IBM VSE/ESA 2.5 Performance Considerations

Wolfgang Kraemer

VSE Product Mgmnt
Dept 3221
71032-04 Boeblingen
WKRAEMER at DEVM
wkraemer at de.ibm.com

Status 2001-07-15

Contents

Contents

Contents

Notes

Notes

All information contained in this document has been collected and is presented based on the current status.

It is intended and required to update the performance information in this document.

It is the responsibility of any user of this VSE/ESA V2 document

- to use the latest update of this document
- to use this performance data appropriately

This document is unclassified and intended for VSE customers.

The VSE performance documents are e.g. available from the INTERNET via the VSE/ESA home page

http://www.ibm.com/servers/eserver/zseries/os/vse

(http://www.ibm.com/s390/vse/ former URL)

Starting with the VSE/ESA 2.4 documentation, these documents are also available on the VSE/ESA CD-ROM kit SK2T-0060.

The following documents are available in Adobe Reader format (.PDF):

'IBM VSE/ESA 1.3/1.4 Performance Considerations'
'IBM VSE/ESA V2 Performance Considerations'
'IBM VSE/ESA Turbo Dispatcher Performance'
'IBM VSE/ESA I/O Subsystem Performance Considerations'
'IBM VSE/ESA VM Guest Performance Considerations'
'IBM VSE/ESA Hints for Performance Activities'
'IBM VSE/ESA TCP/IP Performance Considerations'
'IBM DFSORT/VSE Performance Considerations'
'IBM VSE/ESA CICS Transaction Server Performance'
'IBM VSE/ESA 2.5 Performance Considerations' (this document)
'IBM VSE/ESA Performance on xSeries (NUMA-Q) Enabled for S/390'

The files are
VE13PERF.PDF, VE21PERF.PDF, VE21TDP.PDF, VEIOPERF.PDF, VEVMPERF.PDF, VEPERACT.PDF, VETCPPER.PDF, VESORTP.PDF, VECICSTS.PDF, VE25PERF.PDF, VEXEFSP.PDF

Notes ...

Disclaimer

This document has not been subjected to any formal review or testing procedures and has not been checked in all details for technical accuracy. Results must be individually evaluated for applicability to a particular installation.

Any performance data contained in this publication was obtained in a controlled environment based on the use of specific data and is presented only to illustrate techniques and procedures to assist to understand IBM products better.

The results which may be obtained in other operating environments may vary significantly. Users of this document should verify the applicability of this data in their specific environment.

The above disclaimer is required since not all dependencies can be described in this type of document.

Acknowledgements

Thanks to all who contributed directly or indirectly, be it by measurements, suggestions or in other ways.

Special thanks to Hanns-J. Uhl for the numerous performance measurements done.

All mistakes and inaccuracies in this document are my own.

Please, as in the past, contact me if you have

- suggestions or questions regarding this document
- questions on VSE/ESA performance, not covered in any of the VSE/ESA performance documents

Wolfgang Kraemer, IBM VSE Development, Boeblingen Lab, Germany

Notes ...

Trademarks

The following terms included in this paper are trademarks of IBM:

VSE/ESA	ESA/390	System/390	DB2	PR/SM
VM/ESA	VSE/ESA	ESCON	ECKD	n-ways
CICS Transaction Server		FICON	ESS	AIX
NetFinity				
WebSphere				

The following are trademarks of their respective owners:

UNIX	X/Open Company
LINUX	Linus Torvalds
Windows	Microsoft Corporation
TCP/IP for VSE	Connectivity Systems Inc.
TCP/IP-TOOLS	Barnard Systems Inc.
IpBridge,IpServer	Data21 Inc.
eXadas eData Engine	Cross Access Corporation
Web/VSE-Host	IntelliWare Systems Inc.
ViaSQL	Viaserv Inc.
OPTI-...	Barnard Systems Inc.
BIM-...	Ben Moyle Inc.
Java	Sun Microsystems Inc.
Adabas	Software AG
SnapShot	Storage Technology Corporation

VSE/ESA 2.5 Performance Relevant Items

PART A.

VSE/ESA 2.5 Performance Relevant Items

For articles related to VSE/ESA 2.5, you may refer to

'VSE/ESA Software Newsletter' 3rd/4rth Quarter 2000, G225-4508-21, 12/2000

VSE/ESA 2.5 Performance Item Categories

VSE/ESA 2.5 Performance Item Categories

- **All Performance PTFs shipped after 2.4.0 GA**
- **Installation Enhancements**
- **Hardware Support**
- **VSE Base Enhancements**
- **VSE/VSAM Enhancements**
- **VSE/POWER Enhancements**
- **CICS TS related Enhancements**
- **e-business related Enhancements:**

  'Interoperability'

No change in TD related performance values:
(MP-factors, Non-Parallel Shares)

VSE/ESA 2.5.1 Refresh available since 2001-03-16.
Note that DY45644 is (only) the AF level of 2.5.1.

VSE/ESA 2.5 Performance Item Summary

VSE/ESA 2.5 Performance Item Summary

Productivity items also included

- **Installation Enhancements**

More generous selection of Env. B installation defaults:

- VSIZE increased to 264M (150M in Env. A)
- some partition sizes
- new dynamic classes classes used: R,S

Up to 32 extents allowed for shipped libraries

	PRD1	PRD2
Base Install	32	32
FSU	32	16

CA-Top Secret for VSE 1.3.0 shipped with the base.
Key protected, needs APAR PQ40529 (PTF UQ45985),
and CA-CIS ('system adapter').

- **Hardware Support**

Support of FICON

Supported by VSE/ESA V2 (if appropriate by a PTF).
(Prefetch not supported)

Support of ESS FlashCopy

True point-in-time copy. Usage similar to RVA SnapShot.
H/W implementation is different for ESS, refer to separate charts.

VSE/ESA 2.5 support is available as part of VSE Central Functions.
IXFP/SnapShot is a priced feature of VSE Central Functions.

Refer to Part 'ESS'.

VSE/ESA 2.5 Performance Item Summary ...

VSE/ESA 2.5 Performance Items (cont'd)

- **VSE Base Enhancements**

More than 10 dynamic classes (per table)

Allows more flexible setup and control of VSE/ESA.
Total throughput not affected (#tasks etc.)

Refer to the following recent charts on VSE Dynamic Partitions and Tasks, now in this document.

Label Processing Enhancements

New operand SLAADDR in LABEL macro speeds up label processing, especially beneficial for deleting labels.

Refer to detailed discussion.

Label Area utilization display

Helps to avoid awkward situations with 'Label Area full'.

Refer to detailed discussion.

STDOPT and SETPARM SYSTEM no more require BG

Productivity enhancements if BG not available for whatsoever reason

VSE/ESA 2.5 Performance Item Summary ...

- **VSE Base Enhancements (cont'd)**

Up to 32 Library extents in VSAM space

New parameter EXTENTS(MAX16|MAX32) for LIBR DEFINE cmd.

Was 16, as it still holds in BAM space.

Still, using few larger extents should be preferred.

For LIBR performance hints, e.g. OPENs, refer to the separate VSE/ESA Librarian part, which now was moved into this document.

SDAID support on n-ways

No processors need to be stopped during trace interval.

CPU-time impact on partitions still depends on scope of trace.

Increased maximum blocksize in DTFMT

Maximum BLKSIZE in DTFMT increased from 32K to 64K

Reduce number of tape-I/Os for most tape appplications (savings in tape space are minor).

More SECTVAL SVC avoidance

A BALR interface (no more SVC75) used for

- BAM channel programs with fixed length records
- Common VTOC Handler (CVH) I/Os

LE/VSE 1.4.1

Refer to separate charts

EZA Interfaces (from OS/390)

- EZASMI macro I/F (HLASM)
- EZASOKET CALL I/F (COBOL,PL/I,HLASM)

VSE/ESA 2.5 Performance Item Summary ...

Ù **VSE/POWER (6.5) Enhancements**

Support of more than 10 dynamic classes

Refer to charts in Part 'VSE/ESA Base Enhancements'

Now, 2 productivity related items:

Access to Active Queue Entries

Entries for already finished 'jobs'.

Entries can be updated by a single task AND browsed by multiple others CONCURRENTLY

Access to In-Creation Queue Entries

Entries for still running 'jobs'.

Show and browse entries still in creation via 'PDISPLAY CRE'
e.g. LST entries before first (or only) segment is finished

Command Driven Output Segmentation

Allow segmenting of SYSLST output via new command.

Also retrofitted VSE/ESA releases before 2.5.

Refer to separate discussion

POWER PNET (also) over TCP/IP

Exploit the Internet or intranets for PNET w/o cumbersome (?) setup via FTP.

Refer to separate discussion

VSE/ESA 2.5 Performance Item Summary ...

Ù **VSE/VSAM (6.5) Enhancements**

Refer to separate discussion

VSAM LSR Hashing

Allows exploitation of huge amounts of data in memory with fast access to data (no CPU-time increase by searches)

VSAM B/R support from Snapped Volumes:

- 'IXFP/SnapShot for VSAM files'
- 'ESS FlashCopy for VSAM files'

IDCAMS can now:

- SNAP total volumes
- create 'synonym backups' via VSAM B/R

Ù **CICS TS Related Enhancements (1.1.1)**

Refer to separate discussion

CICS TS internal changes

Required e.g. for CICS Web Support

CICS TS DSA display

Concise storage summary via IUI, or via IEDC transaction

Use of subpools for all ICCF GETVIS requests

Helps in case of GETVIS problem analysis (orphaned storage)

CICS Listener Enabler in CICS TS

This shipped code allows exploitation of TCP/IP applications with the GIVESOKET, TAKESOKET TCP/IP I/F, used in OS/390 TCP/IP applications.

VSE/ESA 2.5 Performance Item Summary ...

VSE/ESA 2.5 Performance Item Summary ...

Ù **e-Business Related Enhancements**

No Java Virtual Machine on VSE, BUT ...

Connectivity improvements via connectors

- Connector clients (outside VSE)
- Connector servers (inside VSE)

Type of Connectors

- MQSeries connectors

Start CICS transactions, via MQSeries Server for Windows

- CICS connectors

CICS Transaction Gateway

- VSE Java Beans connectors (new)

(Also for 2-tier environments)

For VSAM, POWER, ICCF, LIBR, and VSE console access

- DB2-based connectors

(Based on a DB2 infrastructure)

For access to DB2,

and (new) to VSAM and DL/I (via Stored Procedures)

(More) Support of 'Point-in-time' copies

Besides RVA SnapShot, also ESS FlashCopy.

Meeting an ever increasing need for continuous system availability, imposed by e-business requirements

VSE/ESA 2.5 Performance Item Summary ...

Ù **e-business Related Enhancements (cont'd)**

Items that also apply to releases before VSE/ESA 2.5.

Host On Demand

Enable secure 3270 emulation on any browser via middle tier and TN3270: 'remote access'

Also a 'connector'.

CICS Web Support (Interface)

Invoke a CICS TS transaction directly from a standard web browser.

Shipped with CICS TS 1.1.1, thus VSE/ESA 2.5 only

New TCP/IP release 1.4

Also available before VSE/ESA 2.5, via APAR PQ29053. PTF is UQ44071, available since 2000-06-14.

Integral part of VSE/ESA 2.5, key protected

Support Enhancements:

- VSE/ESA e-business connectors
- CICS Web Support
- DB2 Server for VSE V7.1
- POWER PNET

- Enhaced quality and diagnostics
- New set of documentation.

Discussed in the separate performance document 'IBM VSE/ESA TCP/IP Performance Considerations'

References

References

General

- VSE/ESA Hints and Tips, 4rth edition, April 2000. Available via VSE/ESA home page, under Techn. Support Info (1.4 MB PDF file)

Some Manuals new in VSE/ESA 2.5

- VSE/ESA Release Guide, SC33-6718-00

- VSE/ESA e-business Connectors User's Guide, SC33-6719-00

- LE/VSE Release Guide, SC33-6779-00, 09/2000, 255 pages

- TCP/IP for VSE/ESA - IBM Program Setup and Supplementary Information, SC33-6601-03, 399 pages

- CICS TS Internet Guide, SC34-5765-00
- CICS TS Enhancements Guide, SC34-5763-01
- CICS TS External Interfaces Guide, SC33-1669-01

- VSE/ESA Software Newsletter 2000, G225-4508-20 Contains 7 featured articles on VSE/ESA 2.5.

- VSE/ESA Software Newsletter 3rd/4rth Q 2000, G225-4508-21 Contains 10 featured articles on VSE/ESA 2.5.

  Newsletters also available via
  http://www.ibm.com/s390/vse/vsehtmls/newslett.htm

VSE/ESA 2.5 Conference Contributions

- Guide Share, 2000-04-10..12, Bad Wildungen, Germany

- VM/ESA & VSE/ESA Technical Conference, 2000-05-31..06-03, Orlando, Florida

- VM/ESA & VSE/ESA Technical Conference, 2000-06-26..06-28, LaHulpe, Belgium

- WAVV Conference, 2000-10-07..11, Colorado Springs

VSE/ESA Base Enhancements

PART B.

VSE/ESA Base Enhancements

More than 10 Dynamic Classes

More than 10 Dynamic Classes

Ù **More than 10 dynamic classes (per table)**

Now, 23 classes are allowed (per table), same naming as before.

Allows a more flexible setup and control of VSE/ESA.

Total throughput not affected (#tasks etc.).

Related Subject

Ù **Task Related Considerations**

Task types

Task usage displays

Refer to the following charts on dynamic partitions and tasks, now in this document.

VSE/ESA Dynamic Partitions and Tasks

Some Background on Dynamic Partitions

Ù **Definable Dynamic Partitions per table:**
Done in DTR$DYNC.Z (dynamic class table)

Up to 10/23 classes, out of (C-E,G-Z)

- Now, 23 allow to be more flexible/granular in the definitions and, thus, in dispatching.
- Benefits if partition sizes are better adjusted:
  - less VSIZE reserved (VSIZE costs disk space for the PDS, except if unused on RVA).
  - NOT affected are ...
    Real storage, Shared Space, and CPU-time consumption

Up to 32 partitions per class

- Sufficient

Ù **Number of Concurr. Active Dynamic Partitions:**

Up to about 200, so far high enough.
Depends on

- Number of remaining VSE tasks for dynamic partitions
Refer to task considerations.

- Remaining VSIZE

- SYS NPARTS value specified
Includes up to 12 static partitions.

Roughly 1K System GETVIS-24 is required per concurrent partition.
Number accepted is limited by available VSE main-tasks.

VSE/ESA Tasks

Some Background on VSE Tasks

Ù **Up to 255 VSE tasks in total**

- Main- , Sub-, and System tasks
Includes the VSE main-tasks (1 per partit.) and all subtasks.

It also includes up to 32 VSE System Tasks.

Private sub-tasks are not included in this 'limitation'
(e.g. 'pseudo-tasks' used by TCP/IP, POWER, ADABAS ...)

Ù **Up to 31 sub-tasks per partition**
If VSE subtasking is used, do not assume any specific dispatch sequence of sub-tasks. Naturally, the dispatching considers the availability of the non-parallel system state, e.g..

Ù **PSF (Print Services Facility) needs 1 sub-task per PSF printer**
Only up to 31 PSF printers per PSF partition

Ù **Direct message for task shortage**
The POWER partition directly displays a message, (1QA0I) that no more (sub-)tasks are available to be ATTACHed in its partition.

Ù **More on partition/task capacity**

As long as no page-I/Os are introduced ...
(in VSE or as a V=V guest in VM)

- more concurrency potentially increases total throughput

- the overhead by VSE task switching is not high
(if at all only slight increase per partition)

VSE/ESA Task Related Displays

Some Task Related VSE Displays

Ù **SIR command (VSE tasks)**

```
...
TASKS ATT.= xxxx   HIGH MARK = xxxx    MAX = xxxx
...
```

This supervisor display shows the current number as a snapshot, the high-water mark, and the maximum achievable value of VSE tasks that can be ATTACHed:
MAX = MIN (208, 256-32-NPARTS)

NPARTS main-tasks are 'pre-attached', so are no more available for attaching sub-tasks.

> A VSE task bottleneck may start only as soon as HIGH MARK has approached the MAX value.

Ù **D STATUS (POWER tasks)**

```
...
MAX. NO. OF TASKS ACTIVE AT ONE POINT IN TIME:  xx TASKS
...
```

This POWER display relates to POWER tasks and NOT to VSE tasks. So, they cannot match to other displays of task high-water marks.

Essentially, these are tasks for POWER functions, plus >1 tasks per PSTARTed partition. Depends on number of spooled devices (Refer to VSE/POWER Diagnostic Reference Manual, Appendix D).

Ù **SIR MON statistics**
Bound conditions in the SIR MON statistics for the TD may give hints to VSE task shortage

Ù **STATUS AR-command**
No specific task indication is shown.

When no more tasks can be ATTACHed, it is up to the application what action is done. If it is decided to wait for a task (instead of cancel), the task is set in normal WAIT (formally shown as 'I/O bound').

VSE/ESA 2.5 Label Enhancements

VSE/ESA Label Processing Enhancements

Ù **New operand SLAADDR in LABEL macro**

„ **Avoids the SVC4 (LOAD) for $IJBSLA**

Provides the start address of this phase

Of performance benefit especially for all cases of repeated use of the LABEL macro
e.g.
- when using LSERV to print/display labels
- for OEM disk/tape management systems (which often add/delete labels)
- for any deletion of labels (see the following explanation)

Label Delete Activity

This new LABEL macro is especially beneficial for deleting labels, since this function still does not exist:

Instead, all labels of an area have to be read and all others to be written back.

„ **Recommendations**

Í **Use the new LABEL macro with SLAADDR wherever/whenever possible**

If the new label macro is used at compile time (also important for specific vendor programs), it will run also on other VSE/ESA V2 releases (with or without the SLAADDR operand).

You may check whether your vendor exploits that function!

Í **Faster label processing, if exploited**

VSE/ESA 2.5 Label Enhancements ...

VSE/ESA Label Enhancements (cont'd)

Ù **Label Area utilization display and capacity:**

New counters for number of label area segments (LASs, 2K each)

(in case of Label Area on Virtual Disk):

A new function CPCTYLBL of the LABEL macro provides

```
#LASs:   available,   currently used,   max used
```

Each LAS
- consists of 4 512 byte blocks as an 'allocation unit' for the Label Area
- may contain between 19 and 24 labels ('label records').

This allows e.g. vendor programs to avoid 'Label Area full' conditions in programs with massive label handling.

It is not intended/recommended to use this in order to reduce the size of the label area, since ...

- this is only a negligible part of DSIZE out of VSIZE on the Page Data Set
- it is really annoying seeing 'Label area full'

Directly displayed as part of the SIR command output

```
LBLS.USED= 00045    HIGH-MARK= 00056    MAX = 00717
```

In this example, the definition of the Virtual Disk for the Label Area was

```
VDISK UNIT=FDF,BLKS=2880,VOLID=VDIDLA,USAGE=DLA
```

which, in average may allow up to about 9000 labels.

It is not recommended to specify BLKS > 2880, since this cannot be used for the Label Area.

Í **Better control of Label Area capacity**

Conditional JCL Enhancements

Conditional JCL Enhancements

Ù **New OPTION SLISKIP available, to speed up JCL GOTO statements**

```
OPTION NOSLISKIP: as previous situation (default)

OPTION SLISKIP:   new function, described below
```

This new settable option SLISKIP will allow to ignore at GOTO statements in conditional JCL that * $$ SLI JECL statements have to be inspected in order to find the target label, refer to APAR DY45423.

If used, it significantly can speed up e.g. situations where a program abends (job termination).

Due to its usefulness, this function was also retrofitted via PTF to previous VSE/ESA V2 releases:

Release	PTF
VSE/ESA 2.5	Fct Included
VSE/ESA 2.4	UD51424
2.3	UD51422
2.1/2.2	UD51418

LE/VSE 1.4.1

LE/VSE 1.4.1

For general LE performance information, refer to
- 'LE Performance Hints' in the VSE/ESA V2 Base document

„ **New modification level with new CLC number**

„ **Automatically installed during the installation of VSE/ESA 2.5**

„ **Improved performance for COBOL/VSE**

- Reduced number of SVCs used for date routines

- Enhancements to statically called COBOL/VSE subroutines from a 'main' program (Sysroute of LE/MVS APAR PQ11742)

„ **C-Optimization**

Most of the C-runtime environment modules were re-compiled with optimization included (OPT in the C-compiler).

„ **Run-Time Option HEAPCHK**

New since VSE/ESA 2.3

Provides a checking facility to verify that the heap storage has not been damaged. Intended for test environments.

Shipped default is HEAPCHK(OFF,1,0).

Use HEAPCHK in production only when necessary, as it will use extra CPU time.

„ **Working Storage Optimization**

Working Storage, if below the line, uses 2x8 byte areas (both in CICS/VSE and in CICS TS) per 4K page.
Thus, to max it out, such areas could be specified as multiples of 4080 byte.

For more info refer to the LE/VSE 1.4.1 Release Guide

LE/VSE 1.4.1 ...

LE/VSE 1.4.1 (cont'd)

„ **New subparameter MIN|MAX for TRAP(ON) for LE Batch condition handling**

MIN instructs LE/VSE, not to use any STXIT AB processing for LE/VSE batch condition handling and to use only STXIT PC for internal condition handling:

TRAP setting	(ON,MAX)	(ON,MIN)	(OFF)
LE condition handling	Full	Partial	None
Type of STXIT calls issued by LE batch, when LE pgm is 'called'	AB PC	- PC	- -

AB = ABEND Handling (more frequently used, since done e.g. by SORTs on a per record basis)
PC = Pgm Check Handling

- TRAP(ON,MAX) is default, and corresponds to LE 1.4.0 TRAP(ON).
- TRAP(OFF) should not be used

MAX and MIN are ignored under CICS

Under CICS, no STXITs are used by LE.
Thus, TRAP(ON) is not a performance issue under CICS.

Í **TRAP(ON,MAX) still needed for full LE condition handling**

Í **TRAP(ON,MIN) may be used in certain cases**

(with better performance for STXIT uses, e.g. internal SORTs)

LE/VSE 1.4.1 ...

LE/VSE 1.4.1 (cont'd)

„ **Changes in CEExOPT.A (x= C,D,U) members for run-time option STORAGE**

```
STORAGE(heap_alloc_value,heap_free_value,dsa_alloc_value,
                                              reserve_size)
          e.g.  STORAGE(00,NONE,NONE,0K)
```

This option controls the initial content ('value') of some storage areas:
00 means initilization with 0's, NONE means no initialization.

Also the amount of storage reserved for out-of-storage conditions ('reserve_size'):

- Default value is 8K
- Shipped value for Batch (CEEDOPT) was increased to 32K
- Value for CICS (CEECOPT) should ALWAYS be 0, since area not used under CICS.

For PL/I, STORAGE(00,NONE,00,xK) is recommended for compatibility reasons, in case no initialization is done.

Some performance impact, depending on the amount of STORAGE acquisitions.

„ **Shipped and improved SVA load books**

Relevant LE performance PTFs saving SVA-24 space are included and thus now allow a better customized setup of SVA-24 and SVA-31 loading.

Also, double SVA loads from different SVA load books are avoided.

Refer to separate charts.

For more info refer to

- LE/VSE Release Guide, SC33-6779-00
- Developments Down Under, by Jim Alexander VM and VSE Tech Conf Orlando, 06/2000, session E27
- the LE/VSE home page at http://www.s390.ibm.com/le_vse/

LE/VSE Phases and SVA (VLA)

LE/VSE Phases and SVA (VLA)

Be sure to have understood the performance aspects of VLA and the principal rules for SVA loads (different for SVA-24 and SVA-31).

Refer to Appendix A: VSE Space Optimization under 'VLA General Usage Aspects' in the VSE/ESA V2 Performance Considerations document.

„ **General**

- There is no functional requirement to load any LE phase into the SVA (VLA)
- Loading any phase into VLA may save
  - LOAD I/Os
  - VSIZE, if many partitions use LE concurrently
- Shipped LE SVA load books include both 24 and 31 bit phases. Carefully check the needs below and above!

„ **VLA-24 for LE**

Loading SVA-24 phases for LE in general is NOT recommended before APAR PQ23382.

- The following SVA-24 eligible LE Base phases are only loaded once per job step (not relevant for CICS, only for short batch job steps/pgms):

	Old	New	
CEEBINIT	294K	53K	LE batch
CEECCICS	300K	47K	LE CICS
CEEPIPI	305K	65K	LE pre-initialization
CEEPLPKD	74K		LE DTF builders (mainly)

Do NOT load the old versions of these phases into the SVA-24. Apply the PTF UQ27971 for APAR PQ23382 to get the new ones, and put them into the SVA-24.

Refer to the SVA load books below.

Cont'd

LE/VSE Phases and SVA (VLA) ...

LE/VSE Phases and SVA (VLA) (cont'd)

„ **VLA-31 for LE**

You may load the following LE runtime phases into VLA-31 (they are used once per LE-program execution, and thus of special benefit for Batch (non-CICS) use):

```
CEEEV003  965K C  for VSE/ESA Runtime (recommended in any case)
CEEEV005   15K COBOL   -"-
CEEEV010  196K PL/I    -"-
CEEPLPKA  310K LE      -"-          (new, offload from SVA-24)
```

„ **Shipped LE/VSE 1.4.1 SVA load books**

SVA Load book	Can be used in	24-bit members/size	31-bit members/size	Use
$SVACEE	LE Base	9 / 270K	22 /1037K	Rec *
$SVAIGZM $SVAIGZ	LE COBOL	0 / 0K 32 / 165K	3 / 128K 15 / 344K	Rec Opt
$SVAIBMM $SVAIBM	LE PL/I	1 / 36K 1 / 36K	2 / 202K 27 / 280K	Opt Rec
$SVAEDCM $SVAEDC	LE C	0 / 0K 1 / 0.2K	6 /1776K 18 /2324K	Rec * Opt

- Number and total size of members are approximate. Here LE/VSE 1.4.1 in VSE/ESA 2.5.1 is considered.
- Use only 1 of 2 load books for each language (would cause double loads and waste of space)
- Rec means recommended, provided language is used
- Opt means optional, instead of other book
- * These books are included in SKLJCL0 in ICCF lib 59, and used directly in case of Base Install.

Note that CICS TS only uses these LE modules from the SVA, if SVA=YES is specified, plus USESVACOPY=YES.
But be cautious, in case you both have CICS/VSE and CICS TS in use, regarding the loading of DFH* phases into the SVA.
Refer to the CICS TS document.

- Avoid any multiple SVA load of a phase
- Make PSIZE-31 big enough to avoid 'downloading' SVA-31 phases into SVA-24

VSE/VSAM Enhancements

PART C.

VSE/VSAM Enhancements

VSAM LSR Hashing -Perf. Aspects-

VSAM LSR Hashing

„ **Fast access to buffer in LSR subpool via hashing. Saves CPU cycles by eliminating long searches**

Applicability of implemented hashed access

„ **VSAM LSR (not NSR)**

LSR is also applicable in case of only 1 file per subpool, even if only a single string would be used.

Some info on NSR (which is NOT hashed) vs LSR:

- NSR buffers are subdivided into strings with
  - no lookaside to other strings,
  - no synchronization of updates, except for SHROPT 4 via buffer invalidation
- NSR buffer contents is more stable than LSR (since not shared across files)
- For READ Ahead, more NSR buffers than CIs per CA are not beneficial (No READs across CAs). LSR does not provide READ Ahead for sequential access
- In NSR, index buffers are always separate

„ **Any type of VSAM file (not only KSDS)**

ESDS, VRDS, RRDS, compressed or not, XXL or not (XXL uses the CI-number as 'RBA').

SHROPT 4 is not suited for DIM due to buffer invalidation after a random READ.

„ **Access to data and to index buffers**

„ **Any type of LSR request**

MOVE-mode (the only one in CICS), LOCATE mode.
Record mode or CI-mode.
User Buffering (used e.g. by DL/I and DB2/VSE to manage their own buffer pools) is NOT available under LSR.

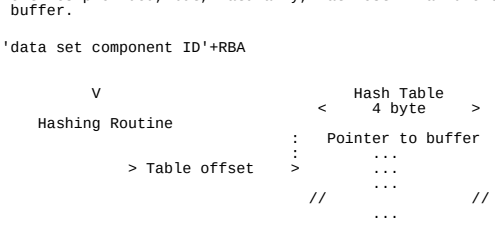
VSAM LSR Hashing -Perf. Aspects- ...

Hash Tables

„ **1 Hash table per LSR subpool**

Each entry (4 byte) contains a pointer to a chain of 1 (or more) data or index buffer control block(s), which themselve point(s) to the buffer address(es).

For ease of implementation and to minimize 'collisions', for NBUFS buffers in a subpool there are in total (2xNBUFS -1) entries provided, but, naturally, at most 1 valid entry per buffer.



In theory, different inputs to the hashing routine may result in the same table offset. This is no problem here, since the buffer control block contains a data set identification as well as the CI ID.

All hashing tables and control blocks may reside in partition GETVIS-31. They actually reside where the LSR buffers are (RMODE31 in the VSAM BLDVRP macro).

„ **Hash Table Maintenance**

Each hash table is 'initiated' at BLDVRP time with all (2xNBUFS)-1 pointers being zeroes (no information on the file(s) or on any data is required).

When buffer contents (i.e. the RBA of the associated CI) changes, the hash table is updated.

í **Building/maintenance of hash tables is fast**

VSAM LSR Searches

„ **Hashing Algorithm**

Refer to the following short article:

'VSE/VSAM Buffer Hashing for LSR Pools with VSE/ESA V2.5'
by Guenter Weigelt, VSE/ESA Software Newsletter, 2000

VSAM LSR Searches

Type of Searches

Type of Search	Frequency	Search length (w/o hashing)	Now Hashed
Record Search	Always	up to NBUFS	X
Exclusive Control S.	Before PUTUs	NBUFS	X
Dublicate RBA Search	After PUTs	NBUFS	X

- NBUFS is the number of buffers in an LSR subpool
- PUT may be a PUTU (PUT after Update) or a PUT New
- Refer to explanations below

„ **Record Search (Data or Index)**

Argument is 'data set component-ID' (includes index/data specification) plus the RBA.

Up until now, buffers had to be searched serially until the needed record was found, or until all buffers of the subpool were examined.

The search always started at that buffer, where the last request (of any type) was satisfied, doing wraparound, if required.

On the average, 50% of all buffers were checked in case of a hit (assuming random access),
in case of a miss, all buffers were checked.

This scanning search was the main consumer of search CPU time and is NOW a hashing search

VSAM LSR Searches ...

VSAM LSR Searches (cont'd)

„ **Exclusive Control Search**

Whenever a record was to be modified (UPDATE), it had to be made sure that this LSR CI is not in use by others.

The search always started at the begin of the buffer chain, as buffers were setup initially.

So far, all buffers had to be searched.

This search NOW is also hashed

„ **Duplicate RBA Search**

Whenever a record was written (PUT), VSAM makes sure that the just updated CI is really the only occurence of these data with this RBA in any buffer.

If not unique, the other buffer(s) is invalidated.

This may be required in case of multi-threading, e.g. if a record was written by a different VSAM string.

This search always was a search through all NBUFS buffers, and now is hashed (in case of LSR), not for NSR.

This search NOW is also hashed

„ **More Aspects**

Only small base hashing overhead

Measurements have shown that LSR hashing only shows a small base overhead. Thus, all accesses to LSR buffer pools are hashed, even for small values of NBUFS.

LSR hashing used whenever available

LSR hashing is used whenever available in a VSAM release

VSAM Hashing Sample Results

VSAM Hashing Sample Results

„ **VSE/ESA Environment**

- VSE/ESA with or w/o the new hashing scheme
- 2003-225 processor
- 2 VSAM files used with same DEFINE CLUSTERs:
  - SHR(2)
  - CISIZE 4K
  - RECORDSIZE 80

	FILEA	FILEB
#index levels	1	2
content	100 records (2 CIs)	100,000 records (1960 CIs)

„ **Tests Performed**

READ/UPDATE/BROWSE/LONGBROWSE: 'RUBL'

Here measured with CICS transactions under CICS TS, BUT also CICS/VSE and Batch applies for LSR hashing.

FILEA:
- R: 10.000 x (READ same record)
- U: 10.000 x (READ same record and update it) Includes 10.000 I/Os for all NBUFS values
- B: 10.000 x (S-BROWSE,READNEXT,E-BROWSE same record)
- L: 10.000 x (READNEXT 10 x the same record sequence)

FILEB:
- R: 10.000 x (READ same record)
- B: 10.000 x (S-BROWSE,READNEXT,E-BROWSE same record)
- L: S-BROWSE,(100.000 x READNEXT),E-BROWSE Includes I/Os to load buffers, depending on NBUFS.
  - 1st time also includes build up of hash table.

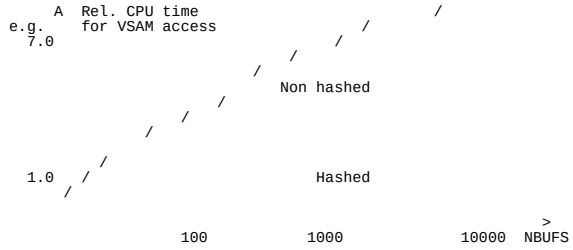
VSAM Hashing Sample Results ...

„ **Some Test Results**

NBUFS	3		1000		3000	
Hashing	no	yes	no	yes	no	yes
FILEA R	603	700	3,794	720	8,826	704
FILEA U	9,664	9,040	16,700	9,050	31,125	9,045
FILEA B	2,250	2,366	5,554	2,301	10,795	2,338
FILEA L	508	505	839	488	1,371	497
FILEB R	708	823	3,172	829	8,778	825
FILEB B	2,379	2,519	4,916	2,456	10,747	2,485
FILEB L 1st	3,847	4,139	9,013	6,886	18,403	11,678
2nd	3,838	4,139	9,127	7,160	8,428	5,736

- All CPU-times in msec
- Emphasis is on variation with NBUFS, not on absolute values or on relative values to other test-cases

„ **Principal Behavior of all LSR Accesses**



í **VSAM CPU-time with hashing is independent of the number of buffers**

Maximum value is 32768 buffers

Regarding potential real-life benefits for applications, refer to next chart

VSAM Hashing Sample Results ...

Potential Real-life Benefits for Applications

Note that the test examples above are worst-case test cases, just doing VSAM LSR requests, without any application logic.

„ **Applicability**

All applications with VSAM LSR, mostly online transactions, but also batch LSR applications exist and would benefit.

„ **Expected Possible Real Life Improvements**

Online example (first day)

Assume e.g. that currently (very roughly)

- 80% of your CICS transactions use VSAM LSR
- the average NBUFS value used is 200
- about 25% of your LSR transaction CPU-time is VSAM

-> about .8x.25 = 20% of your CICS load is VSAM LSR

This VSAM LSR CPU-time may be reduced roughly by a factor of 3 (or whatever factor you can deduce from the test results), resulting overall in roughly 10% lower CICS CPU time per txn.

Online example (more DIM)

With VSAM hashing, the number of buffers can now be increased to exploit large buffer areas

- w/o any increase in CPU-time for VSAM access
- with further reduction of VSAM I/Os for READs and thus of CPU-time

í **Exploit large DIM with even reduced CPU-time**

Buffer Related VSAM Tuning Hints

VSAM NSR Tuning Hints

í **Use larger CI-sizes to reduce the number of buffers in NSR (or non-hashed LSR) subpools**

This is not only applicable to sequentially used files, but also to random access to non-hashed bigger buffer pools.

But avoid to use excessive CI-sizes due to VSAM locking and invalidation on CI-basis during UPDATEs.

VSAM LSR Tuning Hints

í **Ensure that all LSR buffers, hashing tables and control blocks reside above the line:**

Specify RMODE-31=ALL or BUFF in the VSAM BLDVRP macro for batch applications.

Note that

- NONE is the batch default, don't use.
- CICS always allows LSR pools to reside above the line

í **Specify/Allow enough partition GETVIS-31 to actually have all LSR areas above the line**

if the partition has any possible 24-bit constraints

VSE/ESA with Huge Processor Storage

Situations more and more coming up

- **'I have 2 or 3 GB of processor storage available for a single VSE/ESA production. How can I exploit it for Data In Memory (DIM) and benefit?'**

Answer

- **Exploit VSE Virtual Disk, wherever possible**
  Also e.g. big GETVIS or data spaces for SORTs.
- **Apply DIM to all kind of randomly accessed files.**

  Do it in a reasonable way, as suggested below for VSAM KSDS.
  Avoid that saved I/Os and thus saved CPU-time is (over-) compensated by long buffer searches (only possible w/o LSR Hashing).
  If SHROPT 4 cannot be avoided, reduce its overhead.
- **Avoid introduction of paging I/Os**
  As is well known, this would be very disadvantageous for CICS.

  For 2 GB real, a reasonable VSE VSIZE may be about 5 to 10 GB, very roughly estimated.

Suggested DIM exploitation (KSDS)

- **For BATCH files**
  1. **Try to use LSR, if that should be possible**
  2. **Else, use NSR, with up to say 200 buffers (NBUFS)**
  3. **Use OPTI-Cache to cache I/O data globally across VSE partitions**

VSE/ESA with Huge Processor Storage ...

Suggested DIM exploitation (cont'd)

- **For CICS files**
  1. **Try to use CICS Data Tables**
  2. **Use (hashed) LSR with many buffers**
  3. **Use OPTI-Cache to exploit >2 GB real**
- **OPTI-Cache in general**

  Supports S/390 Expanded Storage

  Since Expanded Storage is not supported by VSE/ESA, Opti-Cache handles error situations, transparently to VSE.

  Expanded storage is used as a level 2 cache for partition storage (L1) in VSE.

  Beneficial in order to reduce/avoid SHROPT 4 overhead

  For more info on OPTI-Cache, refer to
  http://www.bsiopti.com

For VM/VSE

- **Same rules as above for native VSE**
- **Use VM Minidisk Caching (MDC), if OPTI-Cache is not your choice**
  OPTI-Cache would also reduce total CPU time, compared to VM MDC

VSAM Backup from Snapped Volumes

VSAM Backup from Snapped Volumes

This line item also may be known or be referred to as
'IXFP SnapShot for VSAM Datasets'

- **Any VSAM file on a snapped volume can now be backup'ed for potential later use**

  - for Recovery of Files
  - for Data Duplication

  So far, reuse of VSAM files from snapped volumes was not possible due to VSAM issues with non-unique names, created by copied volumes.
  There MUST NOT BE

  - duplicate catalog names in 1 VSE
  - duplicate VSAM dataset names in 1 catalog
  - duplicate disk VOLIDs with VSAM objects in 1 VSE

  Now, the new Synonym List is used
  (contains the relationship of source and target VOLIDs)

  - to route VSAM OPEN and the Backup READ routines to the snapped target volumes
  - to avoid and correctly handle non-unique names (catalog and disk data remain unchanged)
- **Applicability**

  The implementation within VSAM is independent of the specific type of I/O subsystem.

  It applies both to RVA SnapShot and to ESS FlashCopy

VSAM Backup from Snapped Volumes ...

VSAM Backup from Snapped Volumes (cont'd)

- **Series of User Actions (all in IDCAMS)**

  1. **Use new command 'SNAP'**
     Rapidly snap-copy VSE volumes:

     Do it for all volumes which contain data or catalog information of the VSAM files to be backup'ed:

     - catalog volume: user catalog volume OR master catalog volume
     - user volumes: all volumes with extents of this VSAM dataset (via LISTCAT)

     New VOLIDs for the target volumes are required (to allow the target volumes to become DVCUP for Backup).

     (If the dataset is in the master catalog, this catalog is renamed and connected as a new user catalog to the master catalog)

     Continue production

  2. **Do IMPORT CONNECT**
     Connect copied catalog under different name

  3. **Do Synonym BACKUP**
     VSAM files backup are now exactly on tape as if done from the original file
     (includes correct(ed) catalog information)

  4. **Do RESTORE (optional)**
     VSAM file will be uniquely accessible (under same or a different name)

VSAM Backup from Snapped Volumes ...

VSAM Backup from Snapped Volumes (cont'd)

„ **Performance Benefit**

Only a very small batch window required, since the backup is done from snapped volumes

„ **Performance Hints**

Í **Do not put large heavily updated non-VSAM files on those disks**

This would mean that non-VSAM source track updates cause unnecessarily that the before-image of such tracks are also kept, increasing the NCL on the RVA.

But, note that the I/O subsystems really can hide a lot of background activities to the host side.

Í **Release the snapped VSAM volumes after backup is completed**

Applies to RVA and to ESS.

The ESS (in contrast to the RVA) always needs the full physical capacity of the target volumes to be reserved.

Refer to the FlashCopy charts in the ESS part

VSE/POWER Enhancements

PART D.

VSE/POWER Enhancements

Enhancements besides ...

- More than 10 dynamic classes
- Access to Active and In-Creation Queue entries

Command Driven Output Segmentation

Command Driven Output Segmentation

Ù **Background Info**

POWER output segmentation currently is possible as

- Count-Driven Output Segmentation
  (segment after a certain number of pages or spooled lines)
- Data-Driven Output Segmentation
  (segment when new $LST card encountered)
- Program-Driven Output Segmentation
  (segment when output segmentation macro is encountered:
  SEGMENT, or better, new IPWSEGM)
- Multivolume Tape Segmentation
  (segment at tape end)

Ù **New command for SYSLST output segmenting**

```
PALTER LST,jobname,jobnumber,CQNUM=qnum,SEGMENT=xxx
```

New SEGMENT operand: PAGE segment at next appropriate boundary
CARD same as PAGE
IMM segment immediately

Also retrofitted VSE/ESA releases before 2.5:

VSE/ESA	POWER	APAR	PTF(s)
1.4	5.2	DY45294	UD51277
2.1/ 2.2	6.1	DY45293	UD51276 UD51275
2.3	6.3	DY45292	UD51274 UD51273
2.4	6.4	DY45291	UD51271 UD51270

VSE/ESA 2.5 IUI panel will be adapted, too.

Í **Start printing possible before end of POWER job**

Beneficial in case no segmenting was pre-planned

POWER PNET over TCP/IP

POWER PNET over TCP/IP

Ù **Background Info**

- A new VSE subtask is used under the Turbo Dispatcher for POWER related TCP/IP activities.

  This task has a higher priority than the POWER maintask.

  To help avoid TCP/IP overflow, RECEIVE is given higher priority than SEND.

- Exploits the new EZASIM macro interface

Ù **Performance Hints**

Í **Select a reasonably good PNET buffer size**

BUFSIZE in PNODE macro
(300 - 32000 byte, with 4096 as default)

For CTC (reliable) use a high BUFSIZE.

Í **Run 'double' buffered**

MAXBUF in PNODE macro:
Number of buffers for every PNET receiver and transmitter task.
Default is (3,3)

Usually, up to 8 buffers should be sufficient

Ù **General PNET Performance Hints**

Refer to 'POWER PNET Considerations' (E35/E36) in

' IBM VSE/ESA V2 Performance Considerations'

CICS TS Related Enhancements

PART E.

CICS TS Related Enhancements

Here CICS Transaction Server for VSE/ESA 1.1.1 is considered.

Refer also to the charts on CICS Web Support,
in the e-business part of this document.

For CICS TS 1.1.0, refer to the separate CICS TS performance document.

CICS TS DSA Display

CICS TS DSA Display

- **Quick display of DSA allocation and usage**
  for debugging
  for tuning purposes
- **Can be done via IUI (fastpath 364),**
  or via native CICS TS transaction IEDC
- **Output better/more than with 'CEMT Inquire DSAs'**

```
IESADMDCST              CICS TS STORAGE REPORT     Time: 09:40:19

Applid: DBDCCICS  Sysid: CIC1  Jobname: CICSICCF CICS TS Level: 110
Storage Protection .... INACTIVE    Reentrant Programs .... PROTECT
                                      CICS Trace Table size..   80

Extended DSA:                (All sizes in kbyte)    LIMIT   25600

                            ECDSA   EUDSA   ESDSA   ERDSA  Totals
  Current DSA Size .....  2048    1024    1024    6144  10240
  Current DSA used .....  1876      64       8    5220   7168
 *Peak DSA used .......   1884      64       8    5220
  Peak DSA Size ........  2048    1024    1024    6144  10240
  Largest free area/Free Stor. 0.95    1.00    1.00    0.94
  Times short-on-storage (SOS)    0       0       0       0      0

DSA:                                                    LIMIT    5120

                             CDSA    UDSA    SDSA    RDSA  Totals
  Current DSA Size .....    512     256     512     512   1792
  Current DSA used .....    344       8     456     344   1152
 *Peak DSA used .......     352      28     456     344
  Peak DSA Size ........    512     256     512     512   1792
  Largest free area/Free Stor. 0.88    1.00    0.86    0.86
  Times short-on-storage (SOS)    0       0       0       0      0

PF1=HELP      2=REFRESH   3=END     4=RETURN


The * in front of the 2 lines 'Peak DSA used' mean that these values
are reset in case the DSA size changed
```


VSE/ICCF GETVIS Subpool Usage

VSE/ICCF GETVIS Subpool Usage

Use of subpools for all ICCF GETVIS requests

Besides the 'old' subpool ICCFFN, 2 new subpools are used:

Subpool-ID	Purpose	Remark	Size
ICCFFN	Control program	Fixed in size	252K
ICCFIP	Interactive partitions	Sizes and numbers may be changed	2832K default
ICCFBU	DTSFILE buffers	Size and number may be changed	44K default
- All ICCF GETVIS requests are for 24-bit storage			

Benefits

- **Better visibility of GETVIS-24 requirements**
- **Helps to debug GETVIS creep problems (orphaned storage)**

This enhancement was retrofitted also to VSE/ESA 2.4 (09/99, UQ34180) and to VSE/ESA 2.3 (12/99, UQ38110).

CICS Listener Enabler in CICS TS

CICS Listener Enabler in CICS TS 1.1.1

Code was ported from the OS/390 CICS Sockets Interface and the CICS Listener. Is part of VSE Base.

- **A shipped TCP/IP application base**

  Exploits the GIVESOKET/TAKESOKET TCP/IP services, used in OS/390 TCP/IP applications.

- **Based on CICS TS terminal management**

  Accept input from a terminal and pass data over to a specified transaction (TCP/IP application) for execution

- **Implementation**

  The Listener task is implemented as a never-ending CICS TS transaction (EZAL), which

  - is able to listen at TCP/IP ports via VSE TCP/IP calls
  - can transfer data and control to a child server application (e.g. a ported MVS CICS TCP/IP application using the EZASOKET call interface from OS/390 TCP/IP, supplied since TCP/IP 3.2 for MVS)

- **Multiple Listeners possible**

  Multiple Listeners can be active concurrently, each with

  - a unique CICS TS transaction name (even when cloned)
  - a unique port assigned to it
  - a (usually different) transaction ID to which the listener is to transfer data and control

Misc. CICS TS 1.1.1 Items

Misc. CICS TS 1.1.1 Items

- **REXX for CICS**
- **Recent CICS TS APARs and PTFs**
  - **Usage of CICS TS Index buffers for VSAM**

    VSE/ESA 2.4

    APAR/PTF PQ36602/UQ45475 contains also an item of wrong usage of VSAM LSR index buffers.
- **New SIT Parameters in CICS TS 1.1.1**
  - **LEVSE=YES|NO**

    VSE/ESA 2.5 only.

    This SIT parameter was introduced in order to give some additional VSCR to customer having CICS TS partitions not using any LE programs.

    LE=NO avoids that any LE phases are loaded into the CICS TS partition. YES is the default.

    You may keep in mind that the C runtime environment is used for more VSE functions that one would assume usually.
  - **WEBDELAY=(5,60)**

    VSE/ESA 2.5 only.

    Delays for the 3270 bridge:

    - maximum time in minutes, until a terminal wait state is being purged
    - time in minutes, until state data is kept before CICS performes cleanup

CICS TS 1.1.1 Specific Hints

CICS TS 1.1.1 Specific Hints

- **For TCP/IP services (CICS Web Support) ...**
  - **Increase the EDSA by at least 2M**

    plus about 1M per active Web connection

    This recommendation stems for CICS/ESA for MVS
- **More Info**

  Refer also to the foils on

  - CICS Web Support (CWS)
  - CICS 3270 Bridge
  - CICS Transaction Gateway (CTG)

    (all in this document)
  - Reduced size of IUI Logon user records (APAR PQ44663)

    (in the CICS TS document)

e-business Related Enhancements

PART F.

e-business Related Enhancements

For general info on VSE/ESA and e-business, refer e.g. to

VSE Applications - How e-business fits -, GF22-5137, 2000-01-07, 17 pages (by Jerry Johnston and Anette Stolvoort). A position paper

Available via the VSE/ESA home page.

e-business Glossary

e-business Glossary (Selected Terms)

Java programs

Applet	A Java application program that a Web browser can retrieve from a Web server and execute.
Servlet	A Java 'CGI'-application running in a Web server, which sends the results to a Web browser.

Java building blocks

Java Beans	A platform-independent technology for reusable Java components as a Java class with specific properties. Usually visual components, like push-buttons, sliders etc. JavaBeans can be treated via VisualAge.
Enterprise Java Beans (EJB)	A specific type of Java Beans (an enhanced Java class), being able to use 'services', typical of a server environment. Also called a distributed bean (similar to RPC) with the client part under a browser and the server part under a Web Application Server. The EJB classes are managed by a 'container', a platform dependent component, running in an EJB server. 1 DB2 row or VSAM record may be included in an EJB.
VSE Java Beans	Special Java beans that allow access to VSE-based file systems, submit jobs, and access to the VSE console (an API class library). Do not contain visual components. VSE Java Beans can be treated via VisualAge.

- All these building blocks must be used/called from within a Java application
- Enterprise Java Beans and VSE Java Beans are special types of Java Beans.
- When EJB are used for connectors, VSE Java Beans are part of EJB (server).

e-business Glossary ...

e-business Glossary (Selected Terms, cont'd)

Connector terms

Connector	Middleware to connect 2 platforms: - Middle-tier and VSE/ESA host (3-tier) - Web client and VSE/ESA host (2-tier). Connectors in their direct sense allow the use of customized programs outside of VSE (e.g. applets, servlets, or other). Such connectors consist of connector client and connector server (together 'connector'). Sometimes the connector client is called 'connector'. Software that just allows access to VSE/ESA and existing applications via TCP/IP are sometimes also called 'connnectors': TN3270, HoD, CICS Web Support (CWS), ...
VSE Connector Server	Host part of the VSE Java Beans, in a VSE batch partition, started via the STARTVCS job. Includes the VSE Listener. In widest sense, also CICS, or DB2, or MQSeries Server on VSE could be called 'connector servers'.
VSE Listener	Part of the VSE Connector Server which listens to TCP/IP socket requests for input from VSE Java Beans.

e-business Glossary ...

e-business Glossary (Selected Terms, cont'd)

Some server terms (outside VSE)

HTTP Server	A program that understands HTTP requests that can coordinate and assemble Web pages and deliver them.
Web Server	A HTTP server with a Java engine (applets, servlets)
Web Application Server	A program that usually gets called and gets HTML data from the Web Server. It also can assemble Web pages to be sent back by the web server. Often a web server is directly included.
WebSphere Server	An IBM specific implementation of a web application server, allowing writing of applications with web development tools. It allows to run EJBs.
VSE AppletServer	A (usually) small program running under a web server, which simply routes requests from the client applet to the VSE Connector Server.

Miscellaneous

VSE Navigator	Known from VSE/ESA 2.3, then a sample to access VSE services. New version now exploits Java Beans. Server is a C-program in VSE, client is a Java applet on a workstation.
CCF	Common Connector Framework. An IBM architecture for connection management: A consistent means of connecting to, and interacting with, host resources from any Java environment. It is part of VisualAge for Java. VSE Java Beans use the CCF class library transparently.
DB2 Stored Procedure	A program that accesses DB2 data, but also VSAM and DL/I in case of DB2 VSAM and DL/I Connect (DB2 for VSE 6.1 or later).

e-business Glossary ...

e-business Glossary (cont'd)

Some Interfaces

CLI	Call Level Interface. A new VSAM API ('VSAMSQLfct') written in C for VSE/ESA 2.5, callable from LE batch programs. Very similar to and based on the DB2 CLI. Currently used by C programs by DB2-based connectors for VSAM.
ECI	External Call Interface. A remote procedure call I/F for CICS applications with separate presentation logic. Provided by CICS clients.
EPI	External Presentation Interface. Deals with 3270 data streams as input/output for unchanged CICS 3270 applications. It allows to extract/insert data from/into a 3270 stream.
MQI	Message Queue Interface. An API provided by the MQSeries queue managers to allow programs to access message queuing services.
CGI	Common Gateway Interface. A standard method to invoke programs on a server. It provides a server program with a standard interface to the HTTP headers and input data in a browser request and also a standard method for returning a response.
FEPI	CICS Front-End Programming Interface. Enables a CICS transaction to emulate a 3270 terminal. Used for transactions as intermediaries between the web user and the target 3270 transaction.

More Info on VSE e-business

More Info on VSE e-business

- VSE/ESA 2.5 e-business Connectors User's Guide, SC33-6719
- e-business Solutions for VSE/ESA, by A. Ackel et al, IBM Redbook, SG24-5662-00, 06/2000 (VSE/ESA 2.4 level)
- e-business Connectivity for VSE/ESA, by A. Ackel et al, IBM Redbook, SG24-5950-00, 10/2000 (VSE/ESA 2.5 level)
- VSEConnectors.html via VSE/ESA homepage http://www.ibm.com/s390/vse
- VSE/ESA V2.5 Super Connectors, by Wilhelm Mild, WAVV 2000, Colorado Springs, Oct 2000
- An e-business Roadmap for VSE/ESA Customers, by Jerry Johnston, WAVV 2000, Colorado Springs, 10/2000 z/VM,VSE and Linux Tech Conf, Jacksonville 05/2001
- WebSphere Application Server and VSE/ESA, by Ingolf Salm, WAVV 2000, Colorado Springs, Oct 2000
- CICS TS for VSE/ESA: Web Support and 3270 Bridge by Erich Amrehn et al, IBM Redbook, SG24-5997-00, 11/2000
- CICS Transaction Server for VSE/ESA: CICS Web Support Overview. by Chris Smith, IBM Hursley. z/VM,VSE and Linux Tech Conf, Jacksonville 05/2001

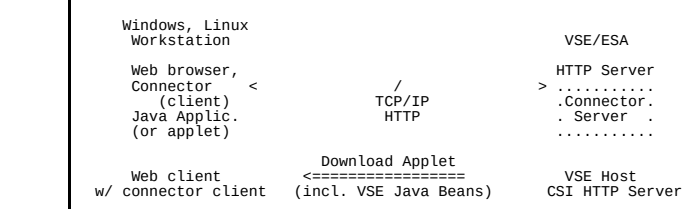
New CICS TS 1.1.1 manuals, with info on CWS:

- CICS TS for VSE/ESA: Enhancement Guide, SC34-5763
- CICS TS for VSE/ESA: Internet Guide, SC34-5765
- CICS TS for VSE/ESA: CICS External INterface Guide, SC33-1669

e-business Connectors for VSE/ESA

2-tier e-business Connector Environments

Web client and VSE/ESA communicate directly

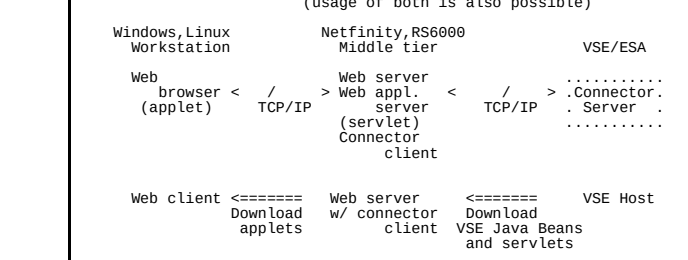


3-tier e-business Connector Environments

Web client and VSE/ESA communicate via a Web server

Middle tier with Web Server and a Web Application Server (usually WebSphere).

Use of VSE Java Beans in applet on browser
or in servlet on Web application server
(usage of both is also possible)



VSE 3-tier Connector Specifics

VSE 3-tier Connector Specifics

WebSphere allows to write applications, using Web development tools, for the middle tier that can access VSE data.

The Web application server coordinates, collects, and assembles Web pages and delivers them to the Web client.

„ **MQSeries connectors**

í **Start any CICS (/VSE or TS) transaction from a web browser and access any file under CICS**

Refer e.g. to:
'What's New in MQSeries', by Christianne Sims,
VM and VSE Tech Conf Orlando, 06/2000

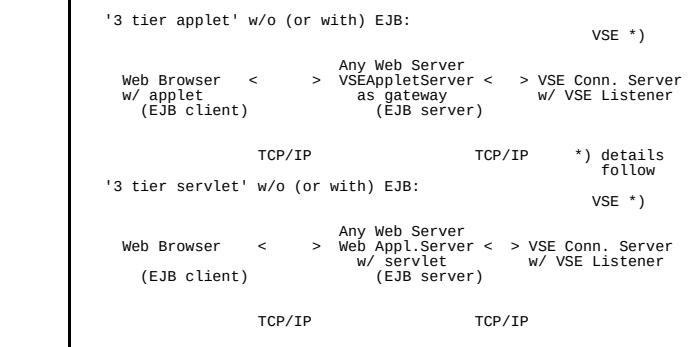
„ **CICS connectors**

í **Start a CICS TS transaction from a web browser and access any VSAM (or DL/I) file under CICS TS**

VSE 3-tier Connector Specifics ...

„ **VSE Java Beans connectors**

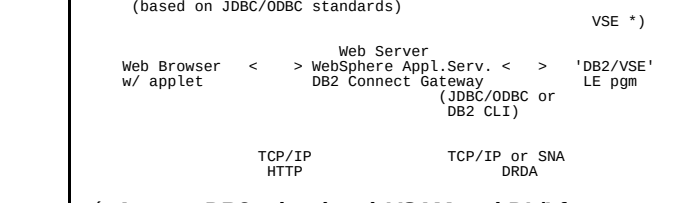
(based on new VSE external services)



í **Access VSAM, POWER files, VSE libraries and ICCF from a web browser**

„ **DB2-based connectors**

(based on JDBC/ODBC standards)



í **Access DB2, plus (new) VSAM and DL/I from a web browser, or start any CICS TS transaction**

VSE/ESA Details for Connectors

VSE/ESA Connector Server Implementations (IBM)

„ **MQSeries connectors**

Connector server runs directly in a CICS partition.

„ **CICS connectors**

Connector server runs directly in a CICS TS partition.

„ **VSE Java Beans connectors**

BSD Sockets used on the Connector Server side.

VSE Connector Server (with listener as a TCP/IP application) runs in a separate VSE partition (only once).
It is a plug-in in C, using VSE interfaces to access data:

- VSAM via VSAM macros (running in the connector part.) to 'mapped' VSAM data
- POWER via POWER SAS
- LIBR via LIBRM(?)
- ICCF via DTSUTIL (read only)

„ **DB2-based connectors**

1 (or >1) Stored Procedure Server(s) act as connector server(s) using DRDA:

- The VSE DB2 database server knows, which Stored Procedure Server must be accessed to initiate a certain stored procedure.

The stored procedures run in a separate batch partition:

- For accesses to VSAM, the VSAM CLI can be used. (Shipped sample(s) is in 'C').
- Also, for accesses to VSAM under CICS TS, a stored procedure can be used as router to exploit the CICS TS EXCI interface.
- For DL/I accesses, calls are done in a CICS TS partition, via DL/I 1.11. (Shipped sample(s) is in COBOL and in Assembler).

Usage of TCP/IP for DRDA communication requires DB2 Server for VSE & VM 7.1, previewed 2000-04-11, announced 2000-09-05.

VSAM Connectors (VSE Details)

VSE Connectors and VSAM (Summary)

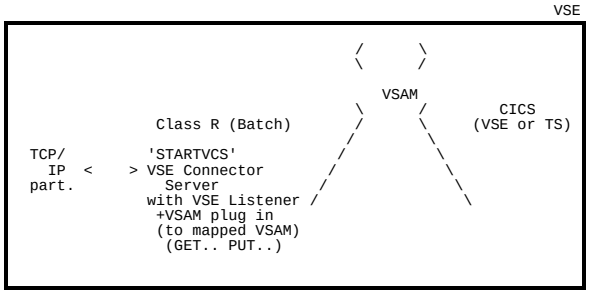
The following type of VSE connectors allow access to VSAM files:

IBM Connector Type	VSE partition of connector server	VSAM file 'owned' by partition doing the WRITE?
MQSeries	CICS(s)	YES
CICS	CICS TS(s)	YES
VSE Java Beans	VSE Listener (class R, default)	NO *1 *2
DB2-based	LE batch pgm(s) a) VSAM CLI I/F b) EXCI used	NO *1 YES

*1 In VSE/ESA 2.5 the VSAM file must be redefined as SHROPT(4), to allow updates from the web.
In VSE/ESA 2.6 SHROPT can stay as it is/was.
*2 Currently, no plug-in shipped with EXCI

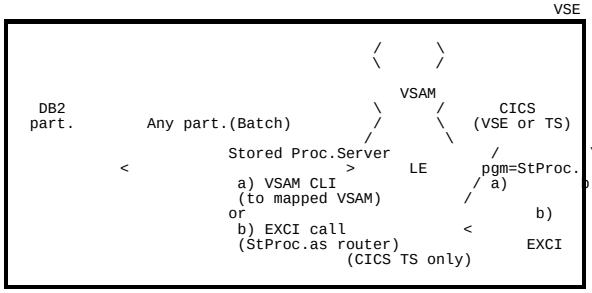
VSAM Connectors (VSE Details) ...

VSE Java Beans Connectors (VSAM)



- **SHROPT 4 required for 'e-WRITEs'**

DB2-based Connectors (VSAM)



- **SHROPT 4 required for 'e-WRITEs', except access via EXCI**

VSAM Share Options with Connectors

No considerations required for VSAM file READs from connectors.

VSAM SHROPT 4 Background Info

- **Using connectors to UPDATE a VSAM file already opened for output, needs SHROPT 4, IF WRITEs are done in an 'extra' partition**
- **SHROPT 4 has big overhead, and thus should be carefully evaluated**

VSAM SHROPT 4 Performance Implications

Bigger pathlength for processing of UPDATE requests
due to VSAM internal locking of control areas.

Internal CA-locks degrade performance
No concurrent updates possible. Locks usually in lock table only.

Each READ must be done from disk
(except GET SEQ Noupdate),
in order to get the latest copy of a record.

> No READ Ahead of CIs (in NSR Sequential mode)

Each WRITE must go to disk
VSAM always must assume updates in other partitions (sequence set, if changed, and data).

> No Deferred Writes possible (in LSR mode)

> LSR buffer lookasides only for higher level indexes, not for sequence set or data

Additional catalog I/Os
At each GET/PUT statistics are updated (+ HI-used RBA, if changed)

VSAM Share Options with Connectors ...

Performance Hints for 'e-WRITEs'

- **Try to use/customize connectors which cause the WRITEs in the file owning partition (mostly CICS), to avoid SHROPT 4.**

  - e.g. via a user written batch EXCI call to start a CICS TS transaction.

  BUT consider EXCI overhead and do not use it for individual READs
  EXCI is also going via DFHIRP, as MRO does.

If that is not possible, try to avoid SHROPT 4 by other means:

- **Use BIM-VSHARE in CICS as 'VSAM traffic cop'**

  VSAM batch I/Os are intercepted and sent via XPCC to a VSHARE component running under the file owning CICS.

  Refer to http://www.bimoyle.com

- **Use SYSB-II for VSE from H&W Computer Systems**
  in Boise, Idaho

  VSAM batch I/Os are intercepted and sent via XPCC to a CICS command level transaction in the file owning CICS.

  Refer to http://www.hwcs.com

Or, minimize/avoid the overhead of SHROPT4:

- **Use BSI OPTI-Cache as a VSAM 'Global Resource Pool'**
  Avoids physical I/O overhead for SHROPT 4, apart from providing 'DIM'.

  Refer to http://www.bsiopti.com

DB2 Stored Procedures

DB2 Stored Procedures (few general remarks)

„ **User-written SQL programs,**

- stored at the DB2 server
- invoked by client applications via the SQL CALL statement, even with dynamic parameters, including procedure name

„ **Can contain most statements that an application program usually contains**

Application logic and SQL statements, local to the data base

„ **Programming language(s) depend on the platform of the DB2 Server**

„ **Client programs can pass parameters to it and receive parameters back**

„ **Helps to reduce the network traffic**

„ **Allows to hide data base design from the end user**

„ **Must be found in the VSE LIBDEF chain of the partition a stored procedure runs in**

„ **Usually, 1 'never-ending' jobstep in a VSE batch partition is used**

DB2 Stored Procedures ...

Specific for the new VSE DB2-based connectors

„ **Only a fixed, not a dynamic amount of data can be passed back at once (VSAM, DL/I)**

This results either in piecemeal or in maximum area return of data.

Í **VSE Stored Procedure connectors (VSAM, DL/I) are better suited for a fixed set or for smaller outputs**

„ **Related Info**

- 'Getting Started with DB2 Stored Procedures: Give Them a Call through the Network' IBM Redbook, SG24-4693-01, 04/98, 492 pages
- 'DB2 Server for VSE Database Administration' SC09-2655, 12/98, 300 pages (Chapter 10)
- 'IBM DB2 Server for VSE & VM, Version 7.1' Announcement 2000-09-05
- 'What is new with DB2 Version 7?' VM/VSE Tech Conf La Hulpe, 2000-06-26 by Guy Prztula, IBM Belgium

Performance Aspects for Web Applications

Performance Aspects for Web Applications

Performance statements and predictions are hardly reliable for transactions via the Internet.
Only for intranets more reliable observations are possible, in general.

Ù **Java Applet Applications**

The following components contribute to overall response times:

1. **Transfer time of the Applet code to the Java capable browser**
2. **Time for initialization and checks of the applet code**
3. **Time for Byte-code translation (if JIT compiler used)**
4. **Execution of the individual Java Applet functions**

Ù **Java Servlet Applications**

TBD

Ù **IBM WebSphere**

Keeps multiple connections up and uses them dynamically in order to save connection setup effort.

Types of e-business CICS Applications

Types of e-business CICS Applications

	Screen Type	Presentation Logic	Screen Conversion to HTML	Use w/ 3270 or TN3270
CICS 3270 appl. (Any CICS)	Native 3270 or BMS maps	Integrated (e.g.COMM.) or Separate (needs COMMAREA)	Dynamically via 3270 Bridge or Generated HTML templates used	YES
CICS COMMAREA appl. *1 (Any CICS)	Series of records, w/o presentation specifics	Separate (outside appl.logic)	COMMAREA to HTML conversion (user provided)	NO
Web API CICS appl. *2 (CICS TS only)	Native HTML, using EXEC CICS DOCUMENT	Integrated (seldom) or Separate (usually)	n/a	NO

- COMMAREA (= communication area) is a CICS area that is used to pass data between tasks that communicate with a given terminal. Can also be used to pass data between programs within a task.
- Separate business logic easily allows both '3270-use' and 'native HTML use'

*1 A COMMAREA with an HTTP request is taken as input, and an HTTP response is built in that area. Often ECI is used by CICS client. COMMAREA based programs allow a free formatting of the browser screen, independent of any 3270 logic

*2 New WEB and DOCUMENT APIs used to process the inbound HTTP request and build the response. No COMMAREA needed

CICS Web Support (CWS)

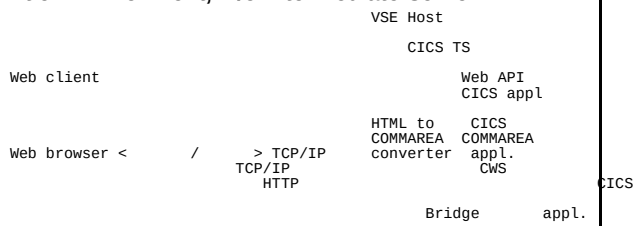
CICS Web Support (CWS)

Shipped with CICS TS 1.1.1 in VSE/ESA 2.5

- **Invoke a CICS TS application from a web browser e.g. without changing any 3270 application**

  CWS so far was known as CICS Web Interface. CWI was split into the listener support and the protocol support.
  Only in widest sense a connector.

- **2-tier Environment, w/o intermediate server**



  3270 Bridge dynamically converts native 3270 screens to HTML and uses generated HTML templates for BMS screens (intercepting CICS 3270 terminal requests)

  Web aware CICS appl. use EXEC CICS WEB and ... DOCUMENT (native HTML)
  No Java application involved

  2 tier recommended for intranets
  (no SSL yet available in TCP/IP for VSE/ESA)

- **3-tier Environment, with intermediate tier**

  Use of HTTP over Secure Sockets Layer (SSL) possible, with an intermediate tier as FireWall (Secure Proxy).

For more info refer e.g. to
- 'CICS3270 Bridge and Web Support' by Colin Boulain, WAVV 2000, 10/2000

CICS Transaction Gateway (CTG)

CICS Transaction Gateway (CTG)

CTG is a workstation application that can accept requests from web browsers and route them into CICS TS.

It is shipped as part of CICS TS

- **Access to existing CICS TS applications from a web browser without any change**

3 tier model:

Replaces the CICS Internet Gateway, and CICS Gateway for Java

CICS Univ. Client provides 'a CICS TS terminal' on the middle tier

HTML conversion is done fully automatic

CICS 3270 Bridge 'alone'

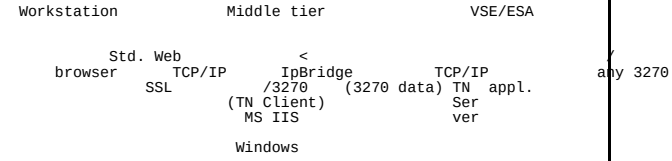
- **Invoke any 3270 txn from anywhere w/o a 3270 terminal**

Data21's Solutions for VSE e-business

IpBridge/3270 from Data21

- **Invoke a CICS 3270 application from a std web browser, without doing any back-end or front-end development**

Automatically converts 3270 data streams into HTML (and vice versa)
3 tier model:



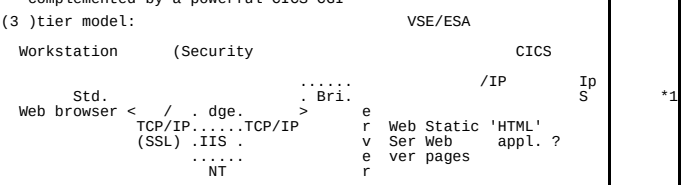
a) As a free Web to Host liteware client for secure access
b) As a nominally priced professional version

IpServer from Data21

- **Invoke a CICS application from a web browser**

  'complemented by a powerful CICS CGI'

2 /(3)tier model:



*1 SSI (Server Side Include) directives are supported, with 'empty' Web pages stored in a VSE library

For more info, refer to http://www.data21.com

IntelliWare's Solutions for e-business

Web/VSE-Host(R) from IntelliWare

- **Access VSE appplications from Anywhere(tm)**

  Any 3270 capable server, e.g. CICS

  no change required in applications

  instant access to 3270 applications

  GUI enhancements possible to any application via WebScreen(tm) facility

3270 screens are dynamically converted to HTML (*1)
They are intercepted (VTAM3270??)

For more info, refer to

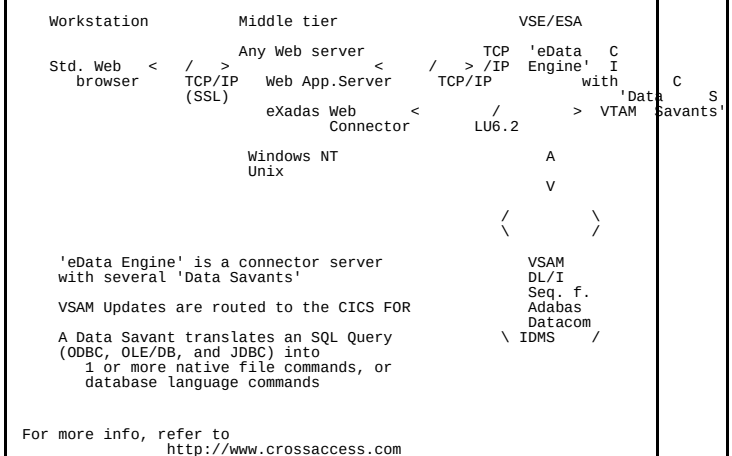
http://www.intelliware.com

Cross Access eXadas eData Solution

Cross Access eXadas eData Solution

„ **Access VSE data from anywhere using SQL without changing any application**

3 tier model:



For more info, refer to
http://www.crossaccess.com

Viaserv's ViaSQL Direct and Integrator

ViaSQL Direct

'Fast and easy ODBC and JDBC access to data and applications on VSE'

ViaSQL Integrator

'All the capabilities of ViaSQL Direct,
plus extended functionality for behind-the-scenes data'

For more info, refer to

http://www.viaserv.com

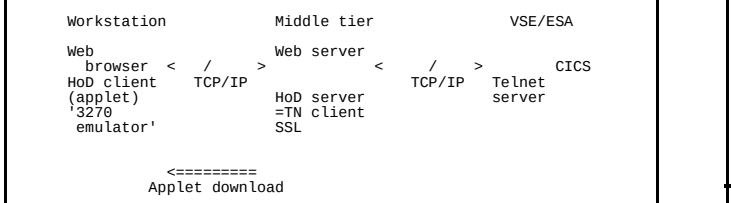
Host On Demand Improvements

Host On Demand (HoD)

„ **Easy access to an existing CICS application**

Applies to VSE/ESA V2 in general.

HoD access in a 3 tier model:



Seen from VSE, Host on Demand is just a TN3270 client.

Use of either
- black classic screen display

or
- configurable GUI screen

„ **More Info**

- http://www.ibm.com/software/network/hostondemand
- IBM SecureWay Host On Demand:
  Enterprise Commmunications in the Era of Network Computing
  SG24-2149-00.
- Host On Demand V4: Planning and Installation Guide
- Web-to-Host Integration Solutions, IBM Redbook, SG24-5237,
  205 pages, 09/98

DB2 Server for VSE

Note

DB2 for VM & VSE performance is discussed in detail in the official DB2 manuals. Refer e.g. to

- DB2 Server for VSE & VM, Performance Tuning Handbook
  GC09-2669, 12/98 (V6.1)
  GC09-2987, (V7.1)
- DB2 Server for VSE & VM, Installation
  GC09-2656, 12/98 (V6.1)
- DB2 Server for VSE, Program Directory GI10-4999 (V7.1)

Only SVA load aspects are discussed here, from a VSE/ESA system point-of-view.

DB2/VSE Phases and SVA (General)

General performance aspects of SVA loads (different for SVA-24 and SVA-31) are discussed in the VSE/ESA V2 Performance Considerations document,

Appendix A:
VSE Space Optimization under
'VLA General Usage Aspects'

„ **General**

- There is no functional requirement to load any DB2 phase into the SVA (VLA)
- Loading any phase into VLA may save
  - LOAD I/Os
  - VSIZE, if many partitions use a phase concurrently
- No SVA load books are provided/required with the DB2 product.

Cont'd

DB2 Server for VSE ...

DB2 Phases and SVA (Specific)

„ **SVA Eligible Phases (ARI*)**

SVA eligible DB2/VSE phases			
	#phases	total size	comment
SVA-24	20	about 3.5M	*1
SVA-31	4	about 1.8M	*2

- 5.1 figures. V7.1, V6.1 figures are similar.
*1 Biggest phase is main phase ARISQLDS (1.44M). Only loaded once in DB2 Server partition, do NOT load into SVA
*2 Phase ARIRXDS is 1.75M

„ **DB2 Recommendations for SVA Load**

Í **Do not load by default any DB2 phase into the SVA, except ARICDIRD**

ARICDIRD is only 1K to 4K for most customers. Placing it into SVA-24 will also eliminate potential functional problems of multiple ACDIRDs. Used by any partition running other DB2/VSE code.

If you detect via PGMLOAD stats or via traces frequent SVA-eligible ARI* loads ...

Í **Assess value of loading individual ARI* phases (customer specific)**

- Do a trade-off between frequency of load and space required BELOW the line
(Loading above the line is always OK as long as a module is used at all)

Í **Make PSIZE-31 big enough to avoid 'downloading' SVA-31 phases into SVA-24**

VSE Connectors and SVA

VSE Connectors and SVA

„ **Performance for VSE connectors improves, if some phases used are loaded into the SVA**

„ **VSE connector server phases**

Phase Name	Size	SVA Location	Purpose	Java Beans	St. Proc.
$IESVSCA	38K	-31	VSAM catalog access	X	X
$IESVSDF	81K	-31	VSAM catalog access	X	X
$IESVSLB	17K	-24	VSAM label access	X	X
$IESVSQL	158K	-31	SQL/IF f. st.proc.	-	X
$IESVSRM	48K	-31	VSAM catalog access	X	X
IESSASVA	.3K	-31	VSE Navigator 'DA'	X	-
CEEEV003	965K	-31	C Runtime (LE)	X	X

- Only IESSASVA must reside in SVA for FUNCTIONAL reasons
- SVA load list $SVACONN loads these phases by default (except CEEEV003)

„ **Nearly all of them may reside in SVA-31 and all are loaded by default**

„ **SVA load even is of benefit when only 1 connector server partition is used**

„ **Further SVA hint**

If EZA interfaces (EZASMI or EZASOKET) are used in several partitions ...

You may put phase EZASOH03 (24K) in SVA-31 (would need SVA=YES for CICS TS use)

Enterprise Storage Server (ESS)

PART G.

Enterprise Storage Server (ESS)

ESS Overview

Enterprise Storage Server (ESS) Overview

2105 I/O Subsystem

- Announcement 1999-07-27: Models E10 and E20 + SODs
- Announcement 2000-03-28: Models F10 and F20 + increased performance and scalability
- Announcement 2001-02-20: Models F10 and F20 with up to 32G cache

Ù **420 GB to 11.2 TB cost-effective high-performance storage across platforms**

Ù **Attaches to hosts with UNIX, OS/400, Win NT etc. and to S/390 operating systems**

Attachment to S/390 is via ESCON (else via SCSI).

Fibre Channel ports incl. FICON (up to 16) available as upgrade option

Ù **Standard caching functions and much more**

- Intelligent cache management (record/track caching)
- Sequential Detect
- Copy Services
- PPRC etc
- ...

Ù **Enhanced availability features**

See separate chart

Ù **S/390 performance accelerators**

See separate chart

Enterprise Storage Server (ESS) Support

ESS Support (Overview)

„ **VSE/ESA 2.1.0 and up in 'transparency mode'.**

No active exploitation of ESS Performance Accelerators

- Parallel Access Volumes (PAV)
- Priority I/O Queuing

Benefits from Multiple Allegiance (MA) expected.

For detailed info, refer to the following charts on ESS

„ **ESS FlashCopy allows instantaneous copying of S/390 volumes**

ESS feature code 183x.

Invocation via VSE/ESA 2.5

For detailed info, refer to the charts on ESS FlashCopy

ESS Enhanced Avail. Features

Enhanced Availability Features

No single point-of-failure.

RAID

RAID-5, striping parity across all DDMs in an array:
6+P+S or 7+P.

Non-RAID disk groups

Additional possibility: Just a Bunch of Disks (JBOD).
For UNIX freaks with 'volatile' data.

Predictive failure analysis

Failure analysis techniques that can predict errors and notifies before they occur

Failover protection

In case a cluster fails, all ports are transferred temporarily to the remaining cluster, thus allowing continued access to data.
It also allows e.g. u-code updates of a cluster concurrently to production with the other cluster.

Sparing

2 spares per SSA loop.

Sparing is automatically called at DDM failure.
Data is reconstructed while access continues.

Concurrent maintenace

A failed component can be replaced w/o shutdown.

ESS S/390 Performance Accelerators

ESS S/390 Performance Accelerators

„ **ESS is supported by VSE/ESA 2.1.0 and up in 'transparency mode'.**

No active exploitation by VSE/ESA:

Parallel Access Volumes (PAV)

ESS feature code 180x.

Application transparent subdivision of 1 logical volume via alias addresses (UCBs) into multiple 'domains'.
READs are simultaneous, also WRITEs to different domains.
WRITEs to the same domain are queued in the ESS.

Replaces IOSQ time in the operating system with (smaller) queuing in the I/O subsystem (ESS).

-> More intelligence in the ESS

Benefits at higher logical volume utilization, usually from more than 1 VSE partition.

Of lower need for VSE due to lower usage of 3390-9s (using more smaller devices instead).
Nevertheless, skewed device utilizations may occur.

Multiple Allegiance (MA)

I/O Subsystem (ESS) accepts requests for the same (shared) volume from another host, thus reducing or eliminating PEND time.

This is a H/W function and transparent to S/W.
VSE should benefit from MA (in case of DASD Sharing).

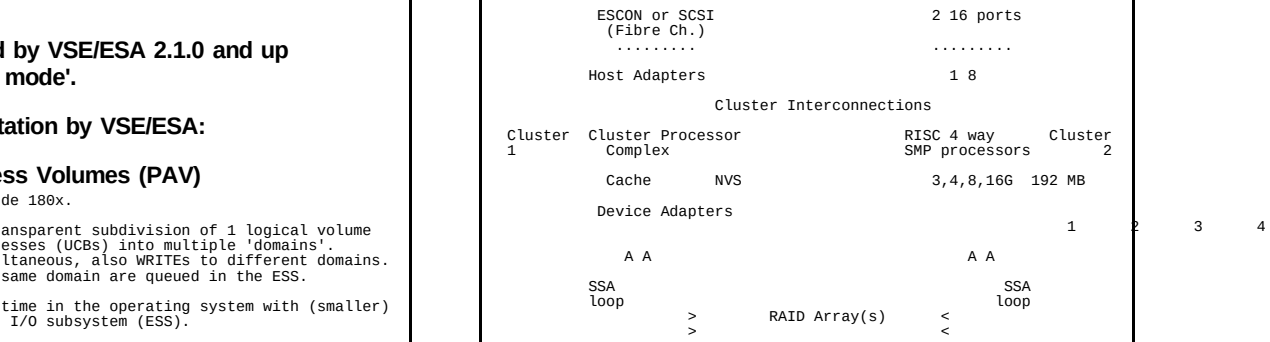
Priority I/O Queuing

Allows a more granular setting of I/O priorities than with PRTYIO: even on channel program base.

Of special benefit together with PAV in case of volume or channel contention

ESS Configuration

ESS Configuration



2 clusters, each with 192 MB mirrored NVS and 3/(4,8,16) GB cache (E/F-models)

(7-day battery back-up).
Small part of cache used as real storage for control pgm etc.

In case of a failing cluster, the other cluster can access data in the failed cluster:
either cluster can handle requests from any host adapter.

For failure protection, any WRITEs are done into both NVSs.

2 to 16 host adapters (HAs) in total

1 to 4 HAs x 4 bays x 1 or 2 ports/HA = 4 to 32 ESCON or SCSI ports

4 Device Adapter (DA) pairs, 2 loops/DA

Up to 8 full duplex loops, with up to 384 DDMs

Cache and NVS are transparent to all platforms, except S/390

ESS Configuration ...

Ù **32 Standard Physical Configurations**

Standard Configuration	DDM type	Capacity (GB)	#8-packs	2105 Models
'Ultra-high performance'	9.1G	420 GB	8	X10, X20
		840 GB	16	X20
'High performance'	18.2G	420 GB	4	X10, X20
		630 GB	6	X10, X20
		840 GB	8	X10, X20
		1050 GB	10	X20
		1260 GB	12	X20
		1470 GB	14	X20
		1680 GB	16	X20
		2170 GB	20	X20
		2660 GB	24	X20
		3150 GB	28	X20
		3640 GB	32	X20
		4130 GB	36	X20
		4620 GB	40	X20
		5110 GB	44	X20
		5600 GB	48	X20
'Capacity'	36.4G	840 GB	4	X10, X20
		1260 GB	6	X10, X20
		1680 GB	8	X10, X20
		2100 GB	10	X20
		2520 GB	12	X20
		2940 GB	14	X20
		3360 GB	16	X20
		4340 GB	20	X20
		5320 GB	24	X20
		6300 GB	28	X20
		7280 GB	32	X20
		8260 GB	36	X20
		9240 GB	40	X20
		10220 GB	44	X20
		11200 GB	48	X20

- DDM = Disk Drive Module
- X10 = E10 and F10, X20 = E20 and F20

- 16 to 384 DDMs in increments of 2 or 4 8-packs
- 36.4G DDMs are 7200 RPM, others 10000 RPM
- X10: 16-64 DDMs, X20: 32-128 DDMs
- 8 8-packs per cage. 1 cage for X10, 2 cages for X20
- No mix of DDM types in standard configurations

ESS Configuration ...

Ù **Logical Configurations**

1 to 16 CU images
(3990-3/6, or native 2105 E10/E20 for OS/390)

Up to 256 (logical) devices (= unit addresses) per CU image, resulting in 4096 S/390 volumes per ESS.

Up to 128 logical paths per CU image

3390-2,-3,-9 + 3380 Track Format Mode
Also 'Custom Volumes' with fewer cylinders possible

Ù **Host Adapters**

- Up to 32 concurrent host transfers

- 18 MB/sec per ESCON channel

Ù **Device Adapters**

- Up to 64 concurrent transfers on internal disk paths

- 2 independent SSA loops per device adapter

Maximum Sequential Bandwidth		
	'Shark 1.0'	'Shark 1.1'
READ	160 MB/sec	320 MB/sec
WRITE	140 MB/sec	200 MB/sec
- Approximate figures, just as an idea - Data rate to an 8-pack is 40 MB/sec		

Ù **Actual Performance Results**

REALLY REALLY good results published. See 'References', especially the ESS Performance White Paper.

General ESS Performance Hints

General ESS specific Performance Hints

General (but non-ESS-specific) hints are e.g. given in

'IBM VSE/ESA I/O Subsystem Performance Considerations'

„ **DDM selection**

Apart from standard configuration and capacity needs,
if you can select between DDMs of different size ...

(DDM selection is more important if workload is cache hostile)

Í **9G and 18G DDMs are faster for single sequential streams**
10000 RPM

Í **9G DDMs usually perform better than half the number of 18G DDMs**
E.g., prefer 16 9G DDMs vs 8 18G DDMs,
if 9G DDMs can satisfy your future capacity needs.

Hint:
Performance with the 36G DDMs is really fine, also.

„ **More Hints**

Refer e.g. to

- IBM ESS Performance Monitoring and Tuning Guide, SG24-5656 (IBM Redbook)

ESS Statement of General Direction

ESS Statement of General Direction

Included in the July 1999 announcement

Ù **Continued performance enhancements through**

more efficient engines

bigger cache sizes

Both has been realized with the 2105 Models F10 and F20 announcement 2000-03-28:

Additional PCI buses, faster IBM RISC microprocessors plus the bigger cache in the F10 and F20 model allow greater throughput

- up to 100% for sequential loads
- up to 25% for random loads

Ù **Higher capacity**

ESS Copy Services and FlashCopy

ESS Copy Services and FlashCopy

Ù **ESS Specialist Copy Services include**

Peer-to-Peer Remote Copy (PPRC)

A synchronous copy solution, where WRITE operations are ensured on both copies (primary and secondary).
ESS feature code 182x.
Needs S/390 software support and invocation.
Supported by VSE/ESA V2.

Extended Remote Copy (XRC)

Not supported by VSE/ESA

FlashCopy

Transparent for S/390 operating systems, if done via ESS Copy Services

Like SnapShot Copy, a point-in-time copy function

Refer to separate charts

Ù **FlashCopy can be invoked**

- by the ESS Specialist web interface

Transparent to the operating system.
Refer to separate chart

- via the operating system

VSE/ESA 2.5 supports FlashCopy calls

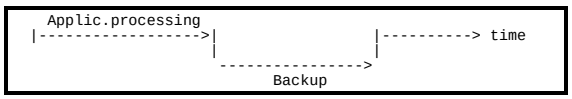
- for volume duplication

e.g. for later backups of VSAM files

- for copying cylinder ranges or file extents

Backup Processes

Traditional Backup Process



„ **Quiesce/stop production during backup**

Provides a logically consistent copy, BUT makes data unavailable

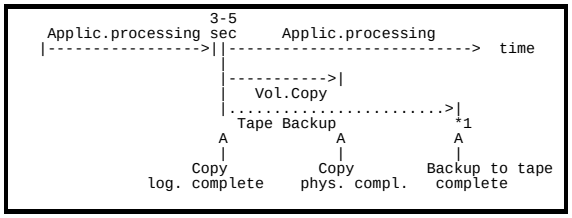
„ **(Do a Backup concurrently)**

Results in a 'fuzzy' copy, only useable, if at all, with log data sets. Also, a data base usually is made up of several logically related data sets.

Possible Backup Process as of today

„ **Use point-in-time copy functions**

Applies to RVA SnapShot Copy and ESS FlashCopy.
Safest action is to bring CICS down and immediately up again.
Then VSAM files really were closed. DB2 allows suspend/resume.



*1 'Backup to tape complete' is independent of 'Copy phys. complete'

- Volume Copy is logically completed after 3-5 sec
- Backup to tape can start immediately

Logically done from target volume (different VOLID)

FlashCopy Implementation

FlashCopy Implementation

Ù **Requirements**

Installation of the ESS FlashCopy feature

ESS needs physical backend storage for FlashCopy

Requires a target and source volume of same size and within the same logical subsystem (LSS)

No relocation of tracks can be done:
CCHH location must be same on target as on the source volume

Ù **Relationships**

While copying, a relationship exists between source and target volume:
The relationship ends when the last track has been physically copied to the target volume.

1 logical volume can have only 1 relationship at 1 point-in-time.

Source and target volume must reside in the same LSS (same SSID), and, naturally, be of same geometry with at least same capacity

Ù **Copying Process**

Physical copying of ALL tracks is done via a background task, with negligible impact on production (if at all)

ESS Specialist (only) allows to specify a 'NOCOPY' option, to disable the background task.

By this, unnnecessary copying (of unchanged source tracks) can be avoided.
But still the full logical space on target volume must be reserved.

Reasonable only for short living copies
(e.g. while backup to tape is done)

FlashCopy Invocation

FlashCopy Invocation via ESS Specialist

1. Select ESS Specialist (-> ESS Main Panel)
2. Select Copy Services (-> Copy Services Menu)
3. From here select:

- Volume

(Display volumes and define them as source or targets and initiate the copy via 'Establish FlashCopy Pair')

Also possible selections:

- Storage Servers
- Paths
- Tasks
- Configuration
- Exit Copy Services

FlashCopy Invocation in VSE/ESA 2.5

Support similar to IXFP/SnapShot for VSE/ESA, via AR commands:

AR commands as close as possible (if not identical) to SNAP, as documented for IXFP/SnapShot for VSE/ESA.
REPORT and DDSR not required/applicable.

More info is provided in the VSE/ESA 2.5 manuals.

Scope of Function

FlashCopy via ESS Specialist
- cannot copy individual track ranges (only full volumes)
- cannot be used to copy a VM partial pack minidisk

Both can only be done via VSE/ESA 2.5 invoked FlashCopy

RVA SnapShot and ESS FlashCopy

RVA SnapShot and ESS FlashCopy

	VDA Architec-ture	Instant 'Copy complete'	Physical copies of tracks done ...
RVA SnapShot	yes	yes	Only when track updated (source or target)
ESS FlashCopy	no	yes	Later, as backgound task. Immediately, when updated
*1 Backup of Target volume can be started after few seconds			

SnapShot and FlashCopy Scenario

SnapShot and FlashCopy Scenario

```
      I
      I   Online Production halted

  t   C
    i   C   Logical Copy:     3   5 sec   (RVA, ESS)
    m   C                 V
    e
      I   Online Production continues.
  V   I   Target volume available,
      I   backup to tape can start:
      I
      I    READs from 'source'            see table
      I    WRITEs  to 'source'            below
      I    READs from 'target'
      I   (WRITEs  to 'target')
```

Activity	SnapShot Copy	FlashCopy
READs from source	BAU	BAU
WRITEs to source	Do a copy to a new source location, keep old as target	Do a phys. copy to target (except already done), before source track is updated
READs from target	Done from source, if not updated meanwhile	Done from target, if already copied, else from source
WRITEs to target	BAU	Do phys. copy from source to target first if not already done
Impact on produc-tion performance - copying	none	Phys. copying of ALL tracks creates ESS background activity
- tape backup	very minor	1st production WRITE to source track slow, if not yet copied
- BAU = business as usual - All RVA updates are done in 'tracks'		

Addt'l Aspects to SNAP-Copies

Addt'l Aspects to SNAP-Copies

RVA SnapShot and ESS FlashCopy

- **Can be used via the VSE AR command 'SNAP', to copy/duplicate**

  Logical volumes
  Specified via CUU or VOLID
  Cylinder ranges
  Files (non-VSAM)
  Via DSN on CCHH (track) boundaries, e.g. BAM file or VSAM space

- **Some conditions for snapping data**

	Source Vol.	Target Volume			
	Status	VOLID	Must exist	Snap Device Status Before	After
ESS Specialist Vol.Copy	?	Can be changed	*1	*1	Up
VSE AR Snap Vol.Copy	Up or Down	*2 Can be changed	No,RVA Yes,ESS	Down	Down
VSE AR Snap Part.Vol.Copy	Up or Down	Not touched	Yes	Down *3	Same?
IDCAMS Snap Vol.Copy	Up or Down	Must be changed	Yes	Up or Down	Up

*1 Volume must exist at least in IOCDS, no VOLID required
*2 If specified via VOLID, no duplicate VOLID is allowed, which either is also UP or also DOWN.
*3 except copy within same device

Addt'l Aspects to SNAP-Copies ...

Addt'l Aspects to SNAP-Copies (cont'd)

- **High care required if double VOLIDs are DVCUP**

  Search sequence not easy to describe and may change.
  Often, search is in ADD sequence of ASI proc,
  BUT manually ADDed volumes are searched first

  So select target VOLID differently for unique VOLIDs.

- **Further Requirements**

  - Same type/size of volume for volume SNAPs
  - Same track/cylinder type for other SNAPs
  - Both volumes must be
    - in the same RVA subsystem
    - in the same ESS logical subsystem (LSS)

- **SNAP related functions need $IJBIXFP in the SVA-24 (functional reasons)**

  a) $IJBIXFP for ESS 36K (part of base code)
  b) $IJBIXFP for ESS,RVA 50K (part of opt. prod.)

  In VSE/ESA 2.5, the 36K phase is loaded by default.

  Can be removed from the SET SDL list, if VSE SNAP not used and this space required

ESS FlashCopy Performance

ESS FlashCopy Performance

The following is my summary and interpretation of the document cited below.

Ù **How fast FlashCopy can duplicate data ?**

Number of FlashCopy Pairs		8	16	32	256
Establish Times	(sec) *1	10	17	19	96
Volume Reserve Times (sec)	First	9	11	11	11
	Last	11	14	15	65

*1 Times from DFSMSdss console messages,
TSO initiation even faster
- Concurrent production does not impact these times

Í **Very fast logical completion of copying volumes**

Number of FlashCopy Pairs	8	16	32	256
#tracks copied/sec				
a) w/o production	1600	2700	3000	3000
b) w/ production	1050	1650	2000	2000

- Total background copy time (w/o production)
1-8 volumes: about 4 min
32 volumes: about 9 min
- All values are approximate values

Í **Fast physical completion of copying volumes**

Impact of production is background copy is uncritical and intended

ESS FlashCopy Performance ...

ESS FlashCopy Performance (cont'd)

Ù **What impact has FlashCopy on production ?**

Number of FlashCopy Pairs	None	8	16	256
Prod. throughput IO/sec				
COPY option	2700	2600	2550	2500
NOCOPY option	"	2650	2600	2600
Prod. IO resp. time msec				
COPY option	5	5	6	8
NOCOPY option	"	6	7	13

- In case of NOCOPY, only changed tracks are copied to the target volumes (= destaged from NVS)

Í **Impact on production throughput was small**

-3%/-7% (NOCOPY/COPY option) for 256 volumes

Ù **Measurement Environment**

ESS Model F20, with 18G DDMs, 8 LSSes, with 16 ranks,
16 ESCON channels to each of 9672-RX3 and 9672-R85.
256 source volumes copied to 256 target volumes.
3390-3 volumes with 3339 cyls and 50085 tracks

Ù **More Information**

Refer to the following document for more info:

'ESS FlashCopy Performance White Paper'
2000-08-04, IBM SSD, 16 pages

A document also for IBM Business Partners and Customers.

ESS References

ESS References

For more info and published actual performance results, refer to

- ESS announcement letter, dated 99-07-27.
Available e.g. via http://www.ibmlink.com

- ESS home page http://www.ibm.com/storage/ess

- ESS Customer Documents CD-ROM SK2T-8770

- IBM ESS Performance Monitoring and Tuning Guide, SG24-5656, 01/2000, 197 pages, by Alison Pate et al

  As IBM Redbook draft ('redpiece') available via
  http://www.ibm.com/redbooks/

- Implementing the ESS in Your Environments, SG24-5420, 02/2000, 269 pages, by Mark Blunden et al (IBM Redbook)

- Implementing ESS Copy Services on System 390, SG24-5680, 06/2000, 280 pages, by Mark Blunden et al (IBM Redbook)

- IBM ESS Introduction and Planning Guide, GC26-7294-03, 03/2000, 124 pages

- IBM ESS User's Guide, SC26-7295, 09/99, 67 pages

- IBM ESS Performance White Paper, Version 1.0, 09/99, 69 pages, by John Ponder et al.
Via ESS home page

- IBM Advanced Copy Services, SC35-0355, 03/2000, 250 pages (includes DFSMS FlashCopy, SnapShot Copy, Concurrent Copy, PPRC)

- Using the IBM ESS with VM/ESA and VSE/ESA, by Bill Worthington, VM/VSE Tech Conf 06/2000, Orlando

- ESS FlashCopy Support z/VM and VSE/ESA, by Bill Worthington Tech Conf Jacksonville, 05/2001

App. A: VSE JA and Reporting Enh.

PART H.

App. A: VSE JA and Reporting Enh.

- VSE/ESA JA in General

- VSE JA Reporting Enhancements
(VSE/ESA 2.4 and 2.5)

Chart 9 and up

VSE Job Accounting (JA)

VSE Job Accounting (JA)

Ù Total CPU time = CPU TIME + OVHD TIME

Any time is related to an individual partition (and job step)

Ù OVHD Time is that time, VSE JA

Cannot attribute to a certain partition

Refer to a following foil, for more OVHD details

Does not attribute to a certain partition.

A more detailed subdivision would impact performance inadequately.

Is distributed by JA across all active partitions

Assuming that every partition has the same ratio of OVHD to CPU TIME

Ù Overhead of VSE JA

Additional CPU time very small

JA=YES vs JA=NO
(JA data are collected even for IPL SYS JA=NO)

Only at job step end

- To go into $JOBACCT routine
- To do whatever user coded $JOBACCT routine does (usually few lines on SYSLST)
- To reset job step counters to 0

Check ICCF library 59 (SKJOBACC) for a VSE/ESA JA skeleton routine.

VSE Job Accounting (JA) ...

VSE JA (cont'ed)

Ù Sample JA Info on SYSLST

```
JOBNAME   = PRINTLOG  USER INFO=       PR    EXEC NAME =  PRINTLOG
DATE      = 11/05/99  PART ID  =    BG
START     = 10:56:23  STOP     = 10:56:28 DURATION     = 5.560 SEC

CPU       =        0.060 SEC          PAGEIN SINCE IPL  =         0
OVERHEAD  =        0.017 SEC          PAGEOUT SINCE IPL =         0
TOTAL CPU =        0.077 SEC

UNIT =       E15   UNIT =     FEC    UNIT =   01F    UNIT =    E16
SIO  =        26   SIO  =       5    SIO  =     5    SIO  =    105
UNIT =       FEE
SIO  =      4083
```

CPU-time related remarks

This routine also displays the CPU-time consumed in the POWER partition (caused by ALL partitions active) during the run-time (elased time) of a considered VSE job step (except during the first job step of a job).

If you would require the POWER CPU-time for a VSE job with 1 job-step, you
- must only have this 1 partition consume CPU-time
- must add a dummy job-step in front

SIO related remarks

SIO figures shown here are SVC0 counts.

So, for a spooled device, it usually is the number of print lines (SYSLST).
For non-spooled devices (logical volumes) it coincides with the number of SSCHs, which may either be

- physically executed

or

- directed to a VSE Virtual Disk

or

- intercepted e.g. by BSI OPTI-CACHE.

or

- intercepted by VM and simulated for VSE

VSE JA Overhead Times

VSE JA Overhead Times

The following examples show activities counted on OVHD, and distributed with equal relative share across all partitions.

Ù Whenever OVHD share is high ... the cause may lie in one of these areas

- Any activity caused by an Attention Routine command
- All paging/page manager activity
- All SSCH instructions
- All intercepts (incl. vendors) of parts counted on OVHD
- All activities done in channel appendages
- Any error recovery for I/O devices
- Any I/O retries/reschedulings
- Part of interrupt handling (I/O, External)
- FETCH directory task (part of PGMLOAD)
- Other system tasks

Ù Supervisor State and JA OVHD

- As is well known from S/390 architecture, certain privileged instructions can only be executed in 'SUPERVISOR STATE'
- Do NOT assume that JA OVHD and SUPERVISOR STATE are the same. There, naturally, is some correlation, but many parts of the supervisor (which completely runs in supervisor state) are attributed to individual partitions as CPU TIME
- Supervisor (and PP) state CPU-times per physical processor can be obtained from the 'Analyse System Activity' screen e.g. of the 9672 Service Processor console

VSE Display Activity (DA)

VSE Display Activity (DA)

VSE Display Activity here refers to the following screens of the VSE System Status Displays:

- **Display System Activity (DSA)**
  (IESADMDA panel, IUI fastpath 361)
- **Display Channel and Device Activity (DCDA)**
  (IESAD... panel, IUI fastpath 362)

Performance related data are displayed at fixed intervals (15 sec default) or at ENTER (interval value =0)

More info in VSE/ESA Operation manual

Ù Display Activity results are based on VSE Job Accounting (JA)

Thus, SYS JA=YES is required (as is normally specified).

Ù Can be only invoked under IUI

ICCF is not required, just CICS and the IUI

Ù Purposes

„ **Get general info on current system status**

„ **Serve as an initial problem or performance diagnosis tool**

- to check CPU utilizations and overall paging
- to follow up progress of test runs
- to observe CICS loads in order to prevent SOS condition during future CICS sessions (CICS section, 'Pages Avail')
- to observe/detect I/O bottlenecks

VSE Display Activity (DA) ...

VSE Display Activity (cont'd)

Ù **1 line in DSA per active job step**

```
 ID  JOB NAME  PHASE NAME  ELAPSED  CPU TIME  OVERHEAD  %CPU  I/O

 F1  POWSTART  IPWPOWER   18:38:30     8.11     2.35          3,436
 F3  VTAMSTRT  ISTINCVT   18:38:09    23.89    31.54          5,466
 F2  CICSICCF  DFHSIP     18:38:06   120.34    34.39    67%  44,682
     ...       ...         ...          ...      ...          ...
```

For details and explanations, refer to 'VSE/ESA Operation'.

Ù **CPU utilization on the DSA screen**
is slightly smaller than actual value (roundings)

Ù **CPU time/utilization figures are reliable, if**

All CPU and OVHD TIME deltas are calculated and added across all active partitions

Measured partitions stay in same job step

í **CPU (and I/O) data only apply to still active job steps**
Active both at start and end of the measurement interval

Ù **CPU-time values approximate virtual CPU time**

- VTIME under VM
- in an LPAR

i.e. the overhead by VM or PR/SM is not included.

VSE Display Activity (DA) ...

VSE Display Activity (cont'd)

Ù **I/O figures are reliable**

on DSA screen
(DSA IO/sec and IO per active job step do not contain POWER spooled records)

on DCDA screens

IO counts are displayed for whole channel or device(s) on channel during the last interval

For spooled devices, e.g. FEE, the figures show the number of spooled records per active job step.

Ù **VSE DA overhead is low**

Base JA overhead
Display overhead at screen refresh
(automatically at intervals, or upon Enter)

í **VSE DA can very reliably determine VSE/ESA native CPU-times**
(if properly used, as indicated above)

í **VSE DA cannot be used to determine LPAR (or VM) overhead**

VSE DA (and JA) reports only the 'virtual' time, not the total processor time

VSE JA and DA Usage

VSE JA and DA Usage

Both tools have a lot in common, and complement each other.

DA is generally for system snapshots and for 'long running job steps', while JA in its native form is for whole job steps.

Ù **Use DA for**

System snapshots
System utilizations and I/O rates (total, per partition) at any instant

Single transaction measurements
(CICS, VTAM, SQL server partition)

Via 'interval end - interval begin value'

(Most accurate results, if only measured transaction is active)

POWER spooling activity
Via 'interval end - interval begin value'

Ù **Use VSE JA for batch jobs and job steps**

(Most accurate results, if only measured partition is active, refer to OVHD explanations)

Add VSE/POWER from DA in addition, if needed and appropriate

VSE JA Data Collection via DMF

VSE JA Data Collection via DMF

Applies to VSE/ESA 2.4 and 2.5

Ù **Summary**

„ **VSE JA data (per job-step) can be collected and**

directed to SYSLST (as in all previous releases)

or, alternatively, NEW!

collected via DMF ('DMF-JA') for later reporting and evaluation
described in the following

Ù **'DMF-JA' Data Collection**

„ **JA data are collected in a DMF data set**

It is the same data set where the DMF Data Handler also writes CICS TS Monitoring and Statistics records.

Separate types of SMF records (type=200) and a separate SYSID (VSE1) are used, so no interference with any CICS TS records occurs.

CICS TS need not be started, just DMF (DFHDFSIP) must have been started before AND must still be active.

As soon as the first record is being accepted in the DMF data space, the following message appears at the console:

*** DMF REC STARTED ***

In case DMF data space can not be accessed:

*** DMF REC STOPPED ***

Also, as for any type of JA information, JA=YES is required.

DMF-JA Data Reporting

DMF-JA Data Reporting

The reporting program allows
(via the DMF data set)

for specified selection criteria
(date range, partition, device ranges ...)

to get chronological JA data for VSE jobs

```
  JOB REPORT: ACCOUNTING INFO FOR INDIVIDUAL VSE JOBS

JOBNAME ST ID DD.MM.YYYY  START.. DURAT.  CPU + OVRHD = TCPU  IO
 ------------------------------------------------------------
HCBILL   2 C1 08.02.2000 14:55:23 0:01:02 3.630 1.920 5.550  3340
H1PAY    2 F5 08.02.2000 15:01:40 0:00:17 2.170 0.730 2.900  2936
 ...

 TOTAL PROCESSED JOBS =   294
```

Í **A history of JA data per VSE job**
(independent of the POWER job name),
optionally sorted by job name

to get VSE job summary data, sorted by job name

focusing on '#I/Os to selected range / total CPU-sec per job'

```
  SUMMARY REPORT: IO/CPUSEC FOR INDIVIDUAL VSE JOB NAMES

JOBNAME = HCBILL ( 2)  AVG IO/CPSEC = 605  COUNT=   24  TOTAL=  28
JOBNAME = H1PAY  ( 2)  AVG IO/CPSEC = 837  COUNT=   17  TOTAL=  19
 ...

 TOTAL SELECTED JOBS=   294    TOTAL RECORDED JOBS =   312
```

For relevance of SUMMARY REPORT, refer to next chart.

DMF-JA Data Reporting ...

Relevance of Summary Report

Ù **Get more appropriate performance data with your own workload/jobs, on a safe(r) statistical basis**

- after change of VSE release
- when migrating to another processor

via IO/CPUSEC values

Physical IO/CPUSEC is often a very robust performance value, which only changes

e.g. - with VSE releases
- when migrating to another processor
- when changing setup for KB/IO (I/O blocking)
- when changing the degree of DIM exploitation (I/O buffering).

Most important, it does NOT depend on the amount of work, which may change daily.

The report allows to exclude e.g. logical (spooled) I/Os, such that SIO values really only apply to SSCHs to disks.

Also you can exclude extreme IO/CPUSEC values from the statistics.

Ù **Examples**

A typical and important VSE job 'BILLING' (which has not been changed regarding I/O blocking or buffering) results in the following values:

```
 JOBNAME = BILLING    AVG IO/CPSEC = 605   before migration
 JOBNAME = BILLING    AVG IO/CPSEC = 594   after  migration
```

Assuming that still between 2 subsequent I/Os the same total (user) work is done, this means that in the average for this job 2% more CPU-time was required after the migration.

When upgrading a processor, you may increase AVG IO/CPSEC by a factor of say 1.5 and more, in average identical to the ITR ratio of the new vs the 'old' processor.

DMF-JA Data Reporting ...

DMF-JA Data Reporting

Ù **Report Generation (Internal steps)**

The JA records stay in the DMF file(s) until re formatted.

```
     DMF File(s)

                          DMFDFOU    Offload of all JA (job         step) info
                             from DMF file(s).
         V                   Mult. jobs req'd for mult. DMF files.

     File ESDS1

                          DFSORT (1) Sort all job         step records by
                                              partition         ID
         V                              job         name
                                  start date and time
     File ESDS2

                          DFSORT (2) Create and Sort job records by
                               start time (date, time) (default)
                               job name
                             SORT E15 exits used
                                to discard non eligible job steps
         V                      to determine job steps/job
                                to add I/O counts within job steps
     File ESDS3                 (for eligible devices)
                                to get job data from job step data
     JOB REPORT
     .... ...
     ....
          /
                          DFSORT (3) Create and Sort job         name records by
         V                                job         name
                               number of job steps
     File ESDS4             SORT E15 exits used
                                to calculate AVG IO/CPUSEC for each
    SUMMARY RPRT                job name occurrence with same
     .... ...                   number of job steps
     ....
          /

    Skeleton SKJOBDMF assumes DFSORT/VSE to be used for sorting.
    Other SORT products may need adaption.
    All ESDSx files are created anew, before used
    All jobs are VSE jobs, NOT POWER jobs
```


DMF-JA Data Reporting Input

DMF-JA Data Reporting Program Input

Ù **Input parameters must be specified via 'PARM='**

In any sequence and number

If an input parameter is not specified, the default is taken

Up to 3 PARM statements, each up to 100 characters

Values of a single parameter (which must be ended by ',END') can be in any sequence

Short example:

```
EXEC ACCTREP,SIZE=384K,PARM='STA=F2,BG,END,RANGE=17022000-13032000,
          ECHN=F,END,RECNUM=5,SKIP=0'
```

Ù **Input Parameter Description**

Only list/use specific static partitions:

"STA=PARTID1,PARTID2,...PARTID12,END"
E.G.: STA=F1,BG,F3,END
(Max : 12 partitions)
(Default: all static partitions)

Only list/use specific dynamic partitions or classes:

"DYN=CLASS1,CLASS2,SPECCLASS3,SPECCLASS4,END"
E.G.: DYN=C,B,G,Q1,S3,END
(Max : 25 dyn inputs)
(Default: all dynamic partitions)

Only list/use jobs started in a specific date range:

"RANGE=DDMMYYYY-DDMMYYYY"
E.G.: RANGE=14051999-16071999
(1 day must also be specified as range)
(Default: all dates)

DMF-JA Data Reporting Input ...

Ù Input Parameter Description (cont'd)

```
How to sort the jobs in the job report:

  "SORT=XXXX"
     (JNAM= Sorted by Jobname, JDAT= Sorted by Jobdate)
     E.G.:  SORT=JNAM
                        (Default: sorted by date)


How many incidents of a job name are required to be listed
in the summary statistics:

  "RECNUM=X"
     E.G.:  RECNUM=5
                        (Default:   10)


How many jobs at the lower and upper end of IO/CPUSEC
should be disregarded in the summary statistics:

  "SKIP=X"
     (X= 0 TO 49) X= value in %
     E.G.:  SKIP=15 (%)
                        (Default:  10)


What channels should not be reflected in the SIO counts
from JA:

  "ECHN=A,B-C,D,END"
     (Channel values, between 0-F,
      ranges are also allowed, E.G. 3-4)
                        (Default:  no exclusion)


What devices should not be reflected in the SIO counts
from JA:

  "EDEV=CAA,CBB-CCC,CDD,END"
     (Device addresses,
      ranges are also allowed, E.G. 220-250)
                         (Default:  no exclusion)
```


Shipped ICCF Members for VSE JA

Ù Shipped ICCF Members for VSE JA

VSE/ICCF members in Lib 59			
	VSE/ESA 2.2/2.3	VSE/ESA 2.4/2.5 *1	VSE phase name
JA to SYSLST - collection and display	SKJOBACC	SKJOBACC	$JOBACCT
JA to DMF	-		
- collection phase	-	SKJADACC	$JOBACCT
- record offload	-	SKJADOFF	DFHDFOU
- reporting phase	-	SKJOBDMF	ACCTREP
- reporting job	-	SKJADRPT	-"-

\- Before overwriting the shipped $JOBACCT dummy phase, it is saved for later use in PRD2.CONFIG

*1 NOTE for VSE/ESA 2.4.0 and 2.4.1:

The JA to DMF collection phase was falsely stored in SKJOBACC, overwriting the traditional (non-dummy) JA to SYSLST routine in library 59.
Without the associated PTF UQ42227 for APAR PQ36445, use member SKJOBACC of VSE/ESA 2.3 (e.g.) to direct JA info to SYSLST.

Ù Suggested sequence of activities

1. Catalog JA routine $JOBACCT (SJADACC)
   (any time before JA record collection)
2. Catalog Report phase ACCTREP (SKJOBDMF)
   (any time before report creation)

//// collect DMF JA records //////

3. Offload DMF records DFHDFDOU (SKJADOFF)
4. Create Report ACCTREP (SKJADRPT)

App.B: VSE/ESA Sequential Disk Files

PART I.

App.B: VSE/ESA Sequential Disk Files

Sequential Disk Files (Summary)

Non-VSAM Files (just a reminder)

Sequential and Direct Access Method

	Disk Types	Access via
SAM	(E)CKD, FBA	DTFSD, DTFPH, DTFCP, DTFDI
DAM	(E)CKD	DTFPH, DTFDA

In the following, SAM via DTFSD is considered, besides SAM ESDS and VSAM ESDS.

Sequential Disk Files (Summary)

Type of Seq. File	DASD Space	Access via	Macros used	(E)CKD Channel Pgms (#SAM LBLOCKs/IO)
SAM (BAM file)	BAM	DTF	BAM	BAM (1)
SAM ESDS (DTF) (VSAM managed SAM)	VSAM	DTF	BAM	BAM (n)
SAM ESDS (ACB) (VSAM managed SAM)	VSAM	ACB	VSAM	VSAM (n x BUFND) *
VSAM ESDS (Native VSAM file)	VSAM	ACB	VSAM	VSAM (n x BUFND) *

\- SAM needs ',SD' in DLBL, all ESDS types need ',VSAM'

\- SAM ESDS can be accessed both via DTF and via ACB. Same record format on disk.

\- 1 SAM Log. Block = multiple Log. Records if Blocked. (Max.) Numbers shown hold for READs (n LBLOCKs per CI)

* 1 VSAM I/O can transfer at most 1 CA

Seq. Disk Perf. Hints (Summary)

Perf. Hints for Seq. Disk (Summary)

For files accessed sequentially, full DIM benefits cannot be achieved:

No reduction of # I/Os by reaccessing data in random access mode, i.e. by keeping lots of data in virtual memory.

BUT, as for most type of file accesses:

> The number of I/Os can be reduced by using a high I/O blocksize, i.e. a high #bytes per I/O

This is more important than a high physical blocksize on disk.

Naturally, if that value is also high, you get
- more data per track,
- and often shorter channel programs.

Both items are also of performance benefit.

- **Use SAM ESDS (via DTF) for higher I/O blocking**
- **Use VSAM ACB access (i.e. VSAM channel pgms)**

  SAM ESDS (ACB) or native VSAM ESDS

  providing
  - native ECKD (avoids ECKD conversion)
  - ECKD with SEQ indication, VSAM READ Ahead possible
  - SECTVAL avoidance
  - BUFND in DLBL
  - multiple SAM Logical Blocks per I/O

  BUT: Needs change in application.

- **For pure workfiles check use of Virtual Disk (FBA)**

  To reduce the number of FBA-I/Os by higher I/O blocksize, may not bring elapsed-time benefit, BUT ... may reduce the CPU-time overhead of Virtual vs. Real Disk.

  Although most programs can use FBA without reprogramming, optional DTF and JCL parameters are available for
  - maximum use of tracks
  - optimal performance

Types of Sequential Disk Files

Types of Sequential Disk Files

- **SAM Files (in BAM Space)**
- **SAM ESDS Files (in VSAM Space)**

  Fct. Benefits vs SAM file:
  - Dynamic space allocation on disk
  - IDCAMS utilities (BACKUP, REPRO ...) etc

  „ **OPENed via BAM DTF (DTFSD or DTFPH)**

  BAM macros used: OPEN, GET/PUT, CLOSE (+EXCP, WAIT for DTFPH)

  OPEN will automatically access VSAM catalog, instead of VTOC.
  No change in user programs required (vs SAM)

  Perf. Benefits vs SAM file:
  - Higher I/O blocking through CIs

  (NOCIFORMAT SAM ESDSs can only be accessed via DTFPH = EXCP)

  „ **OPENed via VSAM ACB**

  VSAM macros used

  Fct. Benefits vs BAM DTF access:
  - VSAM Catalog statistics maintained
  - Direct Access via RBA and Skip Seq or Backwards

  Perf.Benefits vs BAM DTF access:
  - VSAM ECKD Channel programs
  - Higher I/O blocking by multiple CIs

- **VSAM ESDS Files**

  VSAM macros used (OPEN, GET/PUT, CLOSE...)

  Fct. Benefits vs SAM ESDS ACB access:
  - Portable to OS/390

  Perf. Benefits vs SAM ESDS ACB access:
  - None, but possibility to use VSAM LSR (though without Read Ahead possibility)

File Formats

File Formats

- **Terms Used (Hierarchy)**

	SAM	SAM ESDS	VSAM ESDS
- SAM Logical Record	X	X	X
- SAM Logical Block	X	X	X
- SAM/VSAM Control Interval	x**	X	X
- SAM/VSAM Control Area	-	X*	X

* Control Area only used as disk allocation unit and max. blocksize per VSAM I/O
** FBA Control Interval also applicable to SAM on FBA

- **SAM Logical Record**

  An individual record seen by the application

- **SAM Logical Block (LBLOCK)**

```
<          SAM Logical Block (LBLOCK)              >

 B R   SAM Log.   R    SAM Log.          R   SAM Log.
 L L   Record 1   L    Record  2         L   Record 3


BL = Block Length   RL = Record Length
```

- **1 SAM LBLOCK = multiple SAM Log. Records**

  (1 if unblocked)

For SAM:
(also SAM ESDS in NOCIFORMAT)

- **1 SAM LBLOCK = 1 phys. block on CKD disk**

For SAM ESDS and VSAM ESDS:

- **Control Intervals (CIs) are used**

  also for SAM on FBA

File Formats ...

- **Control Interval (CI)**

  This CI-concept for SAM ESDS and VSAM ESDS is the same as the CI-concept for FBA, which is also applied to SAM files on FBA disk.

```
<                        SAM/VSAM CI                          >


 SAM LBLOCK 1    SAM LBLOCK 2   SAM LBLOCK 3          R R R C
or           or             or            Unused  D D D I
 VSAM LREC 1     VSAM LREC 2    VSAM LREC 3          F F F D
                                                     3 2 1 F


RDF  = Record Definition Field
CIDF = Control Interval Definition Field
```

- **Multiple SAM LBLOCKs = 1 CI**

  On FBA, 1 CI consists of multiple FBA 512-byte blocks.
  Since no 'FBA track boundaries' exist for the S/W, VSAM needs not select separate blocksizes as for (E)CKD.

- **1 SAM Logical Block = 1 VSAM Logical Record (SAM ESDS and VSAM ESDS)**

For SAM ESDS:

- Multiple SAM LBLOCKs are contained in 1 CI, usually 1 physical block on (E)CKD disk (sometimes multiple to get better track exploitation)
- SAM ESDS CI-size (if NOT explicitly defined via DEF CLUSTER) is selected during an implicit DEFINE at first OPEN time by VSAM

For VSAM ESDS:

- Multiple VSAM Log. Records (LREC) are contained in 1 CI, usually 1 physical block on disk (sometimes multiple to get better track exploitation)

File Formats ...

Ù Control Area (CA)

CAs are determined from tracksize, primary and secondary allocation, and by cylinder size.
1 CA is also the max. amount of data per I/O.

```
      <                SAM ESDS CA                        >

 CA 1    CI 1       CI 2        CI 3         CI 4      CI 5
                                                      part1


 CA 2  CI 5     CI 6        CI 7        CI 8         CI 9
                part2

      SAM ESDS CIs can span CAs


      <               VSAM ESDS CA                       >

 CA 1    CI 1       CI 2       CI 3         CI 4     unused


 CA 2    CI 5       CI 6       CI 7         CI 8     unused


      VSAM ESDS CIs can NOT span CAs
```

For SAM ESDS:

Control Areas (CA) as in VSAM exist, BUT only are used as increments for disk space allocation ('non-CA-format').
CIs can span CAs, resulting in more data per MB disk space.

For VSAM ESDS:

Multiple CIs are contained in a VSAM Control Area.
CIs cannot span CAs.

Performance Relevant Parameters

Perf. Parameters for Seq. Disk Files (Summary)

	SAM	SAM ESDS (DTF)	SAM ESDS (ACB) *3	VSAM ESDS	Disk Type
DLBL BLKSIZE	X	X	-	-	(E)CKD
overrides (if blocked) DTFSD BLKSIZE	X	X	-	-	(E)CKD
DLBL CISIZE *1	X	-	-	-	FBA
overrides DTFSD CISIZE	X	-	-	-	FBA
DLBL RECORDS,RECSIZE *2	-	X	X	-	all
DTFSD RECSIZE	X	X	-	-	all
DLBL BUFND,BUFSP	-	-	X	X	all
DTFSD IOAREA	X	-	-	-	all
VSAM DEF CLUSTER values	-	-	-	X	all

*1 Only for SAM Output files
*2 Only used for disk allocation
*3 Implicit SAM ESDS DEFINE by VSAM not considered here

Performance Relevant Parameters ...

Performance Parameters for Seq. Disk Files

For applicability of parameters, check previous summary

Ù DLBL parameter (VSE JCL)

DLBL BLKSIZE=n
  Size of the I/O area, allows bigger I/O blocksize than coded in DTFSD BLKSIZE.
  Must be a multiple of RECSIZE in case of FB.

DLBL CISIZE=n
  Control Interval size for SAM on FBA, overrides DTFSD CISIZE.

DLBL RECORDS=(n1,n2),RECSIZE=m
  Specifies the number(s) (and avg. size) of log. records.
  Is only used to calculate primary and secondary allocations.

DLBL BUFND=n (or BUFSP=n)
  Number of VSAM data buffers used for I/Os.
  Can overrule (if larger) BUFND in the VSAM ACB.

Ù DTFSD parameter ('Application')

DTFSD BLKSIZE=n
  Size of the I/O area = size of a SAM Log. Block.
  Must be a multiple of DTFSD RECSIZE in case of FIXBLK.
  Can be overruled by DLBL BLKSIZE if non-VSAM file is blocked.

DTFSD CISIZE=n
  Size of the Control Interval for FBA (SAM).
  In contrast to manual, is also the VSAM CI-size for an implicit DEFINE of this file as SAM ESDS for all type of disks.

DTFSD RECSIZE=n
  Size of a SAM Log. record in case of RECFORM=FIXBLK.
  Can not be overruled by any DLBL specification.

DTFSD IOAREA2=name
  Two I/O areas are used by BAM GET/PUT for overlapped BAM I/Os.

DTFSD VERIFY=YES
  VERIFY is not required/recommended for performance reasons (WRITE-cached I/O subsystems would NOT check on disk anyhow)

Performance Relevant Parameters ...

Ù VSAM DEFINE CLUSTER parameter

An explicit DEFINE CLUSTER can be done for SAM ESDS (DTF or ACB), besides an implicit DEFINE for SAM ESDS (DTF) (next foil).

NONINDEXED

  Required for SAM ESDS and VSAM ESDS

RECORDFORMAT

  Required for SAM ESDS (F, FB, V, VB, NCIF ...)

RECORDSIZE (avg max)

  Required for fixed record format (F, FB), defaults to (4089 4089) for V and VB.
  Max is the largest SAM logical block size and is only used to calculate the CI-size.

RECORDS (primary secondary)

  Number of VSAM records, used for primary and secondary allocation

RECOVERY|SPEED

  Use SPEED in any case.
  Avoids the writing of an empty CI after each data CI.

BUFFERSPACE

  Code this value slightly bigger than the default of 2 CIs.
  Use BUFND in DLBL to further increase.

CONTROLINTERVALSIZE

  Code this value big enough.
  Must be at least 7 bytes larger than the max. VSAM record size. If coded, it avoids the selection
   - of 2K if VSAM RECORDSIZE specified
   - of 4K if VSAM RECORDSIZE not specified

NOWRITECHECK

  Always use this default, never use unnecessary WRITECHECK

Performance Relevant Parameters ...

Ù **Implicit SAM ESDS DEFINE by VSAM**

A (reusable) SAM ESDS file may be implicitly defined during SAM ESDS OPEN, where VSAM selects/determines from DTFSD and DLBL

- the record format
- the SAM ESDS CI-size
- the SAM Logical Block size (= VSAM Logical Record size)
- the SAM Logical Record size for FIXBLK.

NONINDEXED, NOWRITECHECK and SPEED

Always used.

RECORDFORMAT

Derived from DTFSD RECFORM

RECORDSIZE (avg max)

max Derived from DTFSD BLKSIZE
= SAM log. blocksize (=VSAM log. recordsize)

avg From DLBL RECSIZE, otherwise = max

RECORDS (primary secondary)

Derived from DLBL RECSIZE and DLBL RECORDS

BUFFERSPACE

Default of 2 VSAM data-CIs is used in the catalog.

CONTROLINTERVALSIZE

Taken over from DTFSD CISIZE (any type of disk).
If not given in DTFSD, VSAM chooses a CI-size.
With DTFSD IOAREA2, >1 SAM logical blocks per CI.

Í **It may be straightforward, to explicitly DEFINE CLUSTER a SAM ESDS file**

Though, implicit DEFINE may be more flexible for workfiles since always done anew.

Sequential Disk I/Os

Sequential Disk I/Os (Real Disk)

Ù **BAM channel programs (i.e. DTF access) ...**

Only 1 SAM logical block (= 1 phys. block on disk) transferred (read/written) per SSCH

WRITE I/Os are done when buffer (BLKSIZE) full

CKD channel programs only

- must undergo the ECKD conversion routine
- cannot use a SEQuential indication
- SECTVAL SVCs used (before VSE/ESA 2.5)

Ù **VSAM channel programs (i.e. ACB access) ...**

Multiple CIs (each mult. SAM logical blocks) can be read per SSCH

VSAM READ Ahead possible for
- native ESDS with NSR (not LSR)
- when OPENed with SEQ AND GET NEXT used instead of GET(argument)

Only 1 CI can be written per SSCH

WRITE I/Os are done usually when CI full

SEQ processing in ACB and adding records at EOF cause VSAM to use 'Delayed WRITE', instead of 'Immediate WRITE' (VSAM always loads and extends a SAM ESDS in SPEED mode)

Optimal ECKD channel programs

Í **Higher I/O blocking possible for ACB access**

Í **More effective channel programs for ACB access**

Seq. Disk Performance Hints

General Performance Hints

Ù **Sequential file performance usually is the better**

the more data are read/written per disk I/O

the bigger the physical blocksize on disk

the better the channel programs used

Í **Try to use SAM ESDS (DTF) instead of SAM**

Í **Try to use SAM ESDS (ACB) or VSAM ESDS instead of SAM ESDS (DTF)**

Both would need an application change vs. SAM.
So, if change possible, you may directly use VSAM ESDS

Hints for SAM files

Ù **Select a large(r) DLBL BLKSIZE**

Does DLBL BLKSIZE > LBLOCK size help for WRITEs??

Since all BAM WRITEs are format-WRITEs, the remainder of the track is always erased. Even more reason to use bigger blocks

Seq. Disk Performance Hints ...

Hints for SAM ESDS files (DTF access)

Ù **See hints for SAM, also using DTFSD**

Ù **Consider an explicit VSAM DEFINE CLUSTER**

Ù **Care for a reasonably big CI-size**

Hints for SAM ESDS files (ACB access)

Ù **Care for a reasonably big CI-size**

Ù **Select a reasonable number for DLBL BUFND**

Ù **Avoid a too small CA-size**

Per VSAM I/O at most 1 CA is transferred.

Selection of RECORDS and RECSIZE in DLBL
Selection of RECORDS and RECORDSIZE (max) in DEFINE CLUSTER

Hints for VSAM ESDS files

Ù **See all SAM ESDS (ACB) hints**

App.C: VSE/ESA Librarian

PART J.

App.C: VSE/ESA Librarian

VSE Librarian Performance

Backgound Information

„ **Librarian blocksize on disk is 1K (1 LB block)**

Applies to ALL VSE libraries.

This blocksize of 1K (data + control info)

- is (too) big for small members,
  but still 'allows' only 1 member in 1 LB block

- is (too) small for DUMP libraries,
  and gives smaller track exploitation for CKD/ECKD
  plus needs more CCWs to be setup

„ **I/Os to VSE libraries transfer multiple LB blocks**

Only in exceptional cases 1 LB block is used
(e.g. the LIBR TEST cmd to check VSE Library integrity)

The number of LB blocks per I/O is limited by ...

- the number of I/O buffers available
- the number of contiguos blocks used for a member
- the end-of-cylinder or extent

... whatever comes first

Í **I/O blocking in LIBR is done well**

„ **I/Os to VSE libraries**

- are done by the Librarian
(EXEC LIBR: BACKUP, RESTORE...)
- are done by Program Load
(FETCH)

If a VSE library is in VSAM managed space, a difference exists only at OPEN time
Only at OPEN, VSAM is included.
Refer to separate chart

Librarian Member Scattering

Scattering of (Sub-)Library Members

Ù **Scattering only occurs if**

„ **A member is >1K**

and

„ **Members are cataloged concurrently (by different tasks) into the same library**

or

„ **Members are cataloged by a single task into a sub-library AND not enough contiguous freespace is available**

in the Free Chain of LB blocks of that library

Í **Production or IBM supplied libraries are not impacted**
(libraries w/o cataloging members after deletes)

Ù **Performance Impact**

„ **Additional I/O at each non-contiguous point**

„ **As a rough estimate**
e.g. '10%' of members scatterd, gives

- about 10% more elapsed time for BACKUP
- about 1% to 2% more elapsed time for execution (FETCH)

Ù **Should Scattering hurt ...**

Í **Do a LIBR BACKUP/RESTORE of the Library**
Sub-library B/R doesn't change anything

Librarian VS Requirements

Librarian Virtual Storage Requirements

Ù **Librarian Internal Tables in SVA GETVIS-24**

At IPL time about the following area is required and kept for Librarian Tables in System GETVIS-24:

```
8.5K + 0.2K x SUBLIB  + 0.8K x NPARTS
```

IPL SYS	default	shipped
SUBLIB	100	100
NPARTS	12	44

So, a shipped system uses about 64K for Librarian table space.

Ù **Other Librarian areas in SVA GETVIS-24**

These areas are used for Level 2 services (Librarian internal) on a permanent base ('touched since IPL'):

- Contains e.g. Library and Sub-library info
  (libr. descriptor, bit map, sub-libr. descriptor)

- These areas are 1K and called 'shared buffers'.

Ù **Librarian I/O buffers in Partition GETVIS-24**

Used for Level 3 services, i.e. for all LIBR commands:

- Each buffer is 1K

- As many buffers as Partition GETVIS-24 allows are allocated

- These buffers are also called 'private buffers'.

Library in VSAM vs BAM Space

Library in VSAM vs BAM Space

- **Libraries in VSAM space are extendable**
- **Only IJSYSRS must reside in BAM space**
- **Only performance deltas at OPEN (and CLOSE)**

  OPENs are done for a file (= Library), and passed from BAM to VSAM, if req'd.

  (VSAM passes the extent info back to the Librarian).

  It is done (only) at first access to a Library, for every extent (BAM and VSAM), 'clustered', up to 16. For VSAM, in addition at each dynamic extension

  OPEN of a library is followed directly by CLOSE,

  except for the last extent if that extent is in BAM space.

- **Some observations (VSAM vs BAM space)**

  - **Librarian BACKUP/RESTORE for VSAM space differs from BAM space only by the small OPEN part**

    About 50 msec CPU-time on a 10 MIPS processor, plus about 50 I/Os

  - **LIBDEFs for VSAM Space require about 2 times the resources as for BAM space**

    Refers to Elapsed time, CPU-time, # I/Os.

    But both are fast enough, so this not an important critera.

Librarian Performance Hints

Performance Hints

- **Leave $IJBLBR.PHASE in SVA-24**

  Phase MUST reside in SVA-24 for functional reasons (about 256K)
  No other librarian related phase is mandatory in the SVA (VLA)

- **Put Librarian phase names into the SDL**

  Recommended when // EXEC LIBR is used often.

  This will save some search effort for the LIBR phase and its overlays, IF the phase resides in IJSYSRS.SYSLIB.
  Just put the phase name into the SDL, not the phase itself into the SVA; i.e. w/o ',SVA'

- **Define few -and large- libraries**

  This will optimize your space management

- **Define many -and small- sub-libs**

  Reduces the number of index-levels for members; and thus the number of LIBR I/Os.
  Applies to all accesses, except for FETCH/LOAD (There, SLDs are being built).

  Use LIBR TEST SUBLIB to see 'NUMBER OF INDEX LEVELS'.
  (Note that this value only changes after a RESTORE, not just by deleting members).
  Refer to Librarian Hints in 'Hints and Tips for VSE/ESA'.

  (Notice:
  NO choice for IBM provided libraries, due to service reasons)

- **Recommended REUSE= specification:**

  AUTOMATIC for sub-libs used for production

  (local directory lists remain valid after delete)

  IMMEDIATE for sub-libs for pgm development

  (only minor performance impact by immediate updates of free chain in VS and on disk, of SLD and SDL)

Librarian Performance Hints ...

Performance Hints (cont'd)

- **Carefully set up your LIBDEFs**

  as short as possible
  as optimal in sequence as possible

  Additional I/Os are required for searches in preceding sub-library directories of the LIBDEF chain

- **Put the most used sub-libraries into PERManent LIBDEFs**

  OPENs are saved.

  A VSE Library OPEN is done only when the first sublib is touched. Subsequent LIBDEFs for the same or other sublibs in that library will not require OPEN processing.

  It only has to be RE-OPENed, if no started partition has pointed to any of its sublibs via LIBDEF meanwhile.

- **For jobs with many job steps, override previous definition with**

  ```
  // LIBDEF obj,SEARCH=(lib.sublib, ...),TEMP
  ```

- **RESTORE (and COPY/MOVE) is faster, if the target library (i.e. all sub-libraries) is NOT accessible from any other partition**

  In case other partitions have pointers via LIBDEF or via ACCESS/CONNECT Librarian commands ...

  More effort has to be spent to use the table for sharing DASD space between sublibraries

Librarian Performance Hints ...

Performance Hints (cont'd)

- **Reorganize libraries by doing BACKUP/RESTORE on Library level**

  Space is shared between sublibs.

  Fragmentation is no problem in general, but may occur.

  Refer to foil 'Librarian Member Scattering'

- **A high number of library extents (up to 16) only cause a performance degradation at actual OPEN time**

  - For FETCH/LOAD,
    OPEN is done only once (not critical)
  - For LIBR commands (incl. LIBDEF),
    OPEN is done also only once.

- **Avoid libraries (that are NOT shared between VSEs) on volumes which are defined as shared**

  OPEN and member WRITE ...
  would be done via the LOCK-file, instead of only via the LOCK table.

  (Member WRITE can re-use the freespace inventory, if the library is also non-shared across partitions)

  Member retrieval (READ) ...
  does not require locking and thus would be same (FETCH/LOAD and LIBR access).

LIBRM Macro Interface Hints

LIBRM Macro Interface Hints

Ù **LIBRM I/F to access data stored in VSE libs**

A user API, using internal LIBR level 2 services

Ù **Some LIBRM functions**

LIBRM OPEN	OPENs a specific Librarian member
LIBRM GET	Retrieves data of a member into the caller's workarea
LIBRM PUT	Writes data out of a callers workarea into a specific Librarian member

Ù **Recordformats**

RECFM= F|S	F=fixed records (80 byte) S=string

Ù **I/O aspects**

LIBRM uses the same LIBR buffers as all other LIBR functions.

LIBRM GET:

If data is not found in the internal LIBR buffers (in partition GETVIS), multiple LIBR blocks can be read, (in 1 SSCH if unscattered)

LIBRM PUT:

Data is immediately written to disk, multiple LIBR blocks in 1 SSCH is possible.

Ù **Concurrency**

1 user program (partition) can at 1 point-in-time only have 1 LIBRM call pending.
When function is done, control is being returned to the caller.

LIBRM Macro Interface Hints ...

LIBRM Performance (cont'd)

Ù **LIBRM Performance Hints**

í **Always use a big enough caller workarea**

(BUFSIZE= in LIBRM OPEN)

This is a pre-req to reduce the number of LIBRM calls and also the #physical I/Os (SSCHs)

í **Always GET and PUT as enough data as possible**

(UNITS= in LIBRM GET/PUT)

UNITS=0 should be used for F, that is, exploiting all workarea.

Biggest overhead is PUTting single 80-byte records.

This reduces the number of LIBRM calls and also the #physical I/Os

í **Tracing LIBRM is possible**

via LIBR TEST TRACE=LEVEL2

VSE/ESA Growth

PART K.

VSE/ESA Growth

VSE/ESA Growth

VSE/ESA Growth -General Remarks-

Ù **Storage Evolution in VSE/SP**

A look far back:

VSE/SP Version	Real Storage in S/370 mode	Virtual Storage	# Partitions	#Address Spaces
1	8 MB	16 MB	7 to 12	1
2	16 MB	40 MB	12	3
3.1	16 MB	40 MB	12	3
3.2	16 MB	128 MB	12	9
4	16 MB	128 MB	12	9

See also 'Growth Potential in the VSE Environment' GG24-3358-00, ITSC BOE, 03/89

Ù **VSE/ESA (compared to VSE/SP) has removed or mitigated many growth inhibitors**

e.g.

- More than 9 address spaces - More total virtual storage - ESCON and ECKD	VSE/ESA 1.1
- 2 GB Real - 31-bit applications + DIM - Data Spaces, Virtual Disk	VSE/ESA 1.3
- n-way support (TD)	VSE/ESA 2.1
- VSCR - LTA offload	most releases many releases
- CICS TS (VSCR) - VSAM LSR Hashing	VSE/ESA 2.4 VSE/ESA 2.5

Ù **Since 2000 we have customers exploiting up to 200 and more MIPS with a single VSE/ESA**

Whatever you can do within VSE (and vendor products) can be done: nice for you, nice for IBM.
It is by far much more than everybody would have thought, then.

We have not built in any artificial limitations, new limits may/will emerge.
If effort to remove a certain limit is too high, we search for circumventions or alternatives

VSE/ESA Growth ...

VSE/ESA Growth -General Remarks- (cont'd)

- **VSE/ESA growth has received/needed some smaller focus after VSE/ESA 2.1**

  - Multi-platform has become a reality for most customers
  - Interoperability was greatly enhanced (VSE/ESA 2.5 and up)
  - OS/390 is 'closer' to VSE/ESA (regarding migration)
  - Only few current growth problems arose

- **OS/390 (z/OS) is the natural target for fast growing VSE/ESA systems/customers**

- **LINUX for S/390 is a very appropriate complement to VSE/ESA**

VSE/ESA Capacity History

VSE/ESA Capacity History

Achievable MIPS values for high end VSE, for average workloads (including some tuning).

- **04/79: >2 MIPS with regard to VSE/POWER**
  Results from S370/158

- **07/79: >7 to 9 MIPS with regard to I/O scheduler**
  Extrapolation based on S370/168

- **05/85: 15 MIPS actual, VSE/SP 2.1**
  Reported in SETI flash

- **12/89: >11-15 MIPS for VSE/ESA 1.1**
  >25-40 MIPS for VSE/ESA 1.3
  VSE/ESA Capacity Analysis study by W. Kraemer

- **12/00: 200+ MIPS actual, VSE/ESA 2.3**

VSE/ESA Hard Limits

Hard Limits for VSE/ESA Growth

- **Not enough VSCR, in spite of all improvements**

  Many system functions, control blocks, non-VSAM buffers etc. still below the line

  Will remain to be important in spite of additional VSCR by CICS TS and in LE/VSE, since further growth in single VSEs is limited

- **VTAM IOBUF areas in System GETVIS-24**

- **Non-Parallel-Share (NPS) limits n-way support**

  Important for those technologies/processors with very limited MIPS per engine.
  Not a problem on recent CMOS e-Servers with 'enough' MIPS per engine.

- **Task limitations**

  Only up to 255 tasks in total for VSE/ESA,
  up to 31 VSE subtasks per partition,
  up to 208 VSE subtasks in total.

VSE/ESA Soft Limits

Soft Limits for VSE/ESA Growth

- **Missing Integrated System Concept/Functions**

  - **for automation**
  - **for performance monitoring**
  - **for automatic hierarchical data management**

- **Missing Functions/Applications**

  - **for Java applications**
  - **for device support**
  - **for middleware-type of apps**

- **Missing Support for S/390 Hi-end Performance Functions**

  - **Parallel Access Volumes (PAV, on ESS)**
  - **expanded storage**
  - **sort of Library Lookaside**

- **Limitations in Number of Users per TCP/IP Partition**

  - **Telnet users**
  - **GPS users**
  - **external sockets**

Other Capacity Related Resources

Some Other Capacity Related Resources

NOT being considered to be a current actual limit

Ù **Number of partitions**

The limit is 12 static + 150-200 dynamic partitions, currently determined by the number of VSE tasks.

VSE JCL has a limit of 212 VSE partitions.

Ù **Real storage**

The limit is 2 GB.

Using an old rule-of-thumb (6MB real / MIPS), the VSE/ESA 2G limit alone would allow about 340 MIPS to be consumed by a single VSE/ESA with still some reasonable DIM.

Note that e.g. OPTI-CACHE exploits expanded storage for VSE/ESA transparently (for its VSE-wide file buffers)

Ù **Total Virtual Storage**

The limit for VSE/ESA is 90 GB.

This is a much higher limit than the 2G real limit.

The maximum size of a VSE Page Data Set is determined by the maximum number of logical devices (15), with 3390-9's this would result in 108 GB.

Ù **Total number of devices (and 'channels')**

The limit is 1024 devices (and 16 'channels').

With the usage of SnapShot and FlashCopy, the number of required logical disks may easily double.

16 'channels' are caused by the use of 'CUU' for a logical device (from 000 to FFF) and by interpretation of the 'C' as 'channel'.

So 'channel' in VSE has no relation to physical channels, thus does not represent any real resource.

Other Capacity Related Resources ...

Some Other Capacity Related Resources (cont'd)

Ù **Total number of logical units**

The limit is 255 per partition, and 12x255 =3060 in total.

Ù **Dispatching**

Turbo Dispatcher has introduced VSE Relative Shares.
Allows all partitions to be put into the (single) VSE balanced group.

Ù **Label Area**

Capacity was increased, and put into native Data Space.

Current limit is about 9000 labels in total,
and 712 label sub-areas.

Ù **Single LTA**

Only a problem if still misused.

Refer to 'Enhanced Label Area and LTA' in
'VSE/ESA V2 Performance Considerations'

EOD/HAND

EOD End of Document

HAND Have a nice day