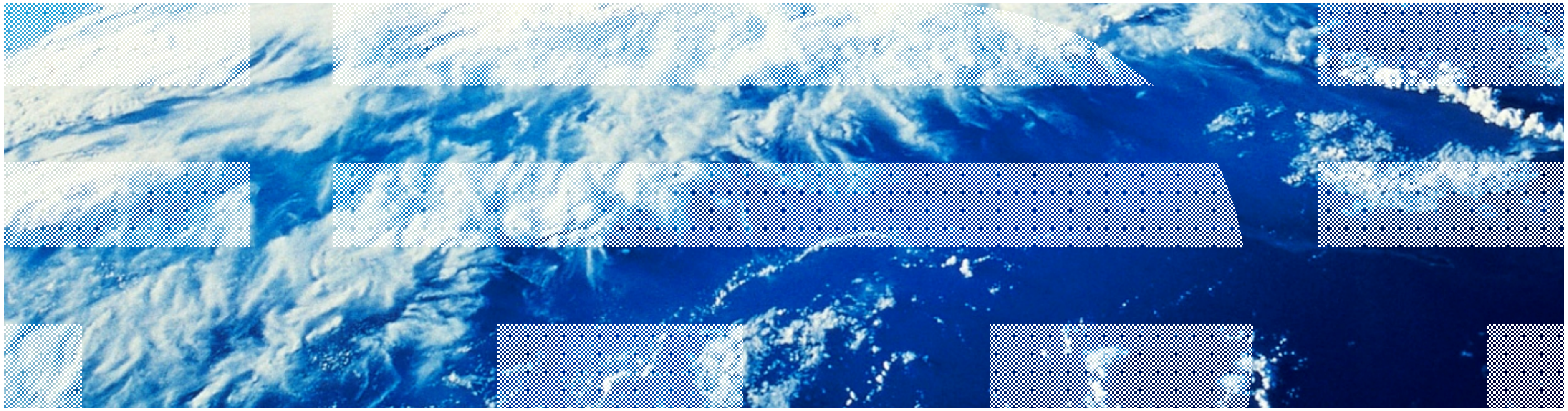
z/VSE V4.3 Performance Update

Ingo Franzki, IBM



Trademarks

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

Not all common law marks used by IBM are listed on this page. Failure of a mark to appear does not mean that IBM does not use the mark nor does it mean that the product is not actively marketed or is not significant within its relevant market.

Those trademarks followed by ® are registered trademarks of IBM in the United States; all others are trademarks or common law marks of IBM in the United States.

For a complete list of IBM Trademarks, see www.ibm.com/legal/copytrade.shtml:

*, AS/400®, e business(logo)®, DBE, ESCO, eServer, FICON, IBM®, IBM (logo)®, iSeries®, MVS, OS/390®, pSeries®, RS/6000®, S/30, VM/ESA®, VSE/ESA, WebSphere®, xSeries®, z/OS®, zSeries®, z/VM®, System i, System i5, System p, System p5, System x, System z, System z9®, BladeCenter®

The following are trademarks or registered trademarks of other companies.

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.
Cell Broadband Engine is a trademark of Sony Computer Entertainment, Inc. in the United States, other countries, or both and is used under license therefrom.
Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.
Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.
Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.
UNIX is a registered trademark of The Open Group in the United States and other countries.
Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.
ITIL is a registered trademark, and a registered community trademark of the Office of Government Commerce, and is registered in the U.S. Patent and Trademark Office.
IT Infrastructure Library is a registered trademark of the Central Computer and Telecommunications Agency, which is now part of the Office of Government Commerce.

* 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.

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

- **z/VSE V4.3 Performance Considerations**
  - Release Overhead Deltas
  - Support for 4 digit device addresses
  - GETVIS constraint relief
  - Dynamic start/stop of CPUs
  - 1 MB frames for data spaces
  - Fast Path to Linux on System z
  - Queue-I/O Assist (QIOASSIST)
  - Crypto Express3 and AP queue interrupt support
  - FICON Express8
  - SNMP Agent support
  - VSAM Redirector executed in subtask under CICS
- **z/VM and Linux considerations**
- **Sizing a System for z/VSE**
- **Performance Measurement Tools**

Supported VSE Releases

VSE Release	Available	End of Marketing	End of Service
z/VSE 5.1	preview		
z/VSE 4.3	11/26/2010		
z/VSE 4.2	10/17/2008	11/26/2010	
z/VSE 4.1	03/16/2007	10/17/2008	04/30/2011
z/VSE 3.1	03/04/2005	05/31/2008	07/31/2009 (out of service)
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)

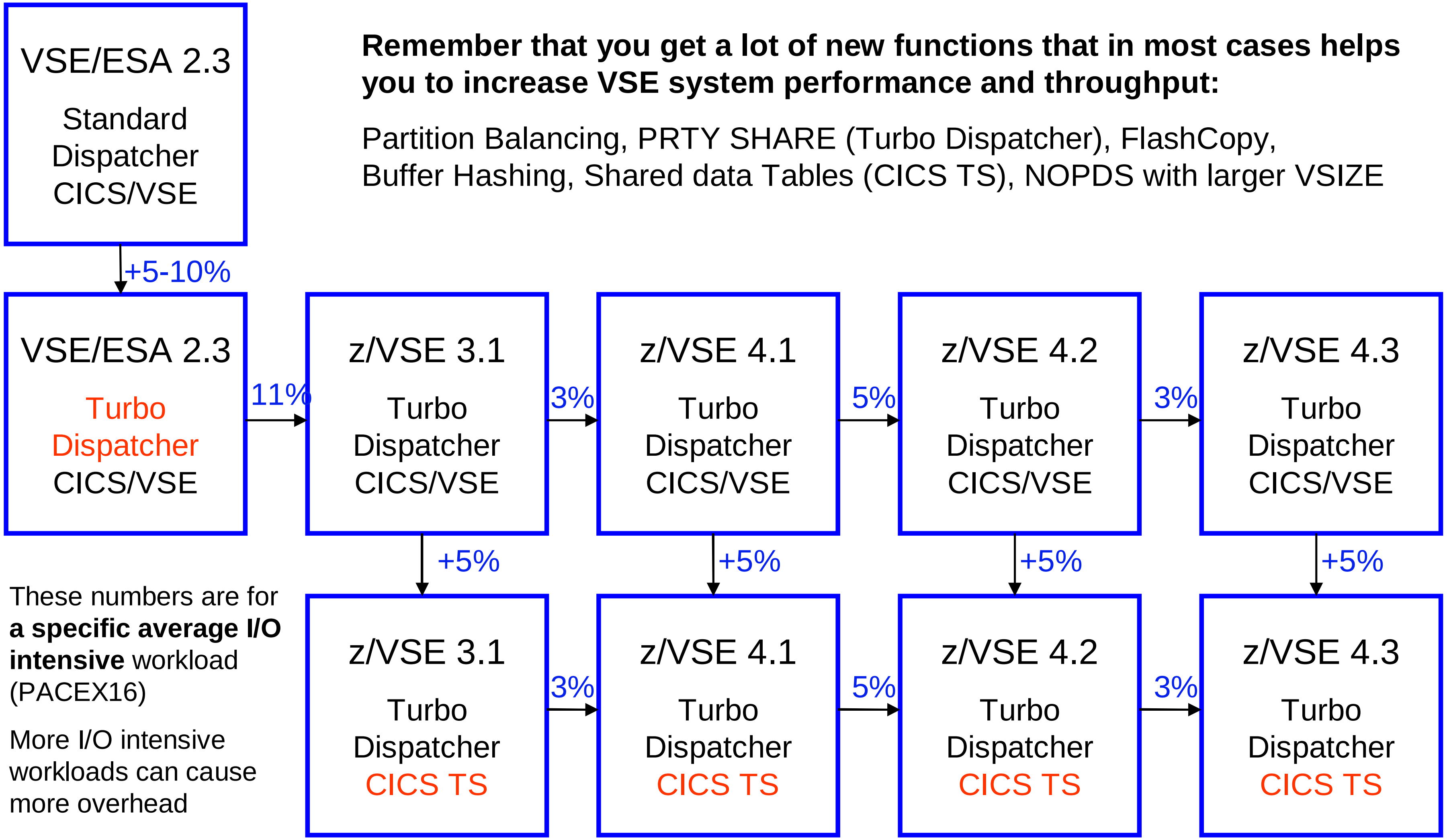
This is very soon!

http://www.ibm.com/systems/z/os/zvse/about/status.html

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

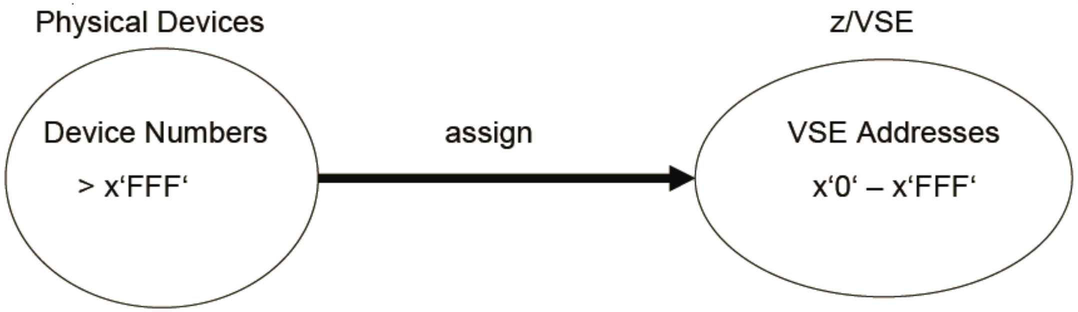
z/VSE V4.3 – Support for 4 digit device addresses

- I/O devices can have physical device numbers in the range of x‘0000‘ to x‘FFFF‘

- z/VSE V4.2 or earlier supported only device addresses from x’000’ to x’FFF’
  - I/O devices with device numbers > x‘FFF‘ were ignored during the z/VSE installation process and cannot be added in the IPL procedure

- z/VSE V4.3 now supports device addresses in the range of x‘0000‘ to x‘FFFF‘
  - I/O devices with device numbers > x‘FFF‘ are recognized during the z/VSE installation process and can be added into the IPL procedure

- If the physical device address of an I/O device is <= x‘FFF‘ nothing changes.
- If the physical device address of an I/O device is > x‘FFF‘ then the user has to assign a so called VSE address to the device



z/VSE V4.3 – Support for 4 digit device addresses

Extended ADD Statement:

```
ADD <physical device address > as <VSE address>,<device type>

ADD <phy_addr1> : <phy_addr2> as <VSEaddr1> : <VSEaddr2>, <device type>

ADD <phy_addr1> .. <phy_addr2> as <VSEaddr1> .. <VSEaddr2>, <device type>
```

Notes:

§ The physical device address must be > x‘FFF‘ if you want to assign a VSE address

§ If the physical device address is lower or equal than x’FFF’ then the VSE address is equal to the physical device address by default. This assignment cannot be changed

§ The VSE address must be unique: You cannot assign the same VSE address twice.

Examples:

§ `ADD 1555 as 555, ECKD`

§ `ADD 1010:1020 as 200:210, 3480`

§ `ADD 1010..1020 as 200..210, 3480`

§ `ADD 300 as 500, ... à NOT allowed !`

z/VSE V4.3 – Support for 4 digit device addresses

Addressing devices in z/VSE Jobs

§ IOCP:

– Control Unit:

```
CNTLUNIT CUNUMBR=8000,PATH=(A1,B1,A2,B2,A3,B3,A4,B4), X
         UNITADD=((00,256)),CUADD=0,UNIT=2105
4 DASDs of type 3390 with the physical device addresses 8000-8003:
IODEVICE ADDRESS=(8000,4),CUNUMBR=(8000),FEATURE=(SHARED),UNIT=3390B
```

§ z/VSE:

– IPL Procedure:

```
ADD 8000:8003 as 800:803,ECKD
```

– Job:

```
// JOB CLRDK
* DLBL DISK,'DISK.FILE.1',1,SD,DSF
// ASSGN SYS012,801
// DLBL UOUT,'DISK.XXXXX',9999
// EXTENT SYS012,,,,2590,5
// EXEC CLRDK
// END
/*
/&
```

In z/VSE Jobs, you always use the 3 digit VSE address !

z/VSE V4.3 – GETVIS constraint relief

§ **I/O Constraint Relief:**

– The z/VSE I/O supervisor routines will run in AMODE(31)

– Depending on the IODEV statement, the control blocks will be allocated either in 24-bit area or in 31-bit area:

- Specifying IODEV=1023 will result in an allocation of the control blocks below the line (24-bit area)
- Specifying IODEV=1024 will result in an allocation of the control blocks above the line (31-bit area)

à *Note: in either way the z/VSE limit of 1024 devices does apply and is as well maintained !*

– *Example:*

```
…
BG 0000 $$A$SUPI,VSIZE=264M,VIO=512K,VPOOL=64K,LOG,IODEV=1024
…
```

– The z/VSE 4.3 system is shipped with IODEV=1024

– If you FSU to z/VSE 4.3, the value remains IODEV=1023

z/VSE V4.3 – Dynamic Starting/Stopping of CPUs

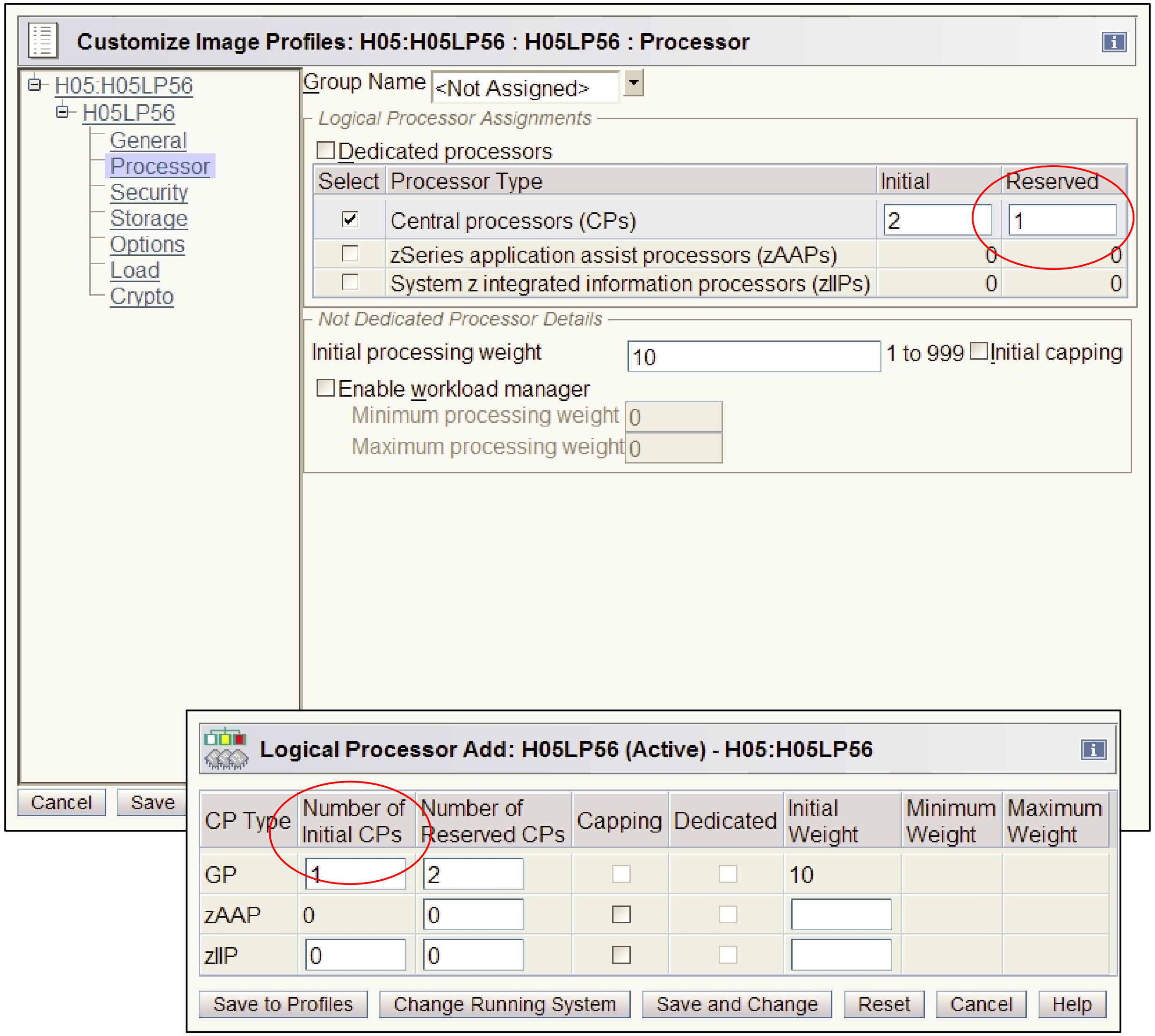
- **z/VSE allows you to start CPUs that were not online at IPL**
- **These can be either:**
  - CPUs that were in a “standby” state at IPL
  - CPUs that were added to the z/VSE LPAR profile after IPL
- Note that z/VSE uses the term CPU to refer to Central Processors (CPs).

- **Using the HMC/SE “Logical Processor Add” task, you can add CPUs to an LPAR after IPL. This feature is available on IBM z10 and z196 platforms**.
  - When CPUs are added to an LPAR, z/VSE automatically updates its CPU configuration.

- **You can then use the:**
  - SYSDEF TD,STARTSBY=cpuaddr command to set CPUs that are in a “standby” state to an “online” state and start these CPUs
  - SYSDEF TD,STOPSBY=cpuaddr command to change the CPU state from “online” to “standby”.
- **These functions allow you to exploit z10 and z196 technology and update the CPU configuration depending on workload needs**

- Note: “Standby” CPUs do not consume any CPU share of their LPAR

z/VSE V4.3 – Dynamic Starting/Stopping of CPUs



Customize Image Profiles: H05:H05LP56 : H05LP56 : Processor

Group Name <Not Assigned>

Logical Processor Assignments

☐ Dedicated processors

Select	Processor Type	Initial	Reserved
☑	Central processors (CPs)	2	1
☐	zSeries application assist processors (zAAPs)	0	0
☐	System z integrated information processors (zIIPs)	0	0

Not Dedicated Processor Details

Initial processing weight 10 1 to 999 ☐ Initial capping

☐ Enable workload manager

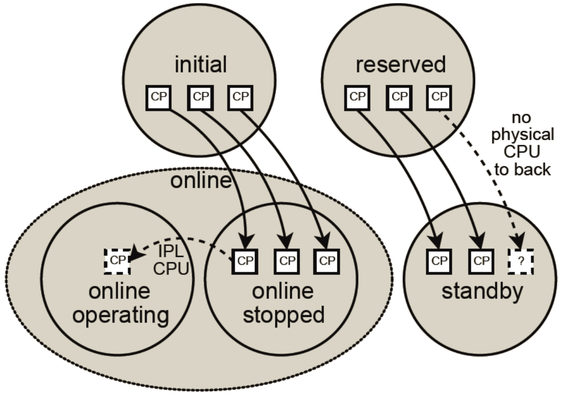
Minimum processing weight 0

Maximum processing weight 0

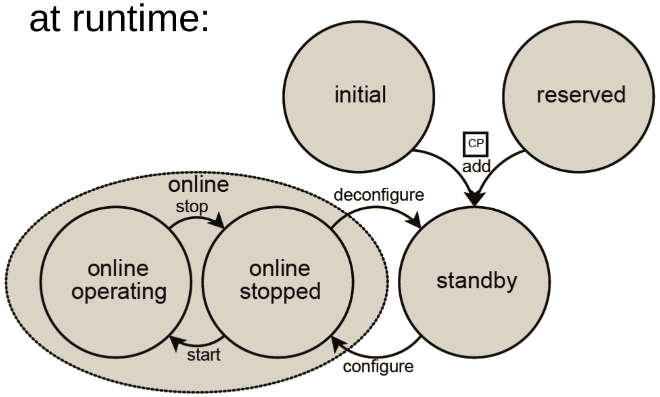
Logical Processor Add: H05LP56 (Active) - H05:H05LP56

CP Type	Number of Initial CPs	Number of Reserved CPs	Capping	Dedicated	Initial Weight	Minimum Weight	Maximum Weight
GP	1	2	☐	☐	10		
zAAP	0	0	☐	☐			
zIIP	0	0	☐	☐			

LPAR CPU state transitions during LPAR activation and IPL:



LPAR CPU state transitions at runtime:



z/VSE V4.3 – Dynamic Starting/Stopping of CPUs

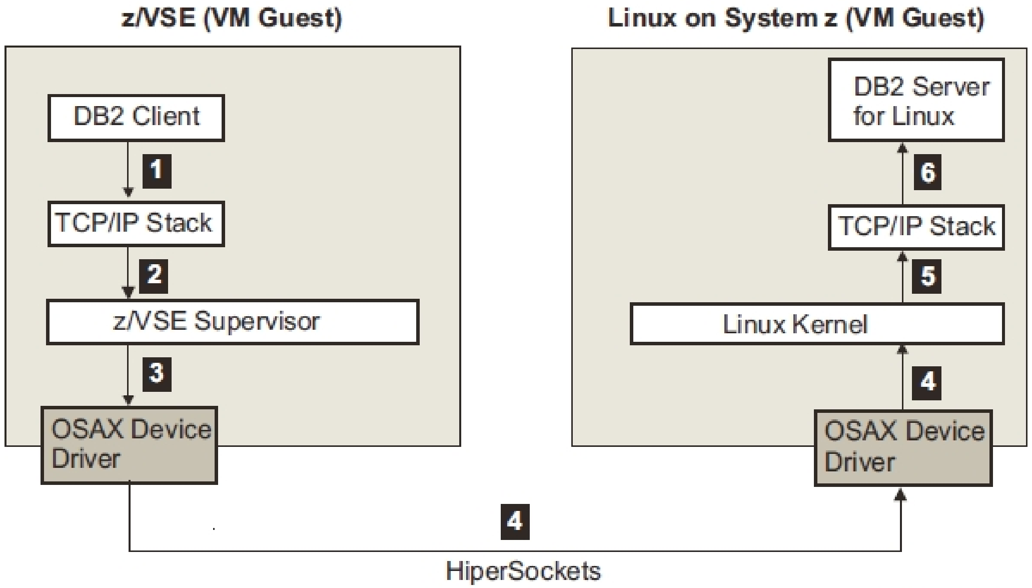
```
query td
AR 0015 CPU STATUS SPIN_TIME NP_TIME TOTAL_TIME NP/TOT
AR 0015 00  ACTIVE 0         43      23070      0.001
[…]

AR 0030 0W03I 00002 STANDBY CPUS HAVE BEEN ADDED TO THE CONFIGURATION

query td
AR 0015 CPU STATUS SPIN_TIME NP_TIME TOTAL_TIME NP/TOT
AR 0015 00  ACTIVE 0         65      23095      0.002
AR 0015 01  STANDBY
AR 0015 02  STANDBY
[…]

sysdef td,startsby=01
AR 0015 1YH7I NUMBER OF CPU(S) - ACTIVE: 1 - QUIESCED: 0 - INACTIVE: 0 -
              STANDBY: 1
AR 0015 1I40I READY
```

z/VSE V4.3 – Fast Path to Linux on System z (LFP)

- **The Linux Fast Path uses an IUCV connection between z/VSE and Linux, where both systems run in the same z/VM-mode LPAR on IBM z10 or z196 servers**
- **It allows selected TCP/IP applications to communicate with the TCP/IP stack on Linux without using a TCP/IP stack on z/VSE**
- **All socket requests are transparently forwarded to a Linux on System z system running in the same z/VM**
- **On Linux on System z, the LFP daemon must run**
  - This daemon fulfills all socket requests by forwarding them to the Linux TCP/IP stack.
- **The fast path to Linux on System z provides standard TCP/IP socket APIs for programs running on z/VSE:**
  - LE/C socket API via an alternative $EDCTCPV.PHASE (IJBLFPLE)
  - EZA SOCKET and EZASMI interface via an alternative EZA interface phase IJBLFPEZ
  - CSI’s (Connectivity Systems, Incorporated) assembler socket interface via the SOCKET macro
  - Other than the basic socket API, no other tools are provided.

- **Possible performance increase due to:**
  - Less overhead for TCP/IP processing on z/VSE (TCP, sequence numbers and acknowledging, checksums, resends, etc)
  - More reliable communication method (IUCV) compared to HiperSockets, which is a network device, with all its packet drops, resends, etc.

z/VSE V4.3 – Fast Path to Linux on System z (LFP)

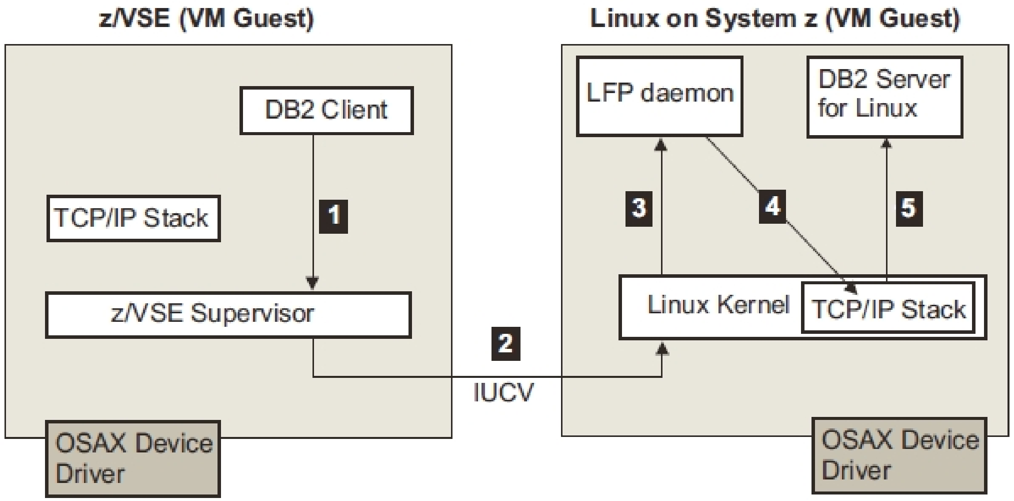


Communication using TCP/IP

- Data is passes from the application to the TCP/IP stack partition (involves dispatching)
- TCP/IP builds IP packets (including TCP, checksums, sequence numbers, etc) and sends it through the OSAX device driver
- Linux TCP/IP stack receives the packets and passes it to the application
- Application receives the data
- Response is going the reverse way

Communication using Linux Fast Path

- Data is passes from the application to the LFP
- LFP builds IUCV packets and sends it through to IUCV channel
- Linux IUCV device driver receives the packets and passes it to the LFP daemon LFPD.
- LFPD translates it into the appropriate socket calls against the TCP/IP stack
- TCP/IP passes the data to the application
- Response is going the reverse way



Performance measurements using Linux Fast Path

Comparison TCP/IP for VSE versus Linux Fast Path:

Workload	TCP/IP for VSE	Linux Fast Path (LFP)	Difference
FTP (BSI FTP server)			
§VSE à Linux (1GB) (NULL file, no I/O)	19 MB/sec 29% CPU (5% App + 24% TCPIP)	72 MB/sec 20% CPU (App)	3.7 times faster 9% less CPU
§Linux à VSE (1GB) (NULL file, no I/O)	21 MB/sec 55% CPU (11% App + 44% TCPIP)	70 MB/sec 20% CPU (App)	3.3 times faster 35% less CPU
Socket Application (running 3 times)			
§VSE à Linux (100MB)	4.6 MB/sec (*3 = 13.8 MB/sec)	14.6 MB/sec (*3 = 43.8 MB/sec)	3.2 times faster
§Linux à VSE (100MB)	9.7 MB/sec (*3 = 29.1 MB/sec)	16.2 MB/sec (*3 = 48.6 MB/sec)	1,7 times faster
	26% CPU (3*1% App + 23% TCP/IP)	9 % CPU (3*3% App)	17% less CPU

Environment: IBM System z10 EC (2097-722). TCP/IP connection via shared OSA adapter.

à Significant benefits in transfer rate as well as CPU usage

z/VSE V4.3 – Fast Path to Linux on System z (LFP)

- **Most existing applications run unchanged with Linux Fast Path**
  - Provided they use one of the supported Socket API (LE/C, EZA or ASM SOCKET)
    - And they do not use any CSI specific interface, features or functions
- **IBM Applications supporting Linux Fast Path**
  - VSE Connector Server
  - CICS Web Support
  - VSE Web Services (SOAP) support (client and server)
  - CICS Listener
  - DB2/VSE Server and Client
  - WebSphere MQ Server and Client
  - VSAM Redirector
  - VSE VTAPE
  - VSE LDAP Support
  - VSE Script Client
  - POWER PNET
  - TCP/IP-TOOLS included in IPv6/VSE product (e.g. FTP Server/Client)
- **Customer applications should run unchanged:**
  - Provided they use one of the supported Socket API (LE/C, EZA or ASM SOCKET)

z/VSE V4.3 – Queue-I/O Assist (QIOASSIST)

- **The z/VM function queue-I/O assist (QIOASSIST) provides performance improvements for V=V guests using real adapters and real networking devices that use the Queued Direct I/O (QDIO) facility**

- **z/VSE exploits the queue-I/O assist function for:**
  - OSA Express devices (CHPID type OSD)
  - HiperSockets devices (CHPID type IQD)

- **To use the queue-I/O assist (QIOASSIST) function in z/VSE, use the z/VM CP command:**
  - SET QIOASSIST ON

- **After you have enabled the queue-I/O assist function, each z/VSE DEFINE LINK,TYPE=OSAX command will then automatically exploit the queue-I/O assist function**

- **If you do not wish to use the queue-I/O assist (QIOASSIST) function in z/VSE, disable it using this z/VM CP command:**
  - SET QIOASSIST OFF

z/VSE V4.3 – Crypto Express3 and AP queue interrupt support

- **Support for AP-interrupts is a new function of IBM System z10 and IBM zEnterprise 196**

- **A hardware interrupt is issued when a response is ready for de-queueing from a card.**
  - Removes the need for the formerly used polling mechanism
  - User can switch between polling and interrupts (default: polling)
  - Using interrupts increase throughput for certain workloads without increasing CPU load

- **Not available under z/VM!**

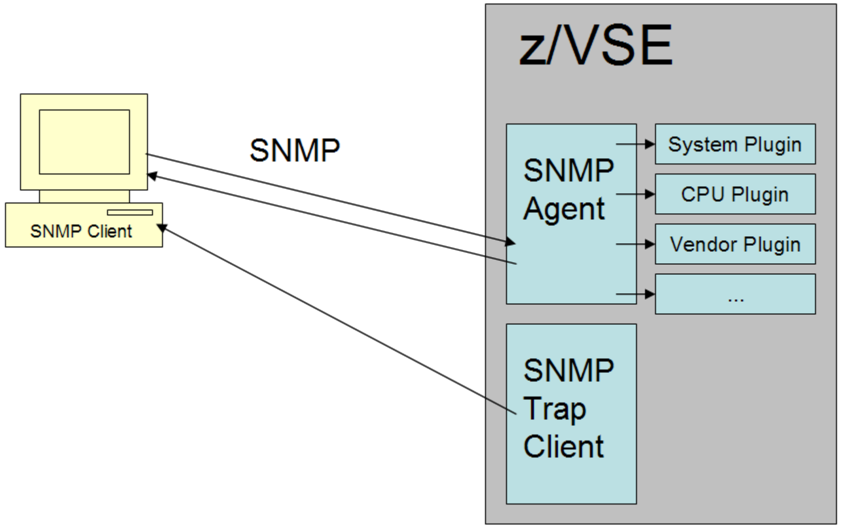
- **Supported cards are:**
  - Crypto Express2 and
  - Crypto Express3

- **The VSE crypto device driver provides new commands:**
  - APEAI, enable AP interrupts for all APs
  - APDAI, disable AP interrupts for all APs

z/VSE V4.3 – SNMP Monitoring Agent support

- **z/VSE Monitoring Agent enables customers to monitor z/VSE systems using standard monitoring interfaces (SNMP V1)**
  - It also includes an open interface, which enables customers or vendors to use own programs (plugins) to collect additional data

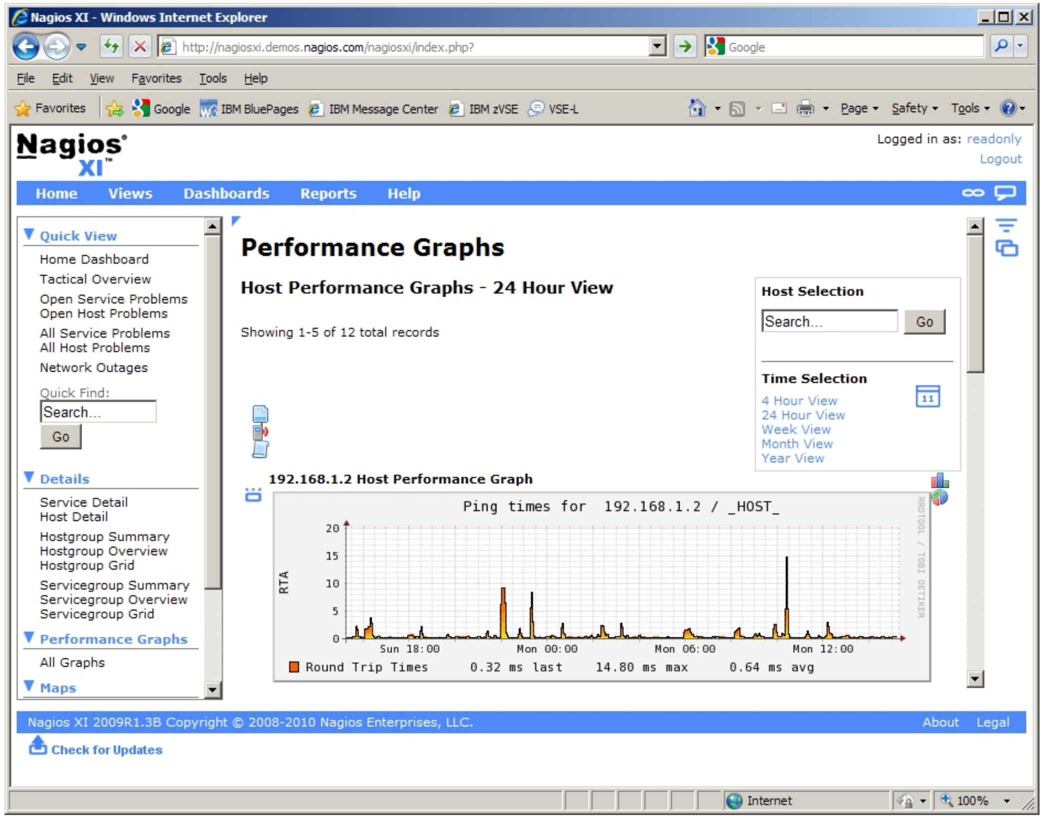
- **Data collected by the IBM provided plugins contains**
  - Information about the environment (e.g. Processor, LPAR and z/VM information)
  - Number of partitions (static, dynamic, total, maximum)
  - Partition priorities
  - Number of CPUs (active, stopped, quiced)
  - Paging (page ins, page outs)
  - Performance counters overall and per CPU
  - CPU address and status
  - CPU time, NP time, spin time, allbound time
  - Number of SVCs and dispatcher cycles



z/VSE V4.3 – SNMP Monitoring Agent support

- **A MIB (Measurement Information Base) is provided describing the data collected**
  - IESMPMIB.Z in PRD1.BASE (plain text member)

- **Standard SNMP based monitoring tools can be used to collect, display and analyze z/VSE performance monitoring data**
  - e.g. Nagios (www.nagios.org)



Performance Graphs

Host Performance Graphs - 24 Hour View

Showing 1-5 of 12 total records

192.168.1.2 Host Performance Graph

Ping times for 192.168.1.2 / _HOST_

■ Round Trip Times 0.32 ms last 14.80 ms max 0.64 ms avg

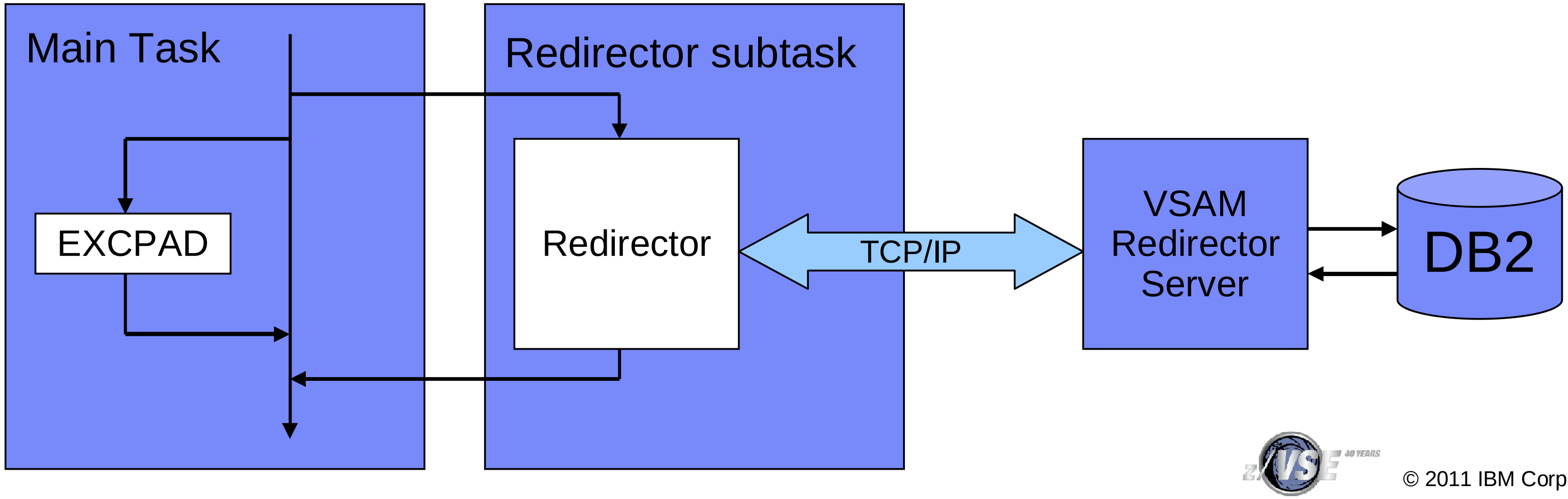
Nagios XI 2009R1.3B Copyright © 2008-2010 Nagios Enterprises, LLC. About Legal

Check for Updates

- **z/VSE SNMP Trap client**
  - Sends SNMP V1 traps to inform one or more monitoring stations or servers about important events
  - For example:
    - The end of a job stream is reached.
    - An error has occurred during a job stream

z/VSE V4.3 – Redirector EXCPAD - Overview

- **Prior to z/VSE 4.3 the VSAM Redirector was executed in the same subtask as VSAM and the application (caller)**
  - Redirector activities may be time consuming (network transfers, database operations, ...)
    - During this time, no other activities are possible for this subtask
  - Under CICS, VSAM normally returns back via EXCPAD exit when waiting for an I/O
    - Allows CICS to perform other activities concurrently
- **Since z/VSE 4.3 VSAM executes the Redirector under a separate subtask**
  - VSAM now also returns back to CICS via EXCPAD when waiting for Redirector
    - Allows CICS to perform other activities concurrently
  - This capability is primarily implemented for CICS TS transactions.
    - The Redirector EXCPAD is not used for VSAM files opened by CICS/VSE.



z/VSE V4.3 – Redirector EXCPAD - Benefits

- **Benefits:**
  - **Prior to z/VSE 4.3 heavy use of VSAM Redirector could slow down transaction processing in CICS**
    - Due to VSAM requests block the CICS I/O task when Redirector is active
  - **With the new subtask the VSAM Redirector handling no longer blocks the CICS I/O task**
    - Allowing other transactions to do its work
    - Multiple redirected requests will be queued up for processing in the new subtask

- **The EXCPAD user exit is enabled automatically under the following conditions:**
  - a VSE/VSAM cluster is enabled for the Redirector
  - the EXCPAD exit is defined during the OPEN request

- **VSAM will attach only one Redirector subtask per partition even if multiple redirected files are opened in the partition with an active EXCPAD**

- **Support is transparent**
  - No need to configure or setup anything
  - All types of Redirector activities are processed in subtask (except OPEN/CLOSE)
    - VSAM Redirector OWNER=VSAM or REDIRECTOR
    - VSAM Capture Exit
    - Customer/Vendor implemented Redirector Exit

Agenda

- **z/VSE V4.3 Performance Considerations**
  - Release Overhead Deltas
  - Support for 4 digit device addresses
  - GETVIS constraint relief
  - Dynamic start/stop of CPUs
  - 1 MB frames for data spaces
  - Fast Path to Linux on System z
  - Queue-I/O Assist (QIOASSIST)
  - Crypto Express3 and AP queue interrupt support
  - FICON Express8
  - SNMP Agent support
  - VSAM Redirector executed in subtask under CICS
- **z/VM and Linux considerations** ←
- **Sizing a System for z/VSE**
- **Performance Measurement Tools**

Shared OSA Adapter versus HiperSockets

To connect a z/VSE system with a Linux on System z you have 2 options:

1. Using a shared OSA Adapter

- All traffic is passed through the OSA Adapter
- The OSA Adapter has its own processor
  - Processing occurs asynchronous
  - Processing in OSA Adapter does not affect host processors

2. Using HiperSockets

- Direct memory copy from one LPAR/Guest to the other
- Memory copy is handled by the host processors
  - Processing occur synchronous
  - Consider mixed speed processors (full speed IFLs and throttled CPs)
    - à Memory copy performed by throttled CP is slower than memory copy performed by full speed IFL

Performance tuning for HiperSockets

§ **When using HiperSockets to communicate between z/VSE and Linux, you may run into a “Target Buffer Full” condition**

- This happens when z/VSE sends faster/more than Linux can receive
- Per default Linux has 16 inbound buffers (64K per buffer = 1M per link)
- To increase the number of buffers on Linux, use QETH option “buffer_count=128”
  - Use YAST to configure, or sysconfig scripts
  - Maximum of 128 buffers require 8MB of storage per link

§ **When TCP/IP for VSE encounters this situation (BUSY), it waits 500 msec until it retries to send the packet**

- Any additional packets to be sent are queued up
- Problem can become dramatic, if more than 16 packets are queued up to be sent after BUSY situation
  - The resend will immediately flood the Linux buffers again, leading to the next BUSY situation, and so on....

§ **You can check via QUERY STATS,LINKID=xxxx [,RESET] if you have ever run into the BUSY situation** (RESET resets the counters)

```
C1 0065 0004: IPL615I  Busy mode.............................0   ß see here
C1 0065 0004: IPL615I  Busy mode, longest....................0
```

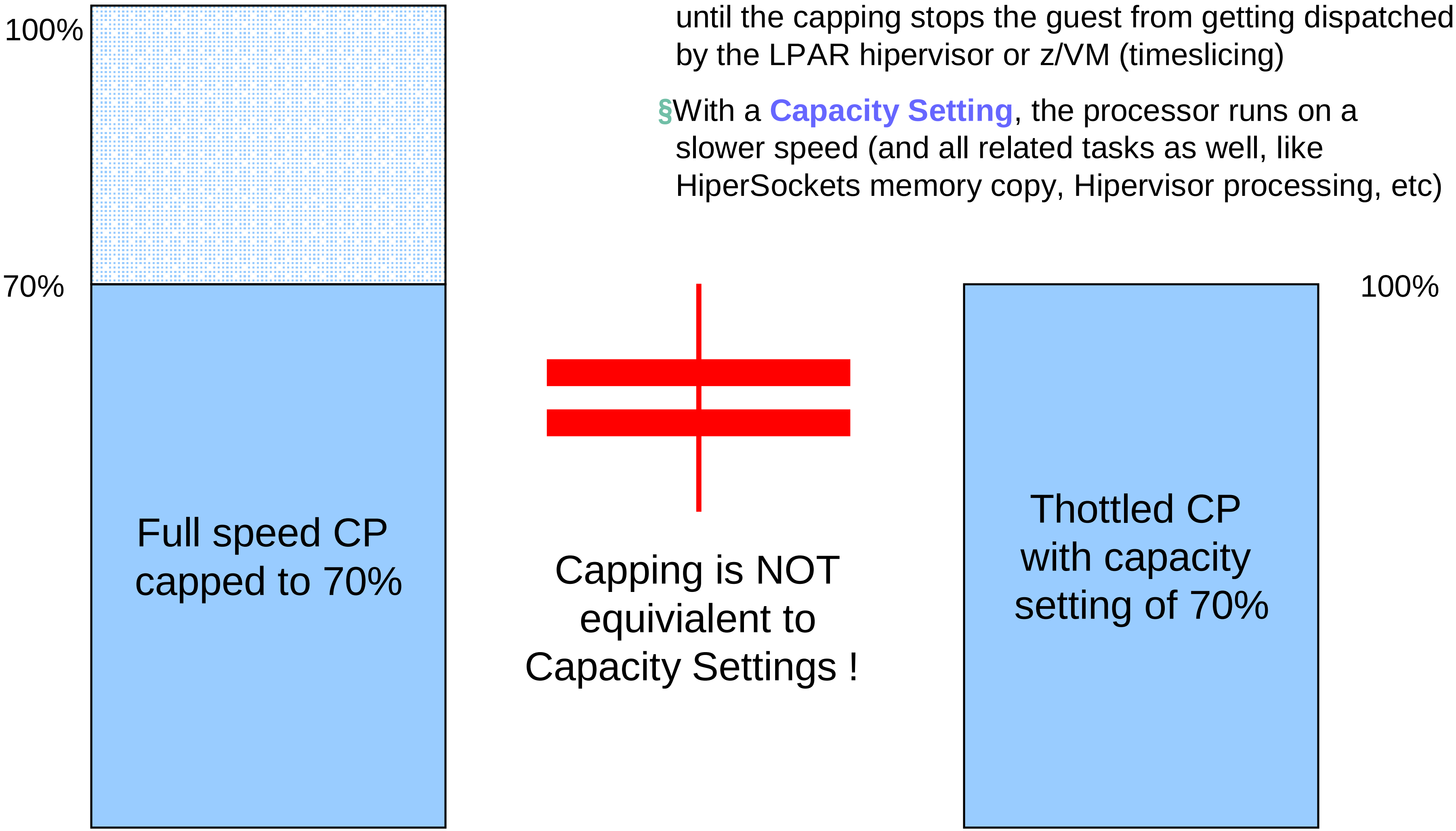
§ **You can configure a shorter BUSY wait time via DEFINE LINK command**

- BUSY=nnn (shortest possible wait time is 100 msec)

Capping versus Capacity Settings

Attention: Do not use Capping to simulate Capacity Settings !

- With **Capping**, the processor runs on its full speed, until the capping stops the guest from getting dispatched by the LPAR hipervisor or z/VM (timeslicing)
- With a **Capacity Setting**, the processor runs on a slower speed (and all related tasks as well, like HiperSockets memory copy, Hipervisor processing, etc)



100%

70%

Full speed CP capped to 70%

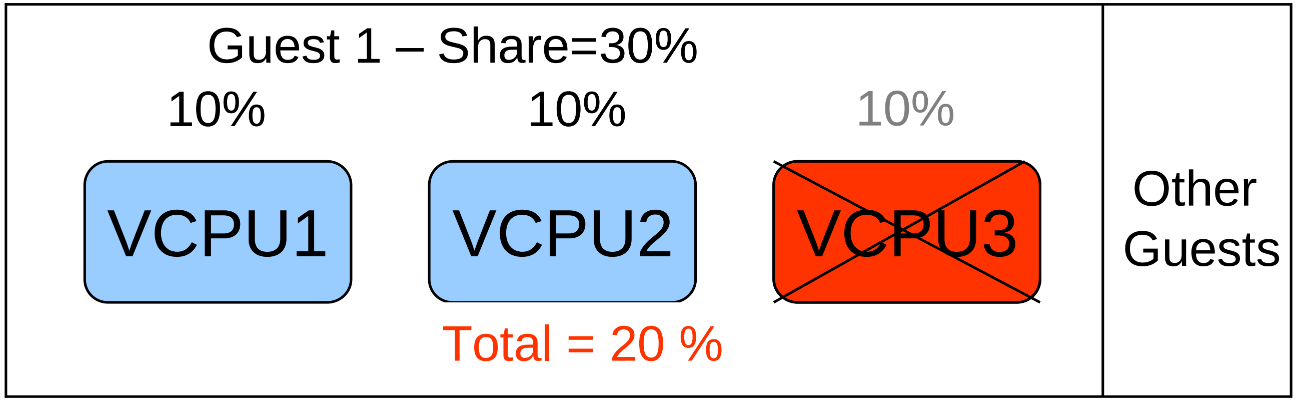
Capping is NOT equivilalent to Capacity Settings !

100%

Thottled CP with capacity setting of 70%

z/VM 5.4 Considerations

<= z/VM 5.3



Guest 1 – Share=30%

10% 10% 10%

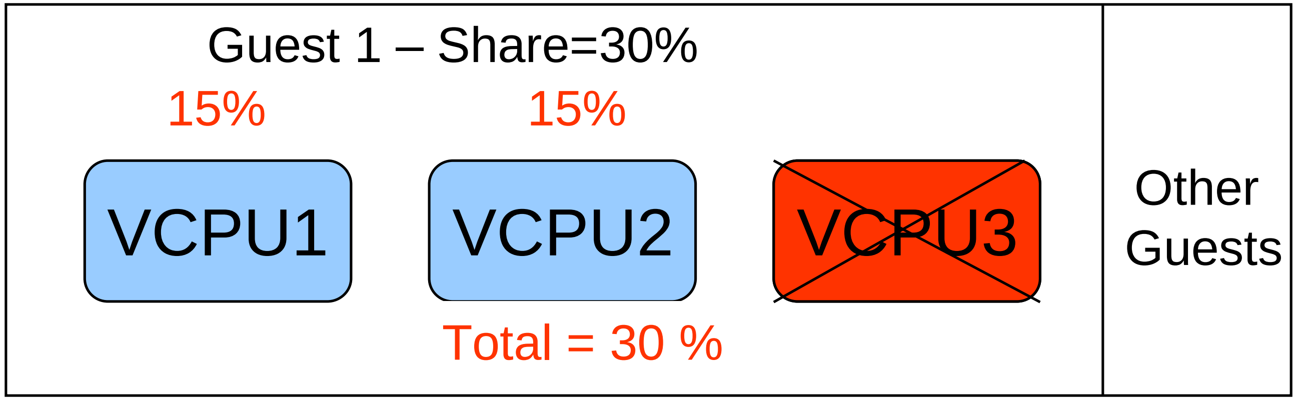
VCPU1 VCPU2 ~~VCPU3~~

Total = 20 %

Other Guests

A guest's CPU share is distributed equally among its virtual processors by dividing its share value by the number of processors, regardless of whether the virtual processors were in a stopped or started state.

z/VM 5.4 or later



Guest 1 – Share=30%

15% 15%

VCPU1 VCPU2 ~~VCPU3~~

Total = 30 %

Other Guests

z/VM V5.4 performs share redistribution whenever a virtual processor is started or stopped and no longer includes stopped virtual processors in the calculation of how much share to distribute to each virtual processor.

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 instead with the z/VSE workloads Batch, Online or Mixed
  - Use free of charge Capacity Planning Services from IBM

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 z processors running various workload environments

- Capacity results are based on IBM's LSPR data supporting all IBM System z processors
  - Large System Performance Reference: http://www.ibm.com/systems/z/advantages/management/lspr/

- For VSE use z/VSE workloads Batch, Online or Mixed

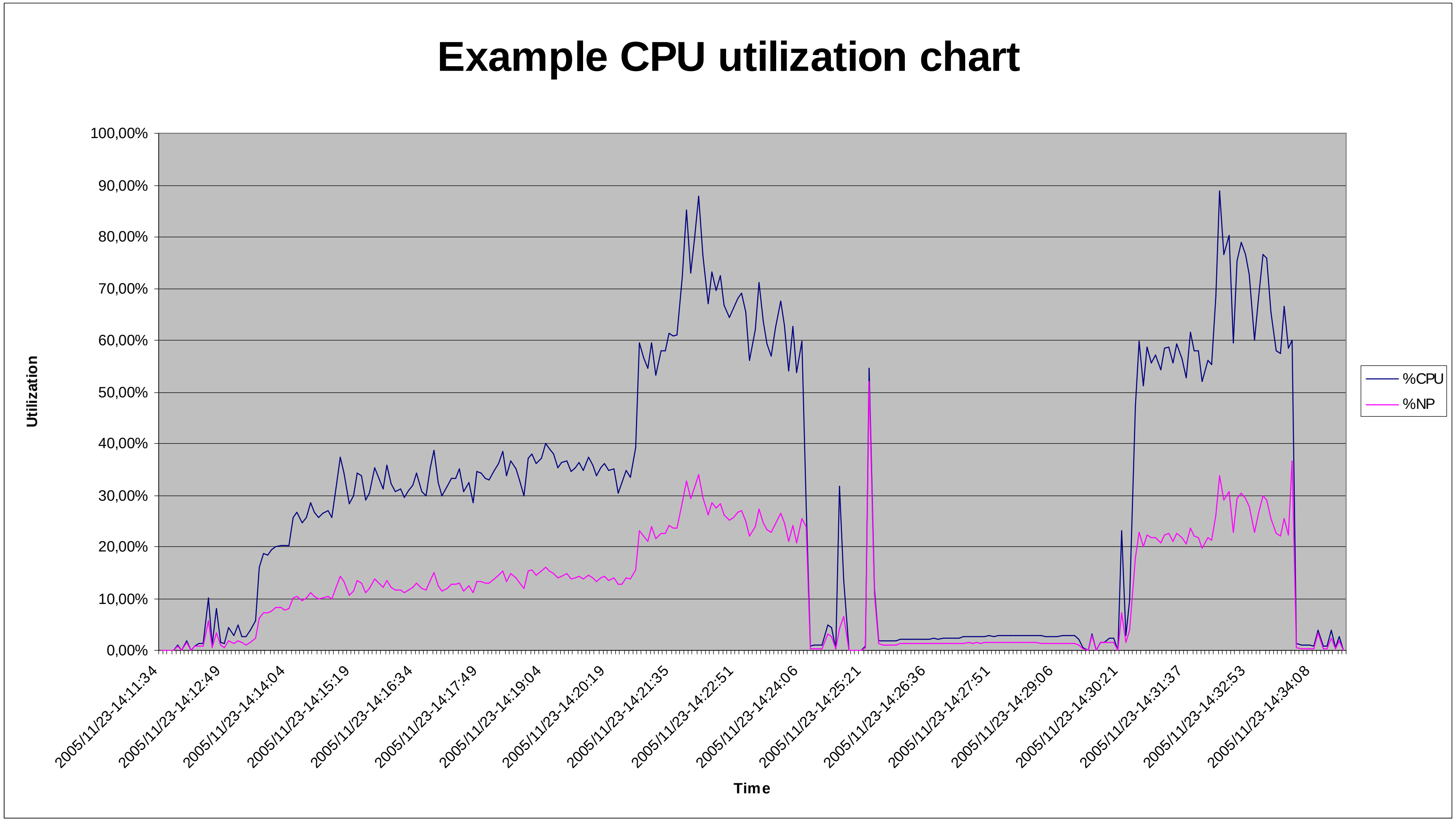
z/VSE CPU Monitor Tool

- 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
- New version available (XML Output) for z/VSE Capacity Planning

- Download
  - http://www.ibm.com/systems/z/os/zvse/downloads/tools.html
  - ‘As is’, no official support, e-mail to zvse@de.ibm.com

z/VSE CPU Monitor Tool



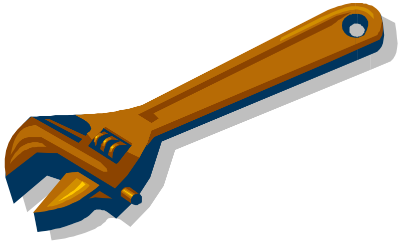
z/VSE Capacity Planning Offering

- **A new z/VSE Capacity Planning Offering is now available**
  - Available for Business Partners
  - and Customers
- **Performance data collection is based on a new version of the CPUMON Tool**
- **Analysis is done using zCP3000**

- **Contact techline@us.ibm.com** and ask for z/VSE Capacity Planning Support

z/VSE monitoring tools

- System Activity Dialogs (SYS fast path 361 and 362)
  - Displays real-time performance information about the System, CPU, partitions and I/O
- QUERY TD command
  - Displays information about CPU usage on the console
- SIR SMF command
  - Displays I/O related performance information on the console
- Job Accounting Exit (SKJOBACC in ICCF library 59)
  - Prints performance related information (CPU, I/O) to SYSLST after each job step
- MAP and GETVIS commands
  - Displays memory related information on the console
- z/VSE CPUMON Tool
  - Monitors overall system CPU usage and performance counters
- CICS Statistics
  - Prints CICS statistics
- CICS built-in tools like CEMT INQUIRE
  - Displays information about CICS ressources
- A z/VSE performance monitor product for batch and CICS
  - Like CA Explore, ASG TMON, etc.
- z/VSE V4.3 SNMP Agent



Documentation

- z/VSE homepage:
  - http://www.ibm.com/systems/z/os/zvse/
- z/VSE Performance:
  - http://www.ibm.com/systems/z/os/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 z Software Pricing
  - http://www-03.ibm.com/systems/z/resources/swprice/
- IBM's MSU ratings for IBM System z
  - http://www.ibm.com/systems/z/resources/swprice/reference/exhibits/hardware.html

Questions ?

?