VSE/ESA 2.6 and 2.7

Performance Considerations

Ingo Franzki

e-mail: ifranzki@de.ibm.com

VSE/ESA Development

Contents

- VSE/ESA 2.6
  - ƒ Performance Items (Reminder)
- VSE/ESA 2.7
  - ƒ Performance Items (Overview)
  - ƒ HiperSockets
  - ƒ Hardware Crypto support
- z800/z900 Remarks
- Dependencies for VSE/ESA Growth

VSE/ESA 2.6 Performance Items

▪VSE/ESA 2.6 Base enhancements
- ƒ Delete Label Function
  - − saves >90% SVCs for sequential processing
- ƒ LTA Offload for some AR commands
  - − Less I/O by less FETCHes for LTA load
- ƒ SVA-24 Phases moved above the line
  - − $IJBPRTY (6KB)
- ƒ Increased max number of SDL entries
  - − Maximum value now 32765
- ƒ SDL update from non-BG partitions
- ƒ POWER Data file extension without reformat

VSE/ESA 2.6 Performance Items - continued

▪VSE/ESA 2.6 Hardware Support
- ƒ FICON Support (VSE/ESA 2.3 or higher)
- ƒ New 2074 System Management Console
  - − ESCON channel attached
  - − Eliminates requirement for a non-SNA 3174 controller
- ƒ OSA Express Adapter (e.g. Gigabit Ethernet)
  - − Available for G5 and above
- ƒ VSAM Support for large 3390-9 Disks (Shark)
- ƒ Fastcopy Exploitation of ESS FlashCopy and RVA SnapShot

VS *@*

e-business

Hardware Support

- Queued Direct I/O
  - ƒ Designed for very efficient exchange of data
  - ƒ Uses the QDIO Hardware Facility, without traditional S/390 I/O instructions
  - ƒ Without interrupts (in general)
  - ƒ Use of internal queues
  - ƒ With pre-defined buffers in memory for asynchronous use

- Exploitation by TCP/IP for VSE/ESA
  - ƒ see TCP/IP Performance Considerations

IBM.

VS *@*

e-business

VSAM SHROPT(4) Avoidance

- Connectors in VSE/ESA 2.5 require SHROPT(4) when updating VSAM files owned by CICS
- New VSAM-via-CICS Service avoids SHROPT(4) by routing the VSAM requests to CICS
- Communication between batch and CICS is XPCC
- Naming convention for "VSAM-via-CICS files"
  - ƒ Each CICS is treated as "virtual" catalog
  - ƒ Files defined in CICS (via CEDA DEFINE FILE) are visible within this catalog
    #VSAM.#CICS.<applid>

IBM.

indicates "virtual"   APPLID of CICS region owning the files
CICS catalog   within this catalog

VSAM Redirector

- New connector with VSE/ESA 2.6
  - ƒ VSE is client
  - ƒ PC / workstation is server
- Exploits VSAM exit IKQVEX01
- Allows to redirect one of more VSAM files to a PC or workstation
- All VSAM requests of a particular file are redirected
  - ƒ Open / close
  - ƒ Get / put / point / delete / insert
- Transparent for applications
  - ƒ Usable from batch and CICS

VSAM Redirector - Performance Implications

- Is the file redirected ?
  - ƒ No: only at OPEN time (very small overhead)
  - ƒ Yes: at each request
- Network overhead ?
  - ƒ Yes, if file is redirected
  - ƒ Depends on
    - – Number of VSAM requests
    - – Size of records
- Data ownership
  - ƒ OWNER=REDIR
    - – no VSAM I/O

VSE/ESA 2.7 Performance Items

▪VSE/ESA 2.7 hardware support

ƒ z800/z900, Multiprice 3000, G5/G6
ƒ HiperSockets
ƒ Hardware Crypto Support
ƒ 32760 cylinder 3390 support
ƒ 3590 buffered tape mark

▪VSE/ESA 2.7 enhancements

ƒ new TCP/IP for VSE/ESA release 1.5
ƒ $IJBLBR above the line
ƒ II User Status Record above the line
ƒ VTAPE: removed DVCDN/DVCUP
ƒ POWER: reallocate queue file during warm start

VSE/ESA 2.7 Hardware support

▪VSE/ESA 2.7 runns on the following machines
ƒ z800 (2066)
ƒ z900 (2064)
ƒ 9672 Parallel Enterprise Server (G5/G6)
ƒ Multiprise 3000 (7060)
ƒ equivalent emulators (Flex-ES)

▪VSE/ESA 2.7 is based on the hardware instruction set described in the manual 'ESA/390 Principles of Operation' (SA22-7201).

▪With VSE/ESA 2.7 it is assumed that all the ESA/390 instructions and facilities described in that manual can be used.

VS @ e-business

z800/z900 Remarks

- Prior to z800/z900 there is one cache for data and instructions
- z800/z900 has splited data and instruction cache
- Performance implications:
  - ƒ If program variables and code thjat updates these program variables are in the same cache line (256 byte)
    - Update of program variable invalidates instruction cache
    - Performance decrease if update is done in a loop
  - ƒ See APAR PQ66981 for FORTRAN compiler

VS @ e-business

32760 cylinder 3390 support

- With announcement 101-341 at 11/13/2001 IBM announced the new 32760 cylinder 3390 volumes of the IBM TotalStorage Enterprise Storage Server (ESS).
  - ƒ This enhancement of the ESS F models was made available 11/30/2001.
- VSE/ESA 2.7 now supports these volumes.
  - ƒ helps relieve address constraints
  - ƒ improves the disk resource utilization
  - ƒ can be used to consolidate multiple disk volumes into a single address.

3590 Buffered Tape Mark support

- The 3590 control unit provides support for writing tape marks (TM) in buffered mode
- Writing TM's in "buffered" mode should enhance the performance
  - ƒ of all programs, which write many TM's as part of their file creation process (e.g. POFFLOAD)
- All the TM's written during OPEN/CLOSE (label processing) will remain to be written "UNbuffered"
  - ƒ all the programs which write TM's mainly or only during OPEN/CLOSE, will NOT benefit from this enhancement.

$IJBLBR phase moved above the line

- The $IJBLBR.PHASE has been splited into two phases
  - ƒ $IJBLBR.PHASE
  - ƒ $IJBLB31.PHASE
- $IJBLBR.PHASE will continue to reside in SVA-24
- $IJBLB31.PHASE will reside in SVA-ANY (high SVA).
  - ƒ This will free about 180k in SVA-24.

VS @

e-business

II User status record above the line

▪During Logon each II user gets besides others two
 storage areas allocated
  ƒ User_Status_Record USR (904 bytes)
  ƒ Panel_Hierarchy_List PHL (1352 bytes)
  ƒ originally located in the CICS DSA (below)
▪With VSE/ESA 2.7 the USR and PHL has been moved
 to ESDSA (shared above)
  ƒ frees 2.3 KB in DSA below per user.
▪ICCF TCTUALOC=ANY now supported
  ƒ ICCF transaction programs has been changed to
    support a TCTUA (28 bytes) above the line

IBM.

VS @

e-business

**HiperSockets hardware elements
('Network in a box')**

▪Synchronous data movement between LPARs and
 virtual servers within a zSeries server
  ƒ Provides up to 4 "internal LANs" HiperSockets
    accessible by all LPARs and virtual servers
  ƒ Up to 1024 devices across all 4 HiperSockets
  ƒ Up to 4000 IP addresses
  ƒ Similar to cross-address-space memory move using
    memory bus
▪Extends OSA-Express QDIO support
  ƒ LAN media and IP layer functionality (internal QDIO
    = iQDIO)
  ƒ Enhanced Signal Adapter (SIGA) instruction
    −No use of System Assist Processor (SAP)

IBM.

VS @

e-business

HiperSockets hardware elements
('Network in a box') - continued

- HiperSockets hardware I/O configuration with new CHPID type = IQD
  - ƒ Controlled like regular CHPID
  - ƒ Each CHPID has configurable Maximum Frame Size
- Works with both standard and IFL CPs
- No physical media constraint, no physical cabling, no priority queuing
- Secure connections

IBM.

VS @

e-business

Measurement Environment

- z800 (2066-004)
  - ƒ 4 processors
- VSE/ESA 2.7 GA Driver in a LPAR (native)
  - ƒ 1 CPU active (~2066-001)
  - ƒ TCPIP00 (F7): OSA Express Fast Ethernet
  - ƒ TCPIP01 (F8): HiperSockets
- Linux for zSeries in a LPAR (native)
  - ƒ 3 CPUs active (shared)
  - ƒ eth0: OSA Express Fast Ethernet
  - ƒ hsi10: HiperSockets
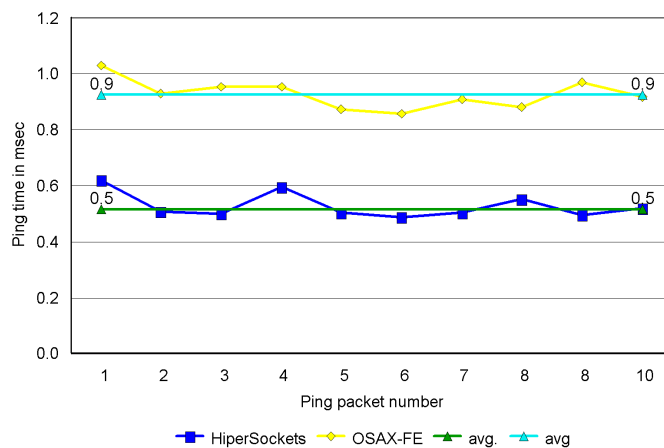
IBM.

Latency (Round trip time) - results

■Measurements has been done with PING command
ƒ Issued at Linux side
ƒ 10 Pings
ƒ PING sends a datagram to VSE
ƒ VSE sends a answer back to Linux
ƒ Time until answer arrives is measured
– Round trip time

Latency (Round trip time) - results

HiperSockets is about 1.8 times faster in terms of latency

VS @

e-business

Throughput (MB/sec)

- Measurements has been done with FTP
  - ƒ Initiated at the Linux side
  - ƒ Transferring 1GB (1000MB)
    - without translation (binary)
    - 1 to 5 parallel streams
  - ƒ PUT: send data to VSE
    - VSE inbound
    - sending a 1GB file to $NULL file (in memory file)
    - No file I/O is done by VSE/Linux
  - ƒ GET: receive data from VSE
    - VSE outbound
    - receiving $NULL file (in memory file) into /dev/null
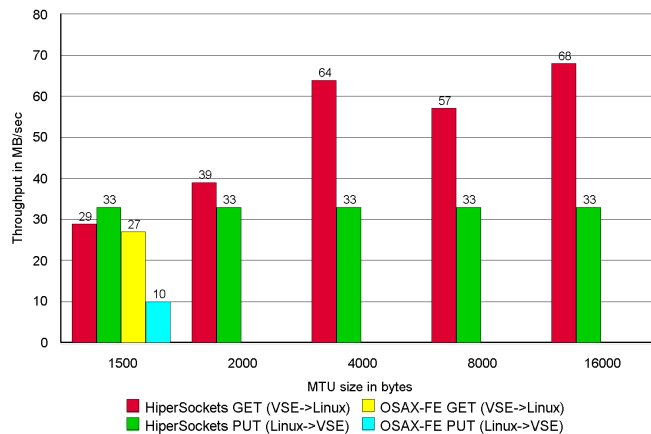   - No file I/O is done by VSE/Linux

IBM.

VS @

e-business

Throughput (MB/sec) - results
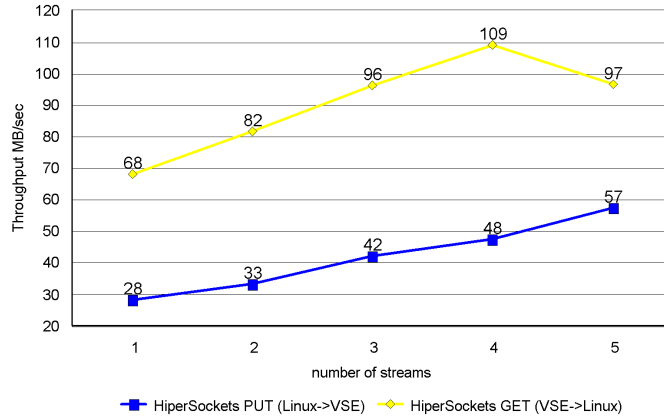

HiperSockets throughput is between 30-80 MB/sec

IBM.

Throughput (MB/sec) - results (2)

Maximum HiperSockets throughput of 109 MB/sec at 4
concurrent connections

CPU time per MB - results
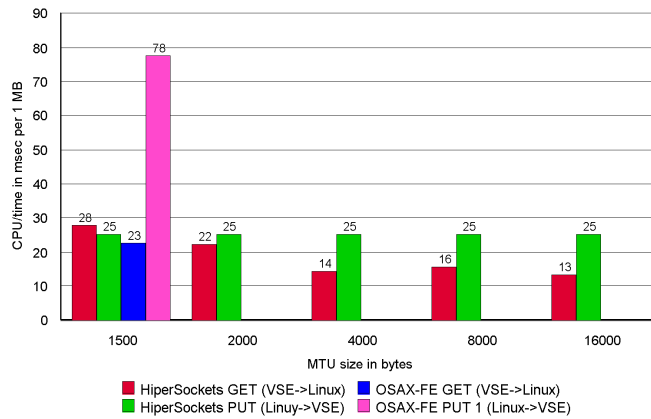
About 15-30 msec CPU time per MB for HiperSockets
(on a z800 2066-001)

Transaction per second

▪Measurements has been done with an ECHO server
  ƒClient on Linux sends 100 bytes to server
  ƒServer on VSE echoes 100 bytes
  ƒPer TCP connection 10000 transactions are driven
  ƒVariations: Number of TCP connections
    −1,2,3,4,5
    −10,50,100
  ƒMeasurements
    −Transactions per second
    −CPU time per transaction

Transactions per second - results
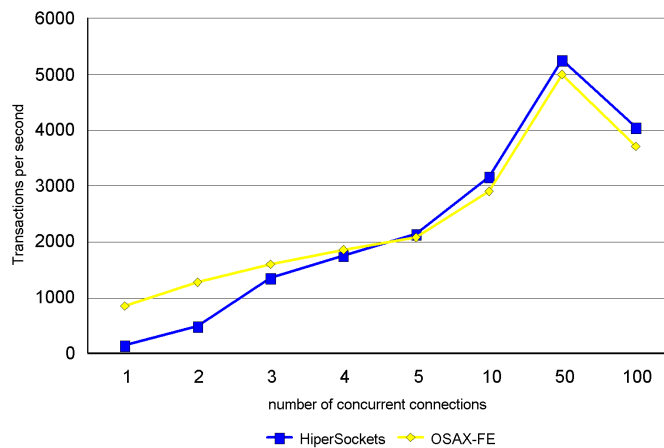
Maximum of 5200 transactions per second at 50 concurrent connections

Transactions per second on 1 connection - results

VS @ e-business

IBM.

```
Transactions per second on 1 connection
1000
 900
 800
 700
 600
 500
 400
 300
 200
 100
   0
      1    2    3    4    5   10   50  100
         number of concurrent connections
```

■ HiperSockets   ◇ OSAX-FE
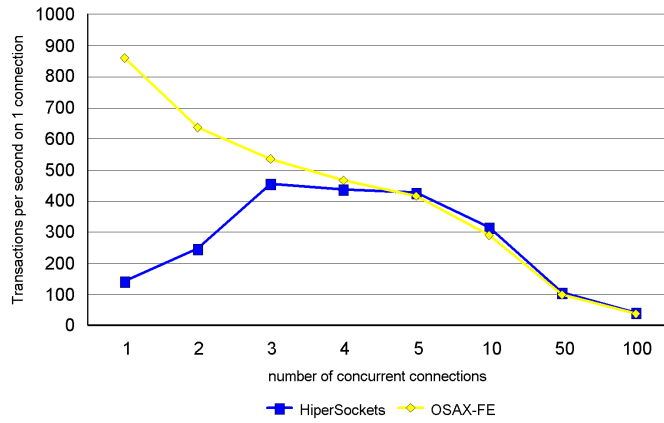
HiperSockets: Maximum of about 450 transactions per second
on 1 connection   (= about 2 msec response time)

CPU time per transaction

VS @ e-business

IBM.

```
CPU time per transaction in msec
0.6
0.55
0.5
0.45
0.4
0.35
      1    2    3    4    5   10   50  100
         number of concurrent connections
```

▲ HiperSockets   △ OSAX-FE
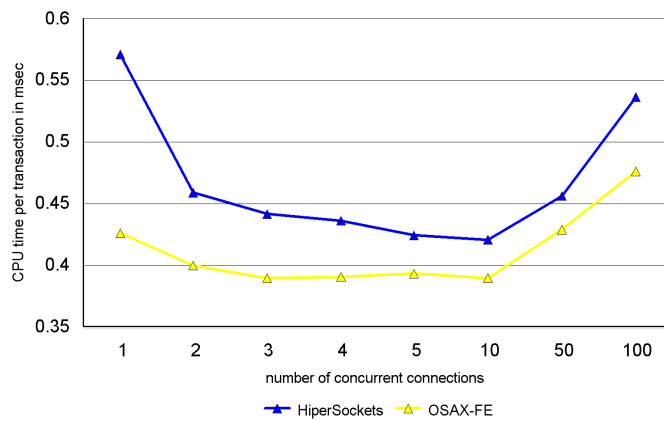
HiperSockets: About 0.45 msec CPU time per transaction
            for 2-50 connections

14

VS@

e-business

Measurement Results - conclusion

- HiperSockets
  - ƒ Throughput
    - − Between 30-80 MB/sec
    - − Maximum throughput of 109 MB at 4 connections
    - − About 15-30 msec CPU time per MB
  - ƒ Transactions per second
    - − Maximum of 5200 Transactions per second at 50 connections
    - − About 0.4-0.45 msec CPU time per transaction

IBM.

VS@

e-business

Hardware Crypto Overview

- Requires VSE/ESA 2.7 and TCP/IP for VSE/ESA 1.5
- Supported crypto cards
  - ƒ PCI Cryptographic Accelerator (PCICA)
    - − Feature code 0862
    - − Available for zSeries (z800, z900)
- The crypty card is plugged into the Adjunct Processor
- Currently only RSA (asymmetric) is supported
  - ƒ Of benefit for Session initiation (SSL-Handshake)
- Also supported with
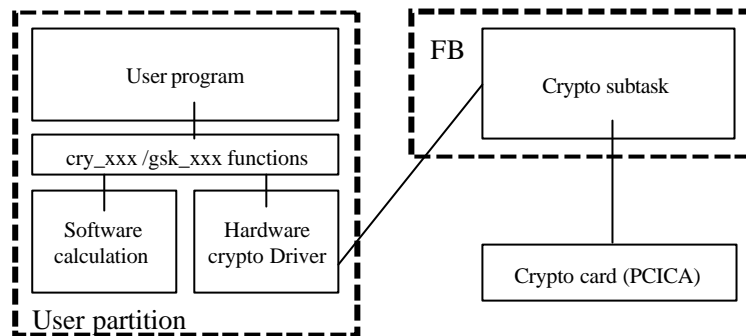 - ƒ z/VM 4.2 + APAR VM62905
  - ƒ z/VM 4.3

IBM.

Hardware Crypto Overview - continued

- New crypto subtask in Security Server (SECSERV) running in FB
  - ƒ Or as separate job if no SECSERV is running
  - ƒ Crypto card is polled by crypto task

```
User partition                              FB
┌─────────────────────────┐   ┌──────────────────────────┐
│  ┌───────────────────┐   │   │   ┌──────────────────┐   │
│  │   User program    │   │   │   │  Crypto subtask  │   │
│  └───────────────────┘   │   │   └──────────────────┘   │
│  ┌───────────────────┐   │   └──────────────────────────┘
│  │ cry_xxx /gsk_xxx  │   │
│  │     functions     │   │
│  └───────────────────┘   │
│  ┌─────────┐ ┌─────────┐ │            ┌──────────────────┐
│  │Software │ │Hardware │ │            │ Crypto card      │
│  │calcula- │ │crypto   │ │            │   (PCICA)        │
│  │tion     │ │Driver   │ │            └──────────────────┘
│  └─────────┘ └─────────┘ │
└─────────────────────────┘
```

Measurement Environment

- VSE/ESA 2.7 running on a z900 (2064-109)
  - ƒ on 1 processor (~2064-101)
  - ƒ with a PCI Cryptographic Accelerator
- Testcase programs on VSE
  - ƒ Crypto operations measurements
    - − calling cry_xxx functions (RSA, DES, SHA, MD5)
    - − each crypto operation is performed 10000 times
  - ƒ Secured data transfer (SSL)
    - − performs SSL handshake
    - − performs encrypted data transfer
    - − counterpart program running on Windows (SSL-client)
- All RSA operations are measured
  - ƒ with Hardware Crypto support
  - ƒ with Software Crypto
    - − (support already available with TCP/IP 1.4/1.5 as shipped in VSE/ESA 2.6)

Measurement Environment - continued

▪Variations

ƒ RSA encrypt/decrypt
- −512 / 1024 bit key

ƒ DES, DES CBC, 3DES CBC encrypt/decrypt
- −software crypto only
- −message length (128, 256, 512 bytes)

ƒ SHA Hash, MD5 Hash, SHA HMAC, MD5 HMAC
- −software crypto only
- −message length (128, 256, 512, 1K, 2K bytes)

ƒ SSL handshake/data transfer
- −01   RSA512_NULL_MD5
- −02   RSA512_NULL_SHA
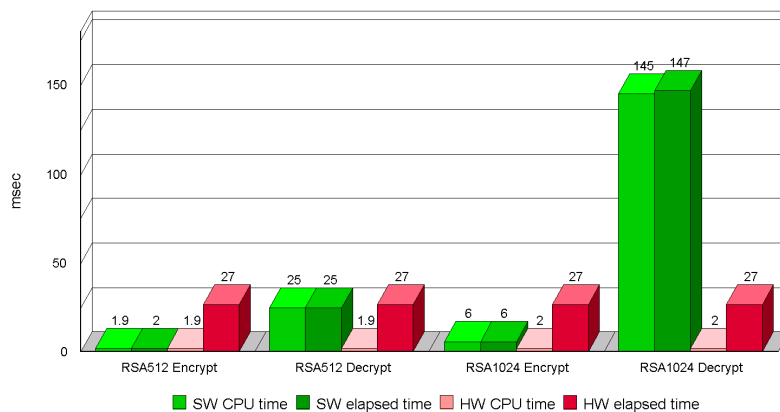- −08   RSA512_DES40CBC_SHA
- −09   RSA1024_DES_CBC_SHA
- −0A   RSA1024_3DES_EDE_CBC_SHA
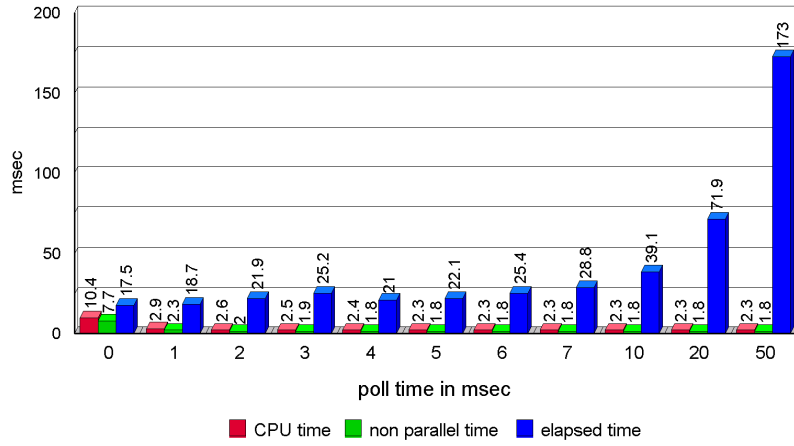

Measurements Results - RSA

HW Crypto:

- CPU time and elapsed time is independent of operation / key length
- RSA operation takes about 2 msec CPU time and 28 msec ellapsed time
- CPU time is always less than software crypto

17

Measurements Results - RSA polltime

msec (y-axis: 200, 150, 100, 50, 0)

Values by poll time:
- 0: 10.4, 7.7, 17.5
- 1: 2.9, 2.3, 18.7
- 2: 2.6, 2, 21.9
- 3: 2.5, 1.9, 25.2
- 4: 2.4, 1.8, 21
- 5: 2.3, 1.8, 22.1
- 6: 2.3, 1.8, 25.4
- 7: 2.3, 1.8, 28.8
- 10: 2.3, 1.8, 39.1
- 20: 2.3, 1.8, 71.9
- 50: 2.3, 1.8, 173

poll time in msec

■ CPU time   ■ non parallel time   ■ elapsed time

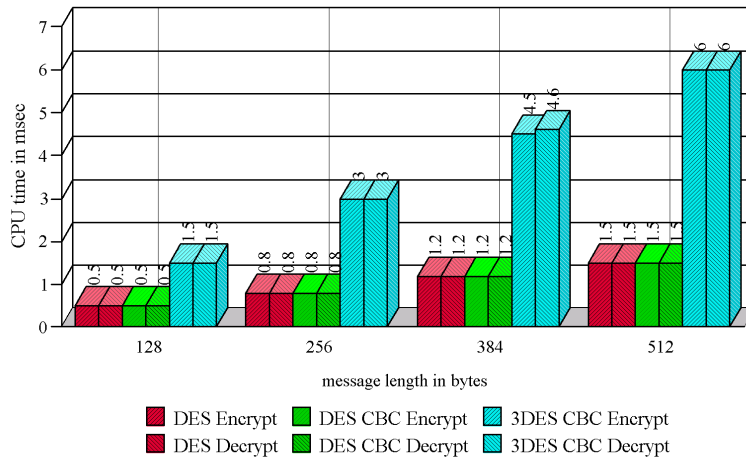Per default a polltime of 7 msec is used.

Can be changed with: MSG FB,DATA=WAITTIME=nn

Smaller values increases CPU time, higher values increases ellapsed time

Measurements Results - DES, DES CBC, 3DES CBC (symmetric)

CPU time in msec (y-axis: 7, 6, 5, 4, 3, 2, 1, 0)

Values by message length in bytes:
- 128: 0.5, 0.5, 0.5, 0.5, 1.5, 1.5
- 256: 0.8, 0.8, 0.8, 0.8, 3, 3
- 384: 1.2, 1.2, 1.2, 1.2, 4.5, 4.6
- 512: 1.5, 1.5, 1.5, 1.5, 6, 6

message length in bytes

■ DES Encrypt   ■ DES CBC Encrypt   ■ 3DES CBC Encrypt
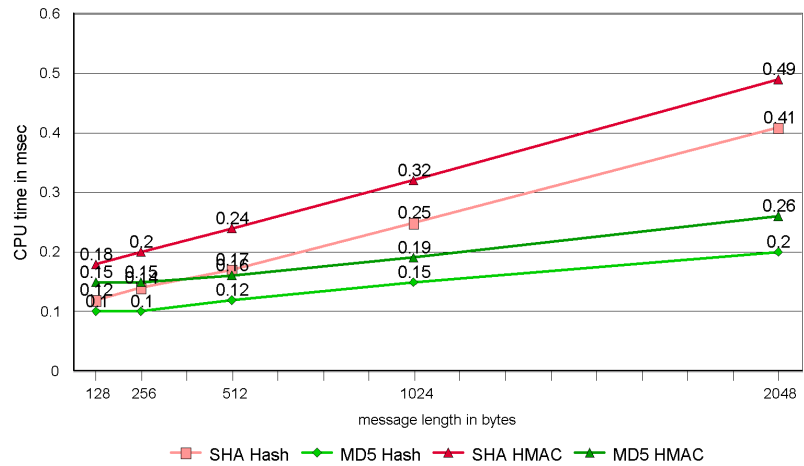■ DES Decrypt   ■ DES CBC Decrypt   ■ 3DES CBC Decrypt

Software Crypto only!

DES and DES CBC takes similar CPU times, 3DES CBC about 3.8 times
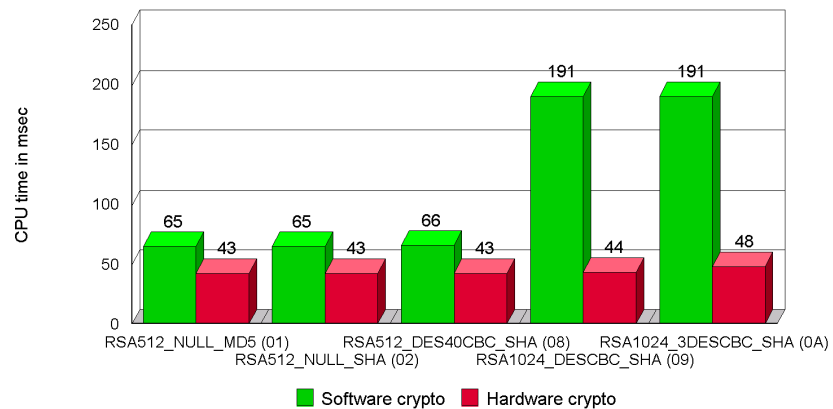
Measurements Results - SHA, MD5

SHA takes about 1.8 times more CPU time compared to MD5
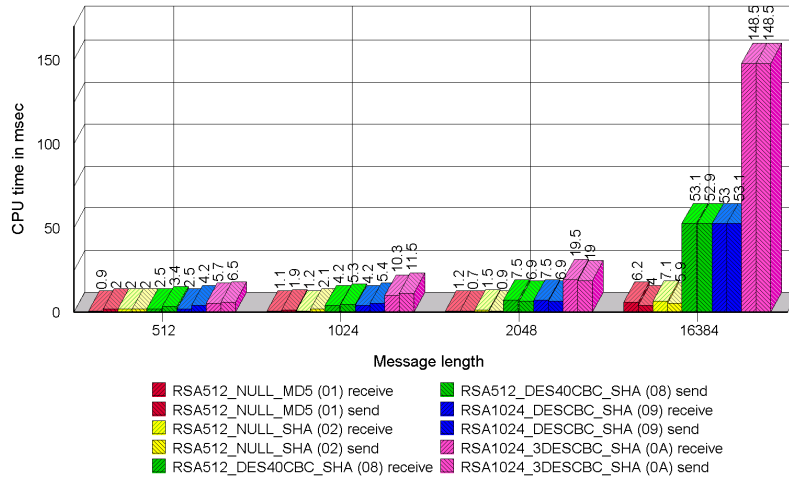Software Crypto only!

Measurements Results - SSL Handshake

HW Crypto:
- CPU time and elapsed time is independent of cipher suite used
- SSL handshake takes about 43-48 msec CPU time (connection establishment)

Slide 1

VS@ e-business

Measurements Results - SSL data transfer

CPU time in msec — Message length

Chart values (CPU time in msec):

512: 0.9, 2, 2, 2.5, 3.4, 2.5, 4.2, 5.7, 6.5

1024: 1.1, 1.9, 2, 2.1, 4.2, 5.3, 5.4, 10.3, 11.5

2048: 1.2, 0.7, 1.5, 0.9, 7.5, 6.9, 7.5, 6.9, 19.5, 19

16384: 6.2, 7, 7.1, 5.9, 53.1, 52.9, 53, 53.1, 148.5, 148.5

Legend:
- RSA512_NULL_MD5 (01) receive
- RSA512_NULL_MD5 (01) send
- RSA512_NULL_SHA (02) receive
- RSA512_NULL_SHA (02) send
- RSA512_DES40CBC_SHA (08) receive
- RSA512_DES40CBC_SHA (08) send
- RSA1024_DESCBC_SHA (09) receive
- RSA1024_DESCBC_SHA (09) send
- RSA1024_3DESCBC_SHA (0A) receive
- RSA1024_3DESCBC_SHA (0A) send

CPU time depends on used hashing (SHA/MD5) and encryption algorithm (DES/3DES)
Software Crypto only!

IBM.

Slide 2

VS@ e-business

SSL data transfer overhead

Non SSL

TCP/IP   Application   TCP/IP

SSL

TCP/IP   Decrypt   Application   Encrypt   TCP/IP
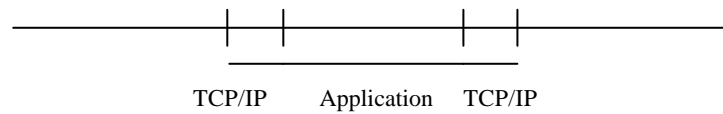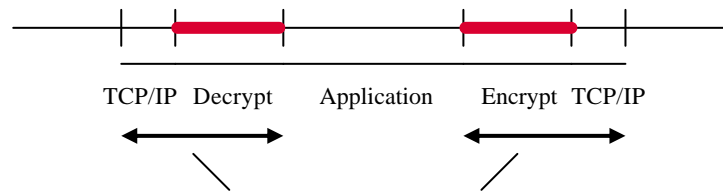
this has been measured

IBM.

VS@

e-business

Measurements Results - conclusion

- HW Crypto
  - ƒ Supports RSA operations only (e.g. used by SSL handshake)
  - ƒ CPU time/elapsed time is independent of operation and key length
  - ƒ Software RSA encryption is faster in terms of elapsed time (on large processors)
    - but hardware crypto saves CPU time
- SW Crypto
  - ƒ CPUtime /elapsed time is very dependent on CPU speed and utilization

IBM.

VS@

e-business

SSL Performance Recommendations

- Use SSL only if there is a need for
  - ƒ If at least one of the follwoing is required
    - Keeping secrets
    - Proving identity
    - Verifying information
- Cipher Suites 01 and 02 has less CPU-time consumption, but NO data encryption
  - ƒ RSA512_NULL_MD5, RSA512_NULL_SHA
- If data encryption is required
  - ƒ Use cipher suites 08, 09 or 0A
  - ƒ 08 uses 512 bit keys, others 1024
  - ƒ 1024 bit RSA keylength is recomended (from a security point of view)

IBM.

VS@

e-business

General Performance Hints for Connectors and TCP/IP

- Reduce amount of data transferred
  - ƒ Transfer only data that is needed
  - ƒ Issue only requests that are needed
- Use connection pooling
  - ƒ Reduce overhead of connection establishment
- Performance of connectors depends on
  - ƒ Network performance
  - ƒ Performance of "server"
  - ƒ Performance of "client" or middle tier

IBM.

VS@

e-business

Dependencies for VSE/ESA Growth

- System dependecies
  - ƒ Many control-blocks etc.. still below the line
  - ƒ VTAM IOBUF areas in System GETVIS-24
  - ƒ Non-Parallel-Share limits n-way support
  - ƒ Number of tasks
    - – Up to 255, 32 per partition, 208 subtasks in total
- Application dependencies
  - ƒ Integrated system concepts/functions
  - ƒ Functions/Applications dependencies
  - ƒ Number of users per TCP/IP partition

IBM.

VS@

e-business

Dependencies for VSE/ESA Growth - continued

- Not being considered to be a limit
  - ƒ Number of partitions
    - − 12 static + 150-200 dyn. partitions
  - ƒ Real storage (max. 2 GB)
  - ƒ Total virtual storage (max. 90 GB)
  - ƒ Total number of devices (3 digit CUU)
    - − Max. 1024 devices (and 16 channels)
  - ƒ Total number of logical units
    - − 255 per partition and 12x255=3060 in total
  - ƒ Label area
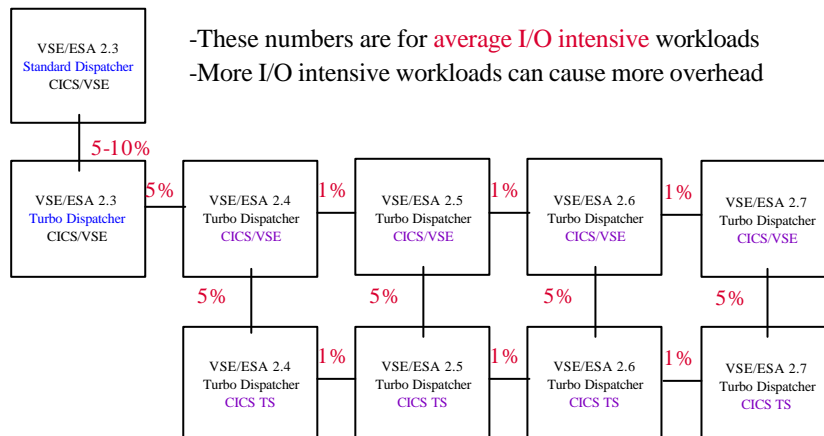   - − Max. about 9000 in total, and 712 in sub areas

IBM.

VS@

e-business

Overhead Deltas for VSE Releases

VSE/ESA 2.3 Standard Dispatcher CICS/VSE

-These numbers are for average I/O intensive workloads
-More I/O intensive workloads can cause more overhead

5-10%

VSE/ESA 2.3 Turbo Dispatcher CICS/VSE	5%	VSE/ESA 2.4 Turbo Dispatcher CICS/VSE	1%	VSE/ESA 2.5 Turbo Dispatcher CICS/VSE	1%	VSE/ESA 2.6 Turbo Dispatcher CICS/VSE	1%	VSE/ESA 2.7 Turbo Dispatcher CICS/VSE

5%   5%   5%   5%

VSE/ESA 2.4 Turbo Dispatcher CICS TS	1%	VSE/ESA 2.5 Turbo Dispatcher CICS TS	1%	VSE/ESA 2.6 Turbo Dispatcher CICS TS	1%	VSE/ESA 2.7 Turbo Dispatcher CICS TS

IBM.

Further Information

▪VSE Homepage:
http://www.ibm.com/servers/eserver/zseries/os/vse/
▪VSE Performance Homepage:
http://www.ibm.com/servers/eserver/zseries/os/
vse/library/vseperf.htm
▪Performance Documents from W. Kraemer
ƒ available on the Performance Homepage
▪VSE/ESA e-business Connectors User's Guide
http://www.ibm.com/servers/eserver/zseries/os/
vse/pdf/ieswue20.pdf

Questions

?