FLEX·*ES*® Tips & Tricks

Presented by Gary Eheman
Fundamental Software, Inc.
WAVV, April 2003

1

Trademarks

The following are trademarks of Fundamental Software, Inc.:
FLEX·*ES*®, FakeTape, Fundamental Software®

The following are trademarks of IBM Corporation:
S/390®, z/Architecture, z/OS, z/VM, VM/ESA®, VSE/ESA

Intel® and Pentium® are trademarks of Intel Corporation

UNIX® and UnixWare® are trademarks of The Open Group

Linux® is a trademark of Linus Torvalds

2

Agenda

- Discuss various topics of interest in optimizing your FLEX▪*ES* experience

- Some performance-related topics

- Some exploitation topics

3

Processors

- Dedicate CPUs to system instance if possible
  - cpu(0) dedicated # specification in config
- When is it not possible?
  - When you need to share the CPU with another FLEX▪*ES* instance
  - When you have only one (laptop)
  - Combination dedicated/undedicated for an n-way single instance should be avoided where possible

4

VM and processors

- VM uses active wait (spins in storage key 3) by default. Runs S/390 CPU (and thus Pentium) at 100% busy by looking for work if otherwise idle

- In general, this is OK for dedicated CPUs

- If sharing the CPU with another instance, use "feature lpar" to cause VM to not use active wait

  - VM does not use active wait if it believes it is running in an LPAR

5

Host paging/swapping is bad

- In general, host (UNIX/Linux) paging can be bad for performance of FLEX▪*ES*

- No big surprise

- Avoid it by not over-committing server memory

- Add more memory if you need more

6

Processor cache size

- Specified in FLEX▪*ES* configuration file
- cachesize(nnnn) is kilobytes of S/390 processor cache
  - nnnn must be an integer power of 2 greater than 64 (1024, 2048, 4096...)
  - Value is per CPU defined to the instance
  - Fulfilled from host server memory
  - Larger sizes can cause server paging (bad)
- Megabytes used per S/390 processor = nnnn * 11/1024
- Example: cachesize (2048) and two S/390 processors
  2 (CPUs) * (2048 * 11/1024) = 44MB
- Cachesize is a big knob if you turn it. You can quickly over-commit memory and start paging if not careful

7

Processor cache misses

- Rules of thumb (ROT) are dangerous. Use with caution

- Try to keep cache misses less than or equal to 4%

- FLEX▪*ES* "display cachestats" command gives statistics
  - flexes> clear cachestats
  - flexes> display cachestats
    ...<other stats snipped>
    Cache misses: 66687978/1122334222 **6%**

  These counters wrap quickly. Clear, then display soon after

8

FLEX▪*ES* DASD caches

- Two levels may be specified in FLEX▪*ES* configuration:
  - Control unit floating pool (optional specification)
    - Specified in 'option' statement for control unit
  - Dedicated to device (subset of C.U. specification)
    - Specified in 'devopt' device options for device
- Default is one cylinder of cache per DASD device. That is 15 tracks for 3390, 3380, 9345
- You can specify more (or less) on a per-device basis. Tune for your known "big hitters"

9

DASD cache effectiveness

- Use "display ckdcachestats *devaddr*" command to display effectiveness

```
flexes> clear ckdcachestats 100
flexes> d ckdcachestats 100
ADDRESS    READS WRITES    CACHE HITS  DEDICATED LINES    LINES USED
   0100        0      0      0 (  0%)               15    15 ( 14%)
   0101        0      0      0 (  0%)               15    15 ( 14%)
flexes> d ckdcachestats 100
ADDRESS    READS  WRITES    CACHE HITS DEDICATED LINES    LINES USED
   0100     4373       0  3931 ( 90%)               15    15 ( 14%)
   0101        0       0      0 (  0%)               15    15 ( 14%)
```

10

Cache control commands

- O.S. cache commands to affect controller cache (DFW, CFW) are ineffective

- But commands to change the status of the cache (on, off, DFW, CFW, etc.) report success if issued

- You *must* tune the DASD cache using the FLEX▪*ES* resource config file
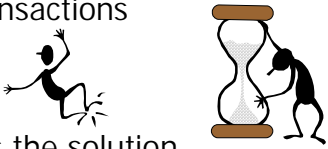
11

PC memory is "cheap"

- For heavy I/O workloads, track caches can dramatically improve I/O response time

- Add more memory to your initial config in anticipation of DASD cache needs

- Memory is inexpensive!
  Be generous to your DASD cache

12

A side effect of fast I/O

- Many VSE customers see their CICS transactions dominate the CPU after first migrating to FLEX▪*ES*
- Because FLEX▪*ES* I/O can be so fast, I/O wait time is practically eliminated for short transactions
- CICS users happy
- Other users frustrated
- Retuning VSE partition priorities is the solution
- This is a "good" problem to have, but **be prepared**
- Many customers find OPTI-WORKLOAD from Barnard Software helpful (www.bsiopti.com/workload.html)

13

Installing Linux for S/390 or z/Series

- Linux for S/390 and z/Series may have difficulty IPLing without 2003-03-17 accumulated patches from IBM Developerworks that fixes the kernel error
    - **Description:** DASD devices will not be set 'online' on slow/fast systems
    Symptom: Root volume does not go online
    Problem: Bottom half raises DASD state before main path
    Solution: Extend scope of irq_lock to include change of state
    Problem-ID: 1919
- It is a Linux problem, not a FLEX▪*ES* problem
- http://www.ibm.com/developerworks/oss/linux390/whatsnew.shtml
- Presents a challenge on new installs of broken Linux kernel code

14

Writethrough *vs* writeback

- Default for FLEX▪*ES* track caches is *writeback.* S/390 gets notification of write complete when it hits track cache but before written to disk
- Track caches are flushed every 5 to 10 seconds
- The importance of a good UPS remains true
- *Writethroughcache* can be specified per device if desired using devopt in config file
- Some volumes are always excellent for writeback
  - PAGE, sort work, T-DISK, etc.

15

Expanded storage

- FLEX▪*ES* supports XSTORE
- Implemented in server memory
- For VM systems, a little XSTORE can still be helpful in the VM paging hierarchy
- For VSE, exploitation of XSTORE requires vendor software
- More "data in memory"

16

S/390 "system hardware console"

- Emulation support for "system hardware console"
- Line mode device accessible from S/390 operating systems via HMC conventional hardware
- Not a 3270
- Emulated via FLEX▪*ES* command line interface
- Added convenience for "single point of control" for controlling emulated S/390 complex
- Authorized users who can use the CLI can use the "hwc" command to communicate with S/390 O.S. that supports this device

17

S/390 hwc example

- For VM, add "System_console" to list of "Operator_consoles" in VM's SYSTEM CONFIG file if you want it as possible IPL console

```
flexes> ipl 100
14:40:12 VM/ENTERPRISE SYSTEMS ARCHITECTURE V2 R4.0 SERVICE LEVEL ...
14:40:12 SYSTEM NUCLEUS CREATED ON ...
<snip>
flexes> hwc query userid
14:40:52 OPERATOR AT FSIVM
Ready; T=0.01/0.01 14:40:52
hwc #cp query operator
14:41:15 CP QUERY OPERATOR
14:41:15  OPERATOR-SYSC
```

18

S/390 hwc and VSE

- VSE can use hwc as IPL console. Use "I" in first position of loadparm on IPL
- flexes> ipl 140 I......
- Use of a 3270 is recommended over hwc for day-to-day operations

Issue the flexescli command "clear messages" occasionally if using the hwc function. A history of all messages typed to the CLI is kept in memory until cleared.

19

More VM exploitation

The OS2/AIX EXECs, OS2 PIPELINE stage, and MOUNT EXEC/MODULE are licensed with IBM's preconfigured VM/ESA and z/VM V3 system CDROMs that are a feature code of those products. They are not found on base VM system install media.

20

Communicating commands

- Issue UNIX/Linux commands and capture results from CMS on VM
- Method uses the same technology as P/390-based systems. Familiar if you migrate from that environment. New ground for others to enjoy
- Two methods can be used:
  - OS2 EXEC (OS2CMD MODULE)
  - OS2 stage of CMS PIPELINES
- Your bullets. Your feet. Aim carefully!

21

Command communications requirements

- Define communications devices in FLEX▪*ES* config
  - cu devad(0x460) path(0) resource cupcserv
    ...
    cupcserv: cu 3274
      interface local(1)
      device(00) 3278 OFFLINE ***devopt 'allowhostccws'***
      device(01) 3278 OFFLINE ***devopt 'allowhostccws'***
    end cupcserv
- Define same in VM SYSTEM CONFIG
  - RDEVICE 0460-0461 Type Unsupported Devclass 3270_Display

- Modify OS2 EXEC to eliminate CPU type check. Rename to a new name if desired (e.g., HOSTCMD EXEC)

22

Command communication examples

- Issue CMS hostcmd EXEC to invoke commands

  ```
  exec hostcmd ls –ltr /usr/flexes/rundir/*.conf
  DEV 0460 ATTACHED TO OPERATOR 0460 BY OPERATOR
  -rwxr-x---  1 flexes flexes  1790 Jan 30 14:23 gary.conf
  DEV 0460 DETACHED OPERATOR 0460 BY OPERATOR
  Ready;
  ```

- Or use OS2 stage of CMS PIPELINES

  ```
  cp attach 460 * 460
  cp set 370accom on
  pipe os2 cmd ls –ltr /usr/flexes/bin | locate /3480/ |  > 3480 files a
  ```

23

Combining power

- Create a shell script
  - cat garystats.sh
    :
    echo 'd cachestats' | flexescli localhost vmtest
- Invoke it from CMS
  - **pipe os2 cmd sh garystats.sh | locate /Cache misses/ | cons**
    Cache misses: 159702/8452700 2%
    Ready;
- The OS2FILE stage of CMS PIPES for file I/O is not supported in FLEX▪*ES*

24

FakeTape

- Tape emulation using disk files
- Multiple formats supported on any FakeTape emulated drive you define
  - FSI's FakeTape format (default)
  - IBM's AWSTAPE format (from P/390)
  - IBM's OMA/2 format (read only)
    - VM and VM service available on CD

25

FakeTape exploitation ideas

- Nightly backups from S/390 to FakeTape, then move those files to another server

- Burn FakeTape files to CD for archival purposes

- Suite of utilities to help you exploit FakeTape
  - faketasc, asctfake, faketbin, bintfake, fakeflat, scsitfake, faketscsi, fakedump, initawstape

26

Mount FakeTape files on tape drives from CMS

- Two ways to mount the file on the tape drive
  - Using the flexescli
    - flexes> mount 500 gary.goodstuff.tape
    - This method requires authorization/access to the CLI
  - Using the CMS MOUNT EXEC/MODULE for P/390 systems
    - Requires *devopt 'allowmountccws'* on emulated drive definition
- Example CMS method:
  - EXEC MOUNT 181 gary.goodstuff.tape

27

Mount FakeTape files on tape drives from VSE

- A program to mount FAKETAPE files on tape drives from VSE batch is available from Barnard Software, Inc.
- http://www.bsiopti.com
  - Click "Downloads"
  - Fill in form for access
  - Download BSIMOUNT after filling in form
- FLEX▪*ES* 6.1.15 or higher required for BSIMOUNT use with VSE on FLEX▪*ES*

28

Internet discussion list

- There is an Internet discussion list (listserv) hosted by UGA for technical discussions
- Over 400 subscribers/participants
- *NOT* a part of any support structure
- Good place to ask technical how-to questions
- Some instructions for getting started can be found at http://support.funsoft.com then click on "E-mail forum"

29

Thank you!

⇒ Come see us in Booth 32

⇒ Visit us online

http://support.funsoft.com

30