VM for VSE Guest Performance

VM & VSE Tech Conference

May 2000 - Orlando

Session G41

Bill Bitner

VM Performance

IBM Endicott

607-752-6022

bitnerb@us.ibm.com

RETURN TO INDEX

Legal Stuff

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 environment 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 references to, or information about, IBM products (machines and programs), programming, or services that are not announced in your country or not yet announced by IBM. Such references or information should not be construed to mean that IBM intends to announce such IBM products, programming, or services.

Should the speaker start getting too silly, IBM will deny any knowledge of his association with the corporation.

Trademarks

The following are trademarks of the IBM Corporation:

IBM, VM/ESA, VSE/ESA, CICS

Overview

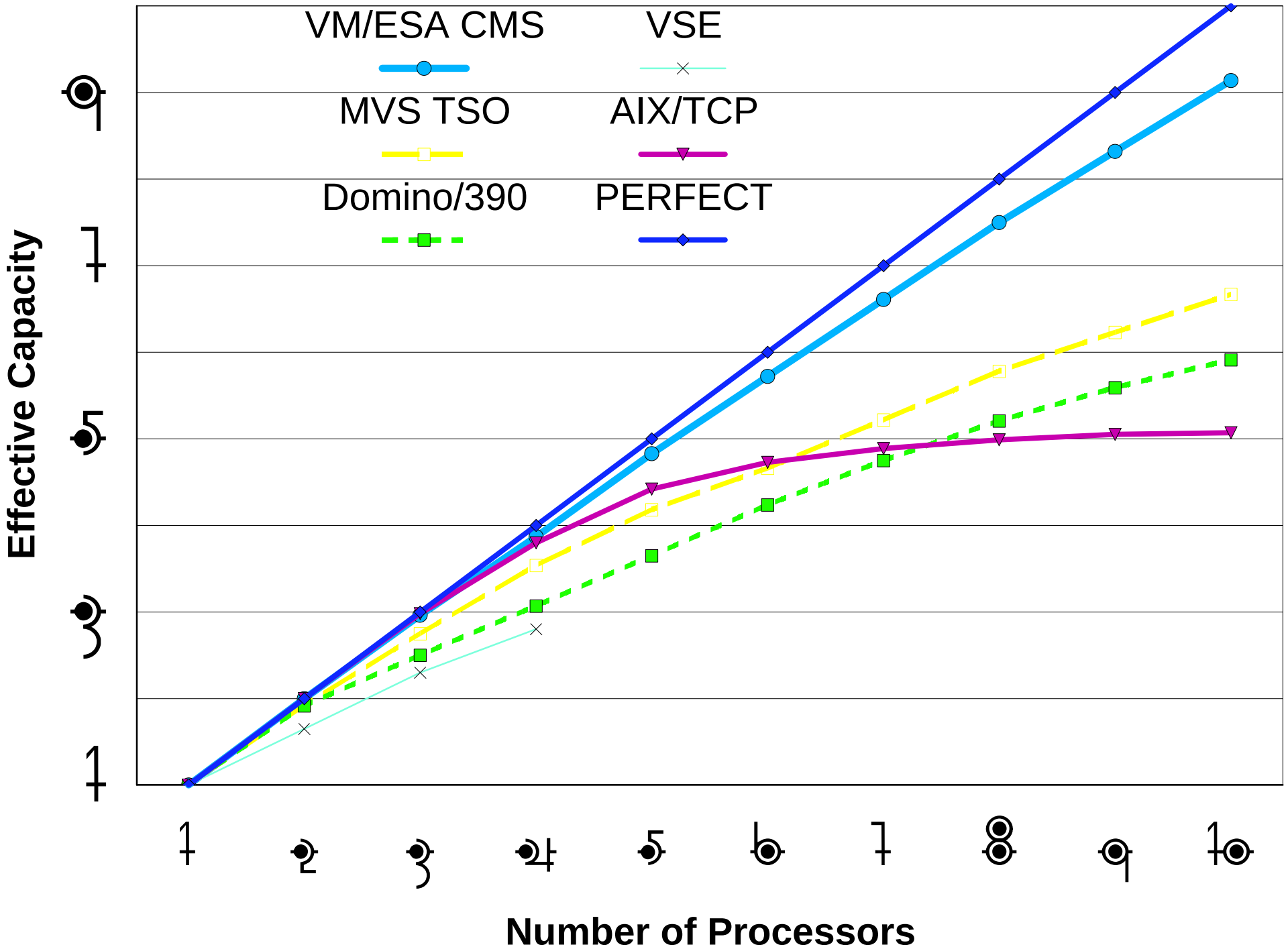
- Is VM Good for VSE Performance?
  - It depends.
- What does it depend on?
  - On what you mean by "performance"
  - Using V=R/F
  - Running VM on Native or LPAR
  - Virtual disk in storage
  - Enhanced minidisk cache
  - Hardware

The Short Answer is Yes

- Performance Value-Add of VM
  - Extend capacity of single VSE, by running multiple VSEs
  - VM/ESA extensions to scheduling such as limit shares
  - Resource sharing
    - Real storage is shared for V=V guests
    - Channels are shared without EMIF
    - DASD devices can be split up into minidisks

The Short Answer is Yes...

- VM paging benefits - demand and block paging, use of expanded storage.
- HW exploitation - greater N-way, expanded storage
- VM/ESA features
  - Virtual disk in storage for lockfile
  - Enhanced minidisk cache
  - DB2 for VM Guest Sharing


VM/ESA CMS
VSE
MVS TSO
AIX/TCP
Domino/390
PERFECT
Effective Capacity
Number of Processors

Nonperformance Reasons

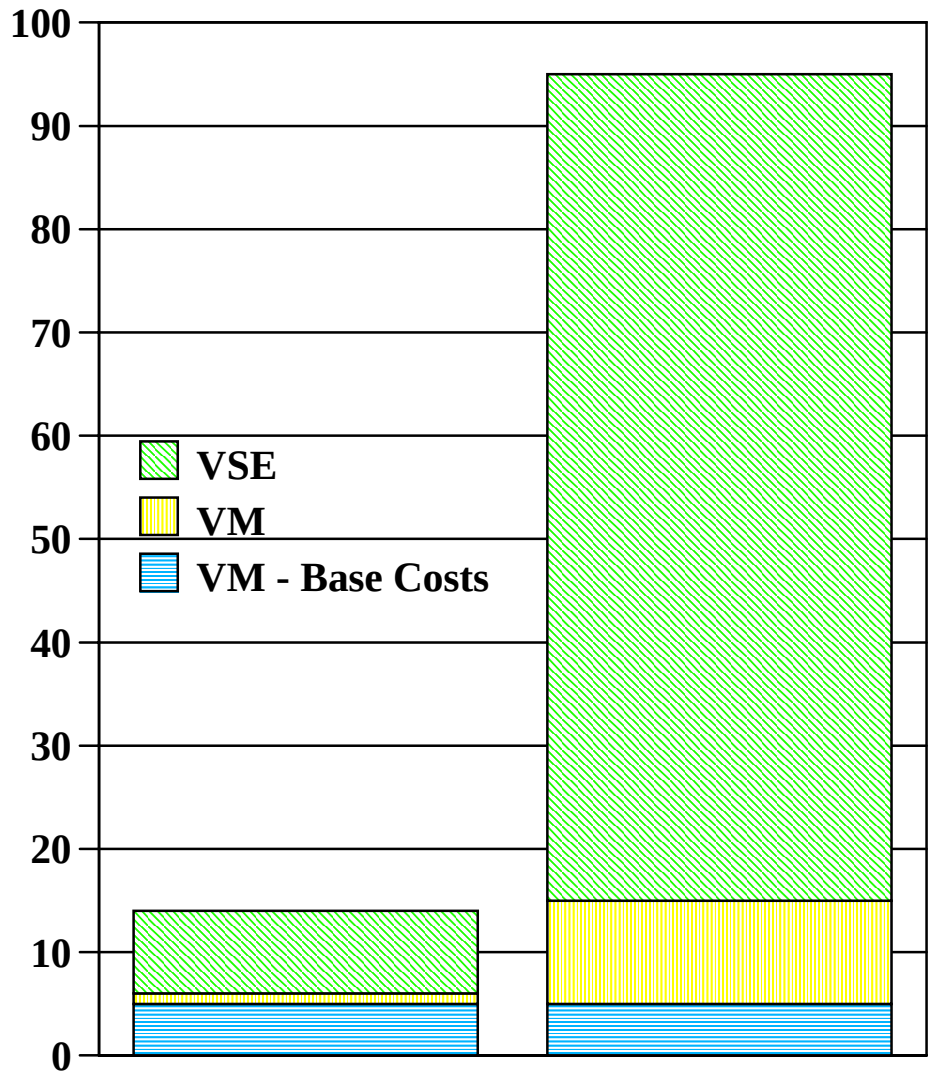
- CMS features
- Isolate production and test
- Migration vehicle
- Resource Management
- Other PPs (e.g. OV/VM)

What do you mean by "Performance"?

- Critical to answering the original question.
- Typically one of the following:
  - ITR = Internal Throughput Rate = a measure of work per CPU second.
  - ETR = External Throughput Rate = a measure of work per wallclock second.
  - CPU Utilization = how busy processor is; tied to ITR.
  - Response Time (Elapsed Time) = how long jobs take; tied to ETR
  - Interactive Users vs. Batch Work
  - How many phone calls you get

CPU Usage by VM/ESA

- Base costs and background work
  - Scheduling and dispatching
  - Accounting
  - Monitor
- Costs proportional to VSE requests or requirments of VM/ESA



CPU Usage - SIE

- Used by VM/ESA to run VSE guest
- Exits from SIE indicate work for VM
- Hardware assists can help avoid SIE exits
- Most common reasons for exiting SIE
  - I/O processing
  - Page fault resolution
  - Instruction simulation
  - Minor time slice expires

Exits from SIE

- Data in memory techniques avoid I/O.
- I/O Assist avoids SIE exit to handle:
  - I/O interrupt processing
  - CCW translation from virtual to real addresses
- CCW translation bypass for V=R guest.
- Minor time slice: SET SRM DSPSLICE
- Avoid Paging
  - V=R/F
  - Reserved pages for V=V
  - Sufficient storage

VM I/O Processing

- I/O Assist
  - V=R/F Guests
  - Dedication Devices
- CCW Translation Bypass
  - Dedicated and full pack minidisks
  - Only some V=R I/O
  - SET NOTRANS ON (SET CCWTRAN OFF)
- Fast CCW translation
  - Applies only to select DASD I/O

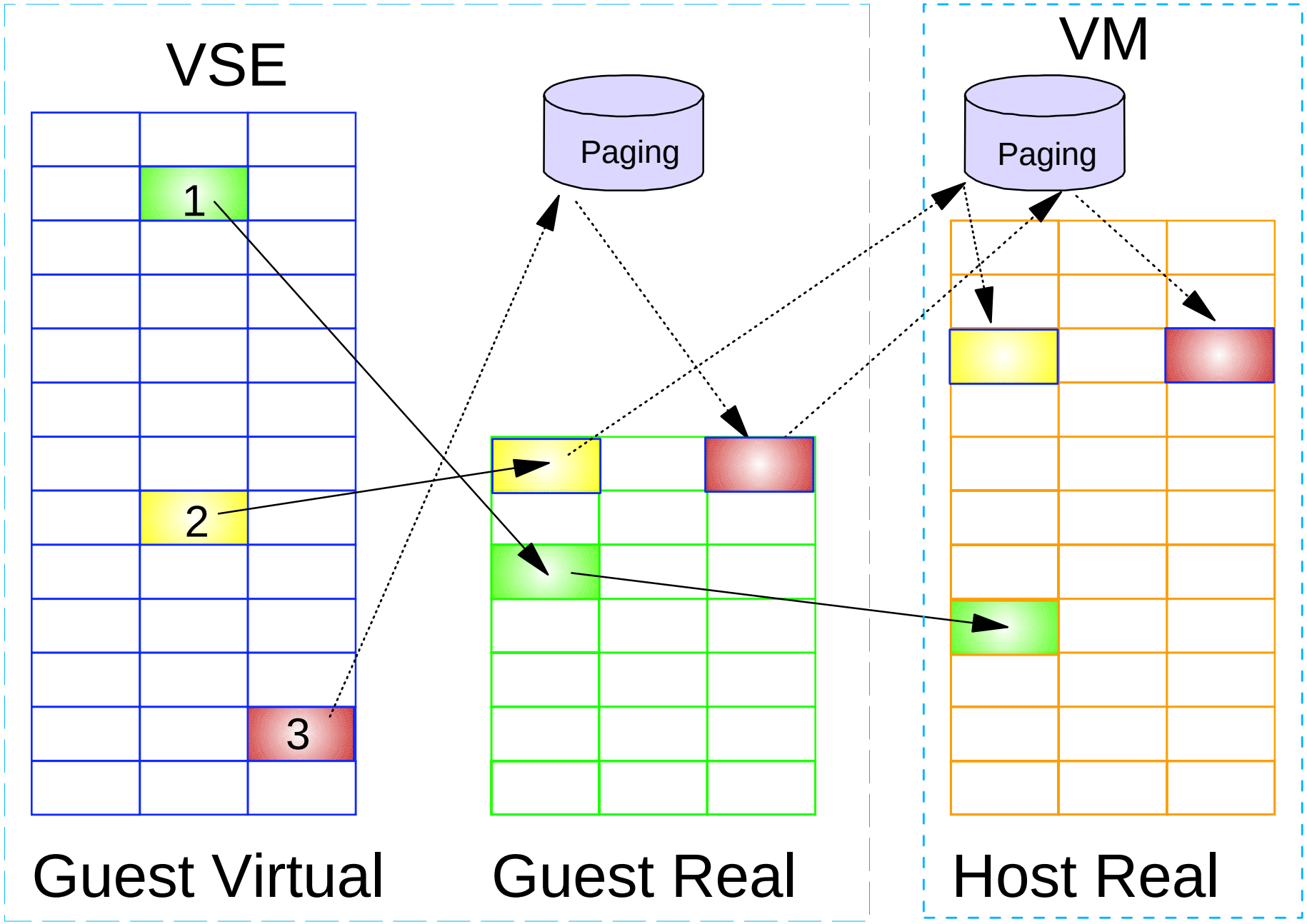
I/O Considerations

- I/O Assist gives best CPU performance
- Dedicated I/O is not eligible for MDC
- For V=R CCWTRANS OFF makes guest I/O ineligible for fast CCW translation
- For VSE Guests, VSE vdisks are more efficient than VM vdisks.
- Both VM vdisks and MDC require sufficient storage
- MDC read performance is as good as VM vdisk performance

DASD Considerations

- Dedicated Devices
  - Required for full I/O Assist
  - Not eligible for MDC
- Full pack minidisks
  - Can be shared between guests
  - Some I/O assist
  - Define via VOLSER or DEVNO
  - DEVNO not eligible for MDC
- Partial pack minidisks can be shared
- FBA Volumes - should start/end on 64 512-byte block boundaries.

Paging Considerations



Paging Considerations

- For V=V guests the potential exists for "Double Paging"
- No VM paging for V=R/F
- The closer the VSE VSIZE is to the defined storage for the virtual machines, the lower the VSE paging.
- PAGEX ON use.
- VM/ESA can use expanded storage for high speed paging device.

V=R/F/V Considerations

- V=R/F potential I/O assist benefit (saves CPU)
- V=F avoids overhead of recovering V=R
- 1 V=R + 5 V=F or 6 V=F
- V=V avoids dedicating storage
- V=R defaults to dedicating processors
- Running VM/ESA in an LPAR -
  - No V=F, only V=R, but without I/O Assist
  - Often better to use V=V and reserve pages

Virtual MP Support

- Define additional processors dynamically
  - Directory include MACHINE ESA 2
  - CP DEFINE CPU vcpu_addr
- Or put everything in the directory
  - CPU 00 NODEDICATE
  - CPU 01 NODEDICATE
- Detaching vCPU resets virtual machine
- For testing: more virtual than real processors

Virtual MP Support

- CP commands of interest
  - QUERY VIRTUAL CPUS
  - CPU vcpu_addr cmd_line
  - DEDICATE and UNDEDICATE

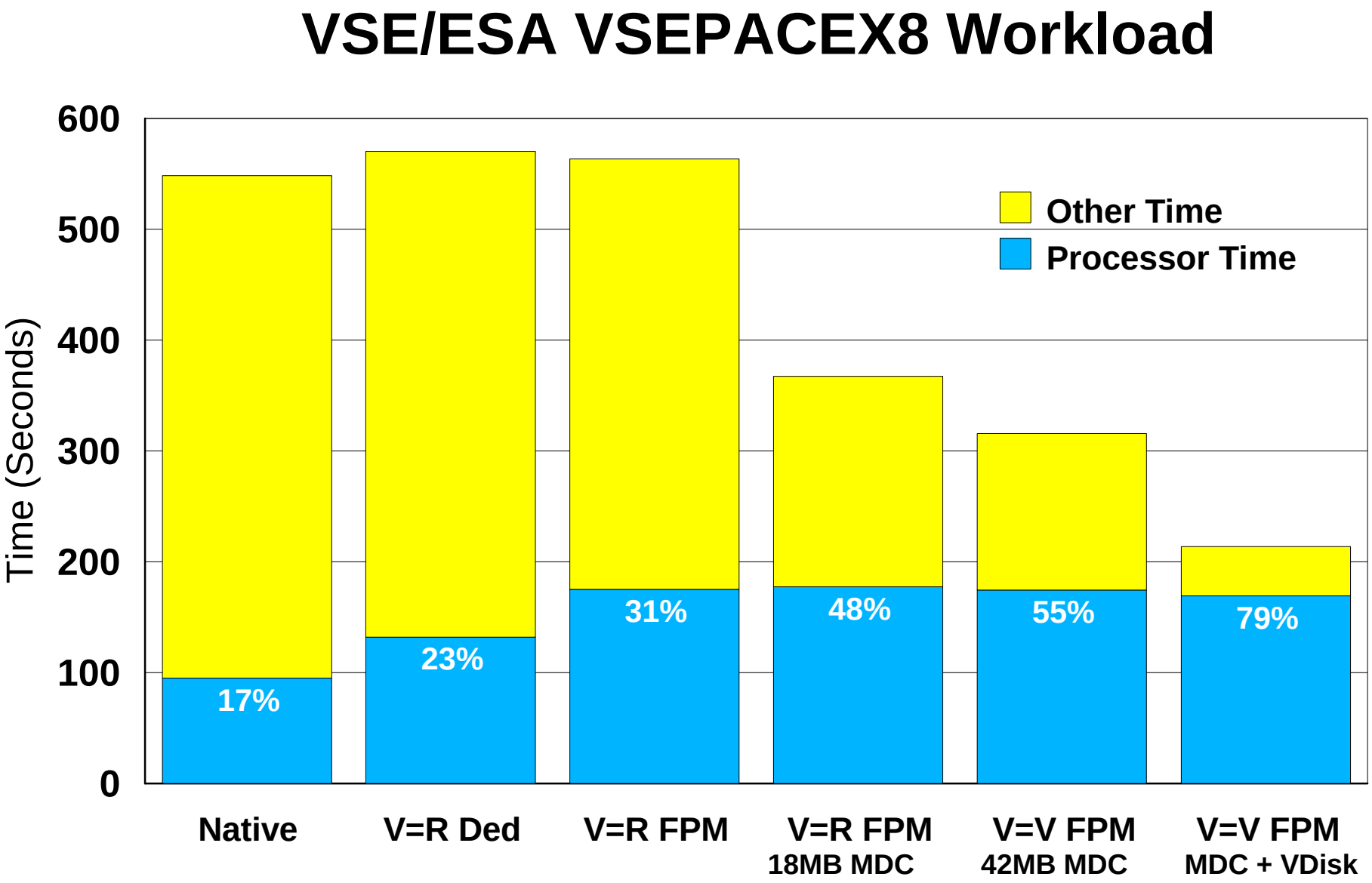
Virtual MP Tuning

- Share setting is for virtual machine, divided amongst all virtual processors
- Processors can be dedicated
- Mixing dedicated and shared processors is not recommended
- Defined but inactive vCPU (stopped state) makes guest ineligible for I/O assist
- Monitor, INDICATE, RTM for all vCPUs
- Potential for >100% (N*100%)
- Dedicated processor looks 100% busy

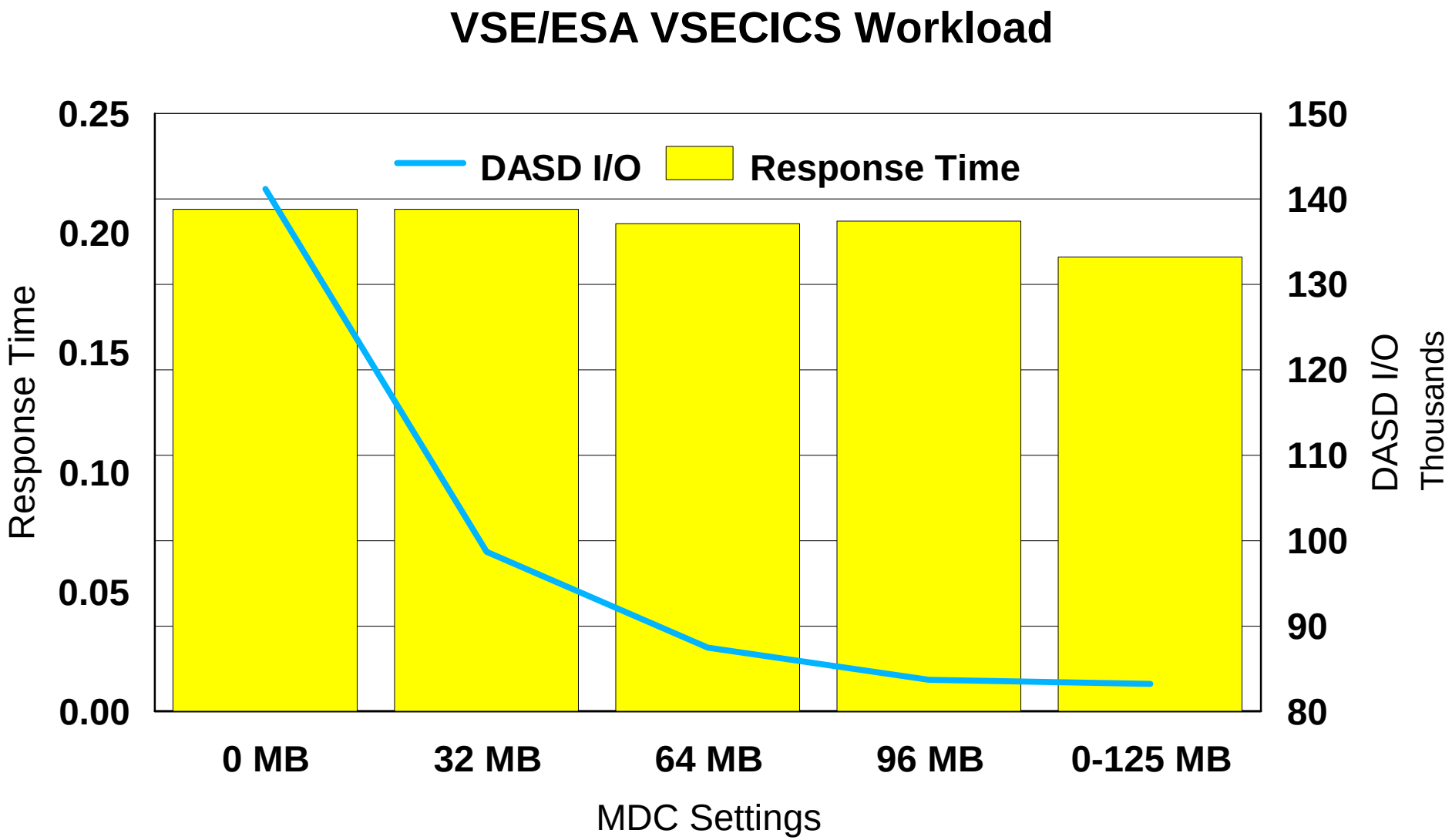
VM/ESA Data in Memory Techniques

- VM Data Spaces
  - Exploited by DB2 Server for VSE and VM (SQL/DS)
- VM Virtual disk in storage
  - volatile FBA minidisk
  - private or shareable
  - perfect for lock file
- Minidisk cache
  - Undedicated 3380, 3390, 9345, and RAMAC
  - SSCH, SIO, SIOF and Diagnose I/O
  - Read-once data generally does not benefit
  - New (2.3.0) Record level MDC does not apply to VSE
  - Do not use MDC for VSE lockfile

Minidisk Cache Benefits Example 1



Minidisk Cache Benefits Example 2



Summary

- Many features to be exploited
- Optimum configuration will depend on
  - What you mean by the term **performance**
  - What resources you have available
- http://www.ibm.com/s390/vm/perf/tips/
  - Common problems/solutions
  - CCW translation matrix
  - VSE Guest Performance
  - Performance related APARs
  - MDC guidelines
  - N-way and CMOS thoughts
- VM is still a friend and helper to VSE