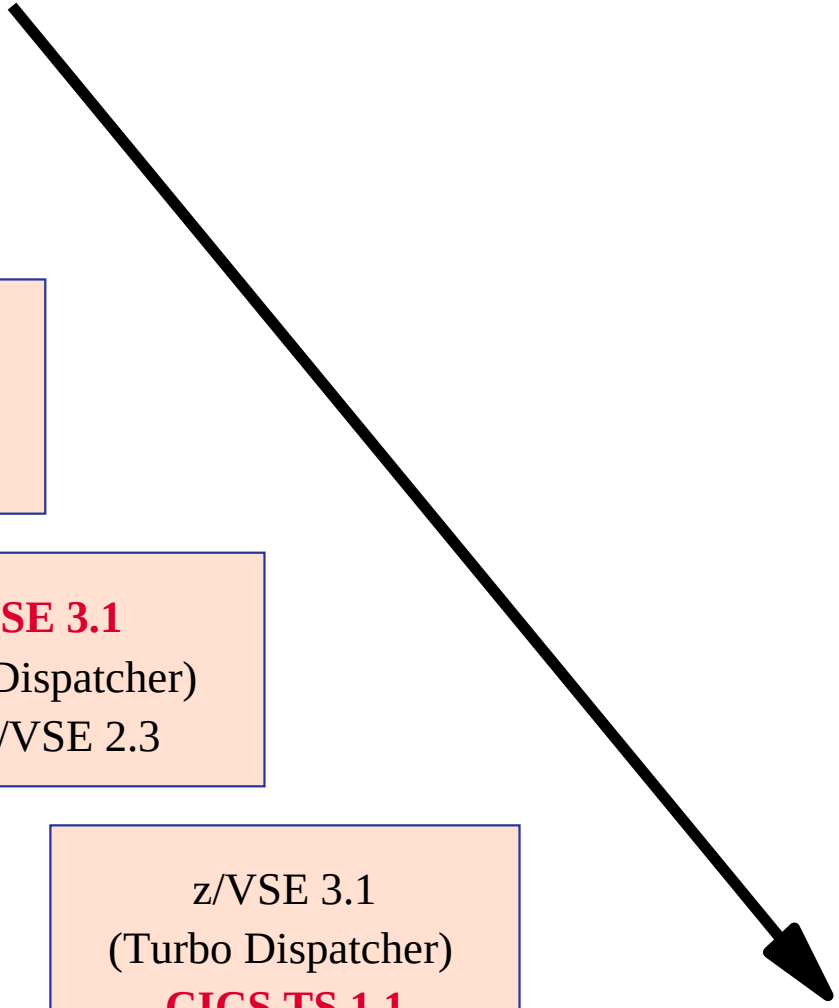
Migration path

VSE/ESA 2.3
Standard Dispatcher
CICS/VSE 2.3

VSE/ESA 2.3
Turbo Dispatcher
CICS/VSE 2.3

z/VSE 3.1
(Turbo Dispatcher)
CICS/VSE 2.3

z/VSE 3.1
(Turbo Dispatcher)
CICS TS 1.1



Change only
one thing at a time!

Allows you to see which step
has introduced a problem.

Performance Tips

- **A partition can only exploit 1 CPU at a time**
  - 2 CPUs do not have any benefit for a single CICS partition
  - Use as many partitions as required for selected n-way
- **Use/define only as many CPUs as really needed**
  - additional CPUs create more overhead, but no benefit
- **Partitions setup**
  - Set up more batch and/or (independent) CICS partitions
  - Split CICS production partitions into multiple partitions (AOR, TOR, FOR)
- **Try to exploit Turbo Dispatcher functions**
  - Priority settings
  - Partition balancing
  - Partition balancing groups

Performance Tips (2)

- **1 CPU must be able to handle all non-parallel workload**
- **Non-parallel code limits the n-Way exploitation**
  - QUERY TD: NP/TOT = NPS (non parallel share)
  - Measure NPS before migration
  - max CPUs = 0.8 - 0.9 / NPS

NPS	#CPUs	NPS	#CPUs
0.20	4.0-4.5 (4)	0.45	1.8-2.0 (2)
0.25	3.2-3.6 (3)	0.50	1.6-1.8 (2)
0.30	2.7-3.0 (3)	0.55	1.5-1.6 (2)
0.35	2.3-2.6 (2)	0.60	1.3-1.5 (1)
0.40	2.0-2.2 (2)	0.65	1.2-1.4 (1)

Performance Tips (3)

- **Non-parallel code limits the maximum MP exploitation**
- **System code (Key 0) increases non-parallel share**
  - Vendor code can have significant impact
- **Overhead increases when NP code limits throughput**
- **Data In Memory (DIM) reduces non-parallel code**
  - less system calls (I/Os)
  - may increase throughput
  - CICS Shared Data Tables
  - Large/many VSAM Buffers (with buffer hashing)
  - Virtual Disks
- **Change VSE/POWER startup to WORKUNIT=PA**
- **Switch tracing/DEBUG off for production**

zSeries Remarks – Split cache

- **Prior to zSeries there is one cache for data and instructions**
- **zSeries has split data and instruction cache**
- **Performance implications:**
  - If program variables and code that updates these program variables are in the same cache line (256 byte)
    - Update of program variable invalidates instruction cache
    - Performance decrease if update is done in a loop
  - See APAR PQ66981 for FORTRAN compiler

zSeries Remarks – Split cache - example

Killer example:

```
*   prepare length
BCTR   R2,0   ADJUST FOR SS-INSTR.
STC    R2,*+5
MVC    RECEIVER(*-*),SENDER
```

STC instruction modifies the next instruction to set the length.

Better code:

```
*   prepare length
BCTR   R2,0  ADJUST FOR SS-INSTR.
EX     R2,MVC01

...

MVC01   MVC    RECEIVER(*-*),SENDER
```

Use EXECUTE instruction instead.

zSeries Performance: Processor Design Considerations:
http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/FLASH10208

zSeries Remarks – Split cache - example

Not causing a problem:

```
         LA    R1,PHASNAME   POINT AT PHASE NAME
         CDDELETE (1)
+*       SUPERVISOR - CDDELETE - 5686-032-06
+        CNOP  0,4
+        BAL   15,*+8
+        DC    A(B'00010010')
+        L     15,0(,15)
+        SVC   65            ISSUE SVC FOR CDDELETE
         DS    0H
```

CDDELETE uses an inline flag byte, but does not modify it

Can cause a problem:

```
         WTO TEXT=DATA
+        CNOP  0,
+        BAL   1,IHB0003A   BRANCH AROUND MESSAGE
+        DC    AL2(8)                TEXT LENGTH
+        DC    B'0000000000010000'   MCSFLAGS
+        DC    AL4(0)              MESSAGE TEXT ADDR
         ...
+IHB0003A DS    0H
+        LR    14,1          FIRST BYTE OF PARM LIST
+        SR    15,15         CLEAR REGISTER 15
+        AH    15,0(1,0)     ADD LENGTH OF TEXT + 4
+        AR    14,15         FIRST BYTE AFTER TEXT
+        LA    15,DATA       LOAD TEXT VALUE
+        ST    15,4(0,1)     STORE ADDR INTO PLIST
+*       SUPERVISOR - SIMSVC - 5686-032
         ...
+        SVC   35            ISSUE SVC 35
 @GE00016 DS    0H
```

WTO uses an inline parameter list, but modifies the parameter list

Note: WTO can be coded with an external parameter list: WTO …,MF=(E,addr)

Possible performance issues with PPRC

- **Issue may occur if**
  - PPRC is used
  - VSE runs in native or in LPAR
  - Not all devices that are defined in IOCP are also defined in VSE ADD statements
- **In case there is an PPRC state change, interrupts are sent to all LPARs where the related device are defined in IOCP.**
  - If the device is defined in VSE ADD, no problem occurs: VSE will process the interrupt correctly.
  - If the device is NOT defined in VSE ADD, the interrupt is ignored by VSE and the interrupt is resent very quickly to that LPAR
    - Results in very high channel activity (up to 100%)
- **Solution:**
  - Define ALL devices in VSE ADD that are defined in IOCP

VSE/POWER POFFLOAD Performance Issues

- **Caused by incompatibility between VSE/POWER tape format and new tape drives**
- **3490F empties cache for FSF used by POFFLOAD LOAD**
  - Install DY46164/DY46245 for VSE/ESA 2.7/2.6
- **3590 synchronizes cache with tape for each WTM**
  - Install microcode FC0520 on A60 controller + VSE/AF APAR DY45817 + AR command TAPE WTM=NOSYNC
  - Unfortunately controller A50 is to small to install FC0520

New POWER POFFLOAD Append function

- **New APPEND option (new with z/VSE 4.1):**
  - `POFFLOAD {BACKUP|SAVE|PICKUP},`
    `queue,tapeaddr,density,class,APPEND...`
  - WAVV requirement WAVV200433
  - Requires 3592 tape unit
- **Great performance benefit when using 3592 tape unit:**
  - Spool time dramatically reduced when using 3592 tape unit (spool time cut half)
  - No wasting of time for skipping existing spool entries on tape when APPENDing using “Space End of Data” command

Tape unit	Search for End-Of-Data to begin APPENDing new Offload Entries	Time to Spool 206 new entries at End-of-data
3592	About 1 second	102 seconds
2490E	307 seconds (not supported)	204 seconds

VTAM 31 Bit I/O Buffer Support

- **VSE/VTAM is providing new 31 bit IO buffer support via PTF**
  - Removes the 24 bit I/O buffer restriction.
  - Moves the I/O buffer pool and all I/O CTC packing buffers above the 24 bit line
  - Support I/O operations in 31 bit mode (using format 1 CCWs).
  - Allow z/VSE customers to grow their communications workloads associated with their business critical applications
- **PTFs**
  - VTAM support - APAR DY46471 - PTF UD52964
  - z/VSE support - APAR DY46396 - PTF UD52873 (AF Base) or UD52874 (Generation Feature)

VTAM 31 Bit I/O Buffer Support

- **To enable the 31 Bit support:**
  - Make sure both PTFs (VTAM and AF) are applied
  - Set VTAM start option IOBUF31=YES (default is NO)
- **Display settings**

```
d net,vtamopts,opt=iobuf31
AR 0015 1C39I COMMAND PASSED TO ACF/VTAM
F3 0003 IST097I DISPLAY ACCEPTED
F3 0003 IST1188I ACF/VTAM V4R2 STARTED AT 17:11:29 ON 11/11/05
F3 0003 IST1349I COMPONENT ID IS 5686-06501-FE6
F3 0003 IST1348I VTAM STARTED AS INTERCHANGE NODE
F3 0003 IST1497I VTAM FUNCTIONAL SUPPORT LEVEL IS INTERENTERPRISE
F3 0003 IST1189I IOBUF31 = YES
F3 0003 IST314I END
```

Documentation

- **z/VSE homepage:**
  - http://www.ibm.com/servers/eserver/zseries/zvse/
- **VSE Performance:**
  - http://www.ibm.com/servers/eserver/zseries/zvse/documentation/performance.html
- **z/VM homepage:**
  - http://www.ibm.com/vm
- **z/VM Performance:**
  - http://www.vm.ibm.com/perf/
- **z/VM Preferred Guest Migration Considerations**
  - http://www.vm.ibm.com/perf/tips/z890.html
- **IBM System z9 and eServer zSeries Software Pricing**
  - http://www.ibm.com/servers/eserver/zseries/swprice/
- **IBM's MSU ratings for the z9 Servers**
  - http://www.ibm.com/servers/eserver/zseries/library/swpriceinfo/hardware.html

Questions ?