

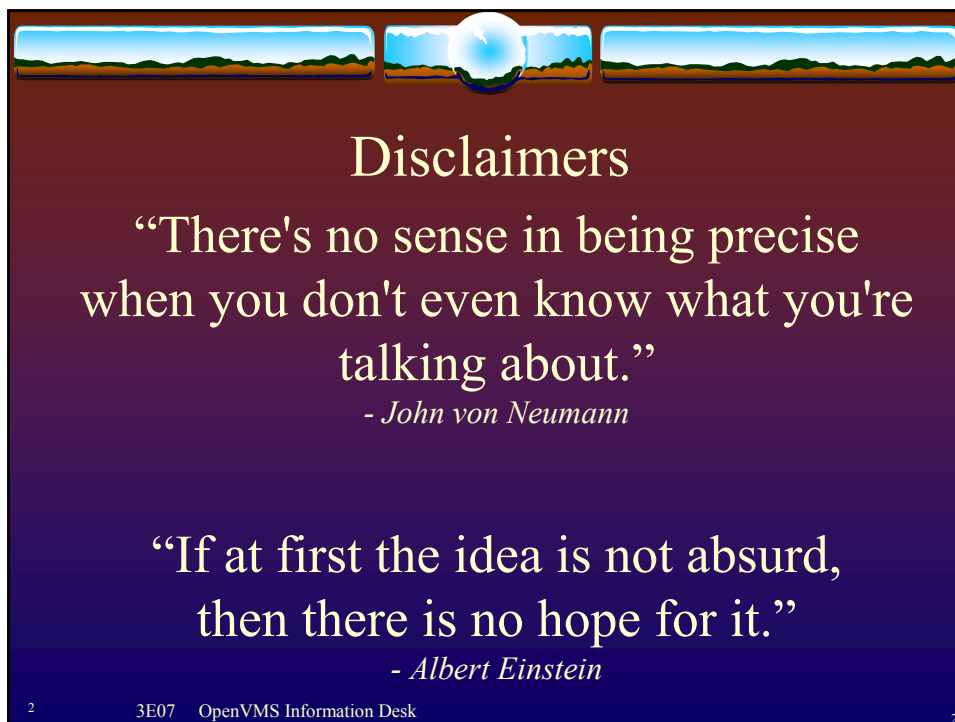
OpenVMS  
Information Desk

*The Sequel*

“Free Advice Worth Every Cent”

Thilo Lauer  
Hewlett Packard GmbH

Manfred Kaser  
Hewlett Packard GmbH



Disclaimers

“There's no sense in being precise  
when you don't even know what you're  
talking about.”  
*- John von Neumann*

“If at first the idea is not absurd,  
then there is no hope for it.”  
*- Albert Einstein*

2 3E07 OpenVMS Information Desk 2




## Our Golden Rules

The best performing code is the code not being executed

The fastest I/Os are those avoided

Idle Cores are the fastest Cores

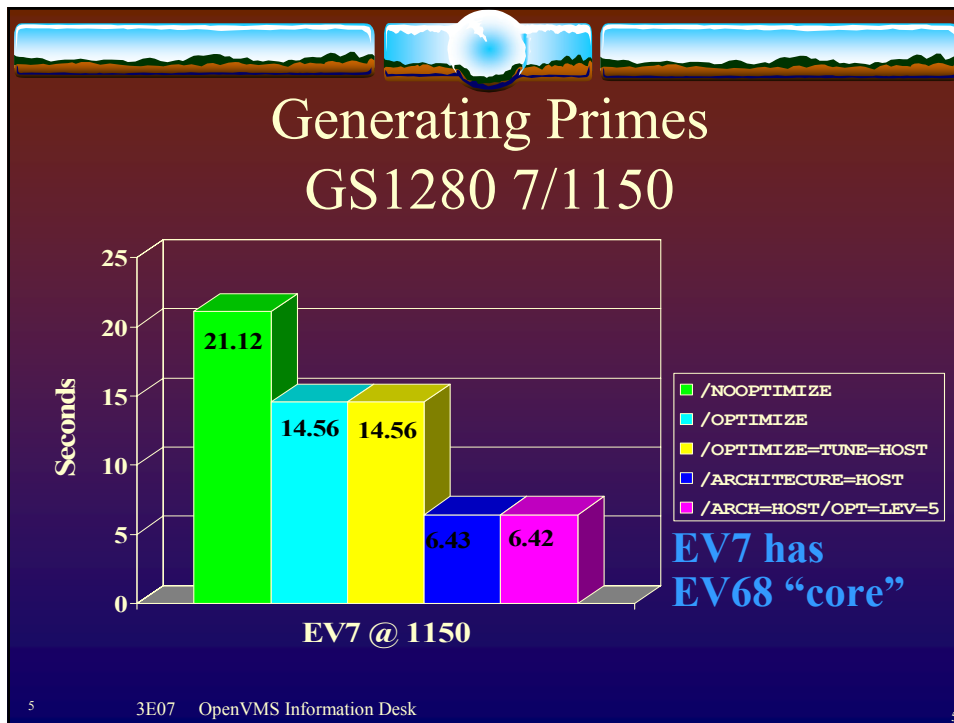
3 3E07 OpenVMS Information Desk 3



## Examples of ...TUNE & /ARCHITECTURE

- ❖ **/OPTIMIZE=TUNE=EV56**
  - ❖ Execute on all Alpha generations
  - ❖ Biased towards EV56
- ❖ **/OPTIMIZE=TUNE=EV6 /ARCHITECTURE=EV56**
  - ❖ Execute on EV56 and later (Byte/Word instructions)
  - ❖ Biased for EV6 (quad issue)
- ❖ **/ARCHITECTURE=EV6**
  - ❖ Execute on EV6 and later (Integer-Floating conversion, Byte/Word & Quad-issue scheduling)
- ❖ **/ARCHITECTURE=HOST**
  - ❖ Code intended to run on processors the same type as host computer
  - ❖ Execute on that processor type and higher

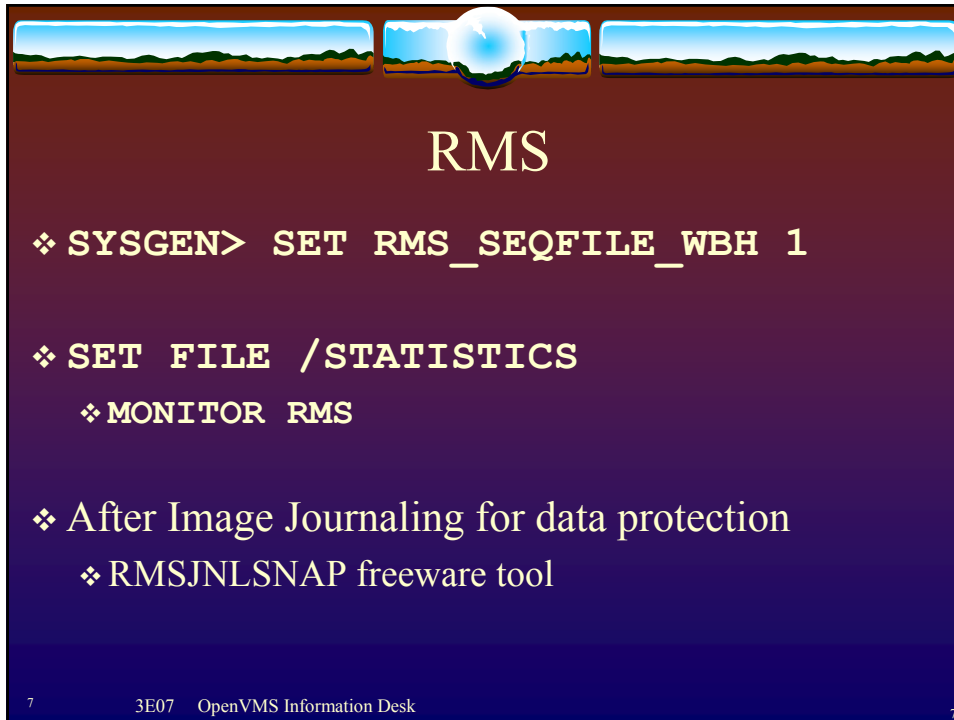
4 3E07 OpenVMS Information Desk 4



## RMS

- ❖ Use larger buffers & more of them
- ❖ FAB/RAB parameters:
  - ❖ ASY, RAH, WBH, DFW, SQO
  - ❖ ALQ & DEQ
  - ❖ MBC & MBF
  - ❖ NOSHR, NQL, NLK
- ❖ SET RMS ...
  - ❖ /SYSTEM
  - ❖ /BUFFER\_COUNT=n
  - ❖ /BLOCK\_COUNT=n

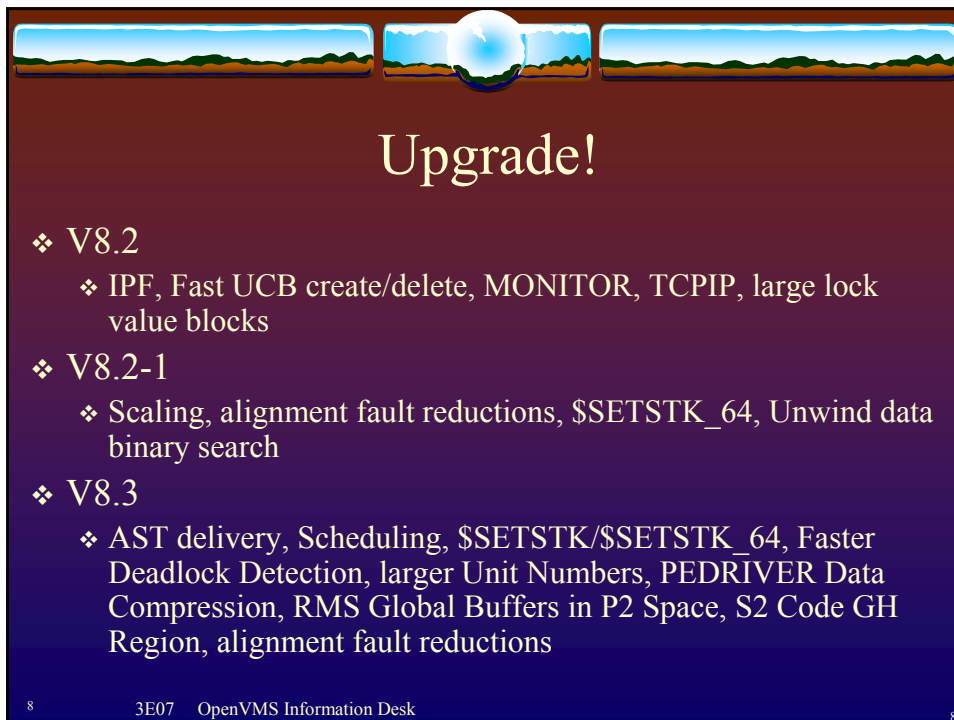
6 3E07 OpenVMS Information Desk 6



## RMS

- ❖ `SYSGEN> SET RMS_SEQFILE_WBH 1`
- ❖ `SET FILE /STATISTICS`
  - ❖ `MONITOR RMS`
- ❖ After Image Journaling for data protection
  - ❖ RMSJNLSNAP freeware tool


7 3E07 OpenVMS Information Desk 7



## Upgrade!

- ❖ V8.2
  - ❖ IPF, Fast UCB create/delete, MONITOR, TCPIP, large lock value blocks
- ❖ V8.2-1
  - ❖ Scaling, alignment fault reductions, \$SETSTK\_64, Unwind data binary search
- ❖ V8.3
  - ❖ AST delivery, Scheduling, \$SETSTK/\$SETSTK\_64, Faster Deadlock Detection, larger Unit Numbers, PEDRIVER Data Compression, RMS Global Buffers in P2 Space, S2 Code GH Region, alignment fault reductions

8 3E07 OpenVMS Information Desk 8




“For an introductory-level 1-semester C course, I'd teach that programmers do 2 things:

1. write new code
2. and maintain old code.

The percentage of your time doing #2 depends greatly on how well you do #1.”

– *Mark Schafer June 22<sup>nd</sup> 2000 12:07*

9 3E07 OpenVMS Information Desk 9



## Initializing Structures


### Which is fastest/efficient?

```
void foo_1() {
    char array[512]={0};
    printf("array=%x", &array);
} 2.23 s
```

```
void foo_2() {
    char array[512];
    for (int i=0; i<512; i++) array[i]=0;
    printf("array=%x", &array);
} 2.92 s
```

```
void foo_3() {
    char array[512];
    memset (array, 0, sizeof(array));
    printf("array=%x", &array);
} 1.61 s
```

10 3E07 OpenVMS Information Desk 10



## setjmp...

```
#include "lib$routines.h"
#include "stdio.h"
#include "setjmp.h"


main(char **av, int ac) {
  int i, env;
  jmp_buf jmpbuf;

  lib$init_timer();

  for (i = 0; i < 1000000; i++)
    env = setjmp(jmpbuf);

  lib$show_timer();
}
```

11 3E07 OpenVMS Information Desk 11



## ...setjmp

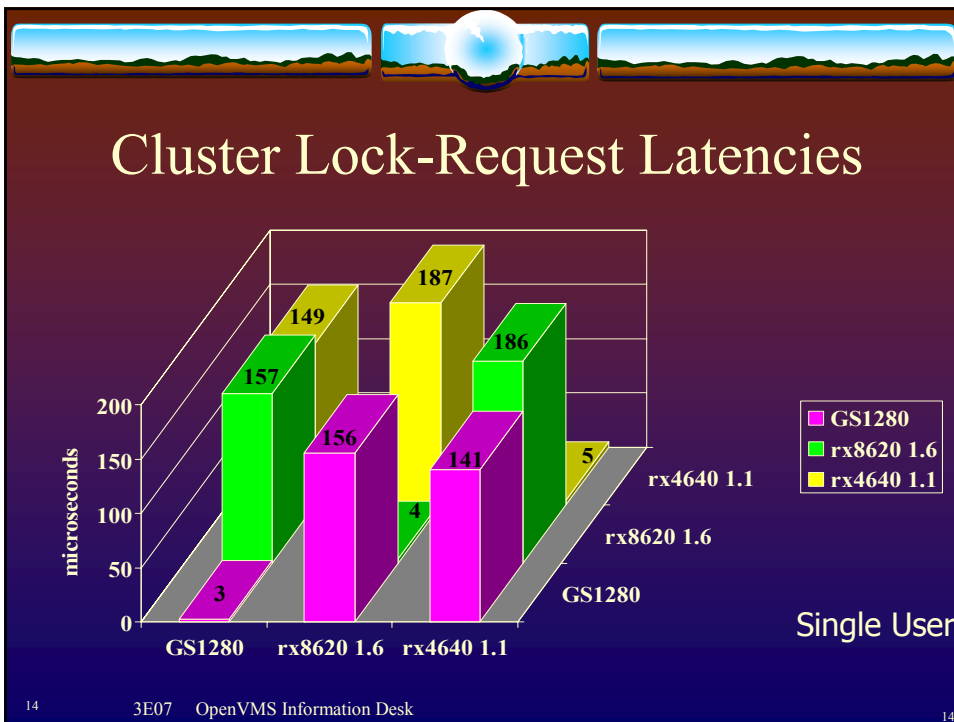
- ❖ Takes 4.3 seconds to execute this program on DS20 (500 MHz)
- ❖ Compiled with `/DEFINE=__FAST_SETJMP` program takes only 0.04 seconds


12 3E07 OpenVMS Information Desk 12

## Performance and Coverage Analyzer

- ❖ Use PCA with your applications!
- ❖ Find where time is being spent – focus first on those areas
- ❖ Identify I/O, System Service, Alignment Faults, PC sampling, etc.

13 3E07 OpenVMS Information Desk 13






## IO vs CPU

- ❖ Advertised:
  - ❖ “OpteronX @ 2GHz”
  - ❖ “64-bit PCI-X @33Mhz”
- ❖ I/O performance is combination of I/O bus type (PCI, PCI-X, etc.), bus speed, bus data path and/or command width, etc.
- ❖ Many times perception that system is "running slow" is more function of I/O contention than CPU overload

15 3E07 OpenVMS Information Desk 15




## EVA/XP Storage

- ❖ Initialize disks with cluster size multiple of 4
  - ❖ OpenVMS Engineering suggests 32
- ❖ Perform sequential write I/O on RAID5 groups...
  - ❖ Multiple of 4 block transfers
  - ❖ Starting on multiple of 4 block VBN
  - ❖ COPY/BLOCK\_SIZE (V8.2, VMS732\_COPY-vxxx)
  - ❖ Avoid excessive async sequential access I/O queues
    - ❖ Throttle your IO load

16 3E07 OpenVMS Information Desk 16






## EVA/XP Storage

- ❖ Assume worst case of 100 I/O per second per physical drive for random IO (assume 0 cache hits)
- ❖ Balance paths
  - ❖ Keep path to a given LUN the same across cluster
- ❖ Blocks are not allocated/initialize until first write


17      3E07    OpenVMS Information Desk      17



## XP storage

- ❖ Best if 8 I/Os per LUN are presented by host
- ❖ OpenVMS methods that can help
  - ❖ BACKUP
    - ❖ Lower values for DIOLM and PQL\_MDIOLM
    - ❖ Redesigned to work with modern controllers
      - ❖ VMS732\_BACKUP\_V0500 (/IO\_LOAD)
  - ❖ WWID throttle IO descriptor to limit the total number of I/Os per FC port
    - ❖ V7.3-2 FIBRE\_SCSI-V0400 and later
    - ❖ `SDA> FC SET WTID /WWID=target_wwid /CAP=cap_value`


18      3E07    OpenVMS Information Desk      18



## MSCP Disk Serving

- ❖ Alpha & I64 MSCP server does not do dynamic balancing
  - ❖ `SET PREFERRED /HOST=<node>/FORCE <dev>`
- ❖ `MSCP_LOAD` in 'mixed size' cluster
  - ❖ 2 for "small" Alpha/IPF servers
  - ❖ 1 for "big" ones (uses default of 340 on Alpha & IPF)
- ❖ `MSCP_CREDITS`  $\geq 64$  for busy/big servers
- ❖ `MSCP_BUFFER`  $\geq 2048$ 
  - ❖ `127 * MSCP_CREDITS` when using host-based shadowing


19 3E07 OpenVMS Information Desk 19



## TCP/IP & DECnet

- ❖ TCP/IP V5.4 or later
  - ❖ Scalable Kernel  
(logical name `TCPIP$STARTUP_CPU_IMAGES`)
  - ❖ Default as of TCPIP V5.5
- ❖ `SET RMS /SYSTEM /NETWORK = 127`

20 3E07 OpenVMS Information Desk 20




## TCPIP

- ❖ TCPIP PPE (packet processing engine)
  - ❖ Available in TCPIP V5.x
  - ❖ Similar to dedicated LCKMGR

```
$ sysconfig -r inet inetcpu_cpuid=n
$ sysconfig -r inet inetcpu_enable=1
$ show system/process=tcpip$inetcpu
```

21 3E07 OpenVMS Information Desk 21




## Virtual Terminals

- ❖ Avoid process deletion at network disconnect (PC crash?)

Add to system startup :

```
$ ! ENABLE VIRTUAL TERMINALS
$ MCR SYSMAN IO CONNECT /NOADAPT VTA0 -
  /DRIVER=SYS$LOADABLE_IMAGES:SYS$TTDRIVER
$ DEFINE/SYSTEM/EXECUTIVE TCPIP$TELNET_VTA TRUE
```


22 3E07 OpenVMS Information Desk 22



## Fibre Channel & Fastpath

- ❖ V8.3
  - ❖ Removal of IOLOCK8 spinlock usage for fibre channel drivers
- ❖ Previously
  - ❖ Fastpath allows concurrency during I/O initiate
  - ❖ Distributed interrupts allows concurrency during I/O complete
  - ❖ However, ISR (interrupt service routine) takes global IOLOCK8... Yikes...
  - ❖ Workaround: assign FGx adapters to same fastpath CPU

23 3E07 OpenVMS Information Desk 23



## SHOW FASTPATH


```

Ryerox> show fastpath
Fast Path preferred CPUs on RYEROX 19-APR-2006 14:29:42.81
hp AlphaServer GS1280 7/1150 with 16 CPUs

Device:                Fastpath CPU:
EWA0                    1
EIA0                    1
EIB0                    1
EWB0                    8
FGA0                    1
FGB0                    8
PEA0                    2
PKA0                    1
PKB0                    1
PKC0                    1

OpenVMS TCP/IP is currently running on CPU 3
OpenVMS Lock Manager is currently running on CPU 4
Ryerox>
  
```

24 3E07 OpenVMS Information Desk 24



## POOL

- ❖ **NPAG\_GENTLE=NPAG\_AGGRESSIVE=100**  
to disable pool reclamation – Current VMS default
- ❖ Leave **NPAG\_GENTLE** and **NPAG\_AGGRESSIVE**  
out of MODPARAMS to benefit from future  
changes


25 3E07 OpenVMS Information Desk 25



## Application Temporary Files

- ❖ Frequently create/delete small temp. files?
  - ❖ Consider caching in virtual memory instead
  - ❖ “Spill” to disk file if needed after some threshold  
(1MB?)
- ❖ Don't be afraid of P2 virtual address space
  - ❖ Keep an eye out for excessive page faulting


26 3E07 OpenVMS Information Desk 26



## Large Sequential Files

- ❖ Rarely read?
  - ❖ Create in Directory marked **/CACHE=NOCACHE**
- ❖ Perhaps for...
  - ❖ Backup Savesets, log files, .LIS/.MAP files, etc.
- ❖ Avoids polluting XFC cache

27 3E07 OpenVMS Information Desk 27




## File With Incorrect LRL?

- ❖ **ANALYZE /RMS /UPDATE\_HEADER**
  - ❖ Used with **/CHECK** or **/STATISTICS**
  - ❖ Determines & sets LRL
  - ❖ Determines & sets file length hints (on ODS-5 volumes)

```

Maximum Record Size: 255
Longest Record: 82
:
File Length Hint (Record Count): 98
File Length Hint (Data Byte Count): 3243
:
:
  
```


28 3E07 OpenVMS Information Desk 28



## Global Sections

- ❖ Memory resident
  - ❖ Shared page tables
  - ❖ Granularity hints (when registered)
- ❖ P2 virtual address space
- ❖ Per-RAD sections on Wildfire

29 3E07 OpenVMS Information Desk 29



## Granularity Hint Regions

- ❖ Use less CPU translation buffer entries
  - ❖ Each maps many pages; reduces TB misses
- ❖ Resident images & global sections with reserved memory

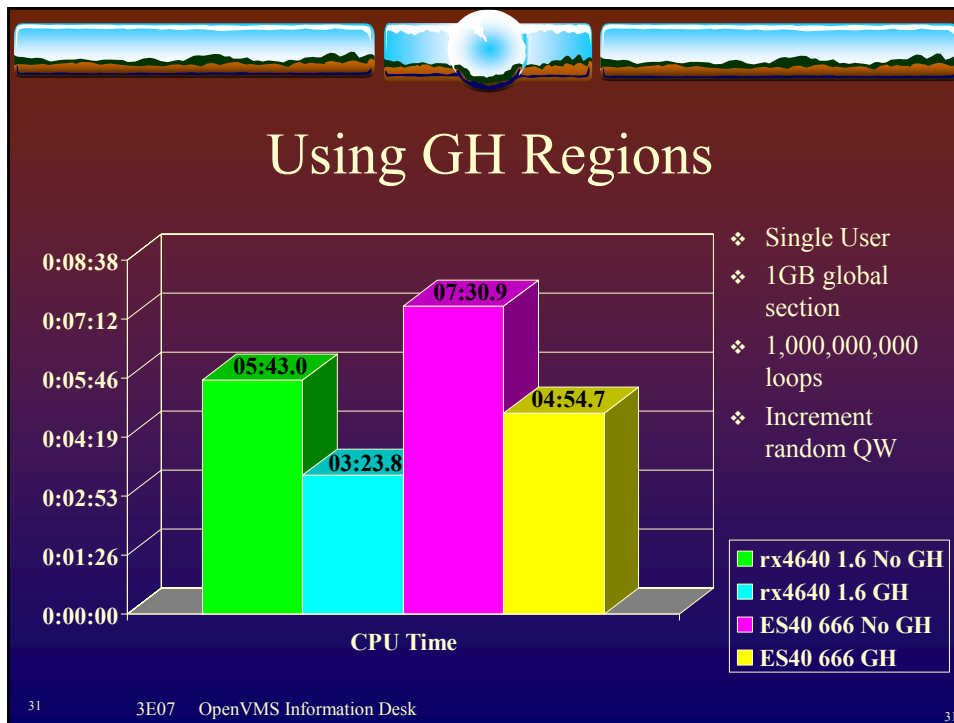
```

Wells TNA27:> MCR SYSMAN RESERVED_MEMORY ADD NJL$SHARED_MEMORY -
              /PAGE_TABLES /SIZE=1100 /ALLOCATE

Wells TNA3:> SHOW MEMORY /RESERVE
Memory Reservations (pages): Group      Reserved   In Use     Type
NJL$SHARED_MEMORY           SYSGBL      138        0   Page Table
NJL$SHARED_MEMORY           SYSGBL    131072     0   Allocated
NJL$SHARED_MEMORY           SYSGBL     8192      0   Allocated
NJL$SHARED_MEMORY           SYSGBL     1536      0   Allocated
Total (1.07 GBytes reserved)                140938     0

```

30 3E07 OpenVMS Information Desk 30




## XFC

- ❖ It isn't 1980 any longer...
  - ❖ Historically I/O sizes maxed at 127 blocks.
  - ❖ Today, utilities are doing I/O up to 256 blocks at a time
- ❖ Consider setting `VCC_MAX_IO_SIZE` to 256
- ❖ `MCR SYSMAN RESERVED MEMORY ADD VCC$MIN_CACHE_SIZE /SIZE=xxx /ALLOCATE /NOGLOBAL /NOZERO`

32 3E07 OpenVMS Information Desk 32






DECram

- ❖ Create virtual disk from system memory
- ❖ When temp/work files can not be avoided
- ❖ Integrated with VMS V8.2
- ❖ May be shadowed with physical disk
  - ❖ Shadowing smart enough to read from memory


33 3E07 OpenVMS Information Desk 33



Software RAID

- ❖ Bind local disks into RAID (0 or 5) sets
- ❖ “Magically” distribute I/O load among spindles
- ❖ Partition RAID arrays into logical units
- ❖ Small CPU overhead vs. I/O distribution
- ❖ Or...Use hardware controllers

34 3E07 OpenVMS Information Desk 34



## CRC

- Significant performance enhancements
  - LIB\$CRC
  - CRC macro instruction


```

$ r crc2
500 buffers of size = 32768
lib$crc latency 228.6628 msec
Total bytes processed = 16384000
Rate = 68.3321 Mbytes/sec

$ r crc2
500 buffers of size = 32768
lib$crc latency 152.2836 msec
Total bytes processed = 16384000
Rate = 102.6046 Mbytes/sec

```

35 3E07 OpenVMS Information Desk 35



## Disk Volumes

- ❖ SET VOLUME
  - ❖ /NOHIGHWATER
  - ❖ /EXTEND=big?
  - ❖ /CLUSTER=<multiple-of-4-or-16>
  - ❖ /LIMIT

36 3E07 OpenVMS Information Desk 36




## Find & Defrag

```
$ DEFINE DFU$NOSMG TRUE
$ MCR DFU SEARCH/FRAG=MIN=8 DEVICE

$ MCR DFU SEARCH /FRAG=MIN=8 /OUT=DFU.TMP
$ MCR DFU DEFRAG @DFU.TMP
```


37 3E07 OpenVMS Information Desk 37



## Analyze High MPSYNC Time

```
sda> spl start trace/buff=5000
.
.
sda> spl stop trace
sda> spl analyze/usa=hold=1
```


38 3E07 OpenVMS Information Desk 38



## Analyze High Locking Rate

```
sda> lck show active ! which files, volumes  
or  
sda> rdb show active ! which Rdb db's  
or  
sda> lck start trace ! which processes  
sda> lck start collect/process  
.  
.  
sda> lck show collect
```


39 3E07 OpenVMS Information Desk 39



## Analyze High IO Rate

```
sda> io start trace  
sda> io start collect/device  
or  
sda> io start collect/process  
.  
.  
sda> io show collect /full
```

40 3E07 OpenVMS Information Desk 40




## Erasing Disks

- ❖ Before erasing data, read “*Secure Deletion of Data from Magnetic and Solid-State Memory*” by Peter Gutmann, Department of Computer Science, University of Auckland

[http://www.cs.auckland.ac.nz/~pgut001/pubs/secure\\_del.html](http://www.cs.auckland.ac.nz/~pgut001/pubs/secure_del.html)

41 3E07 OpenVMS Information Desk 41




## Disk Erasure

- ❖ Easiest way to erase user data on disk
  - ❖ Write pattern(s) to every bit in every block
  - ❖ Repeat until you no longer feel paranoid

```
$ ANALYZE /MEDIA <dev>: -  
/EXERCISE = PATTERN = -  
%XA5A5A5A5A5
```


42 3E07 OpenVMS Information Desk 42



## Data Encryption

- ❖ VMS Encryption kit ships with VMS V8.2
  - ❖ V8.3 adds additional algorithm
    - ❖ Encrypt integrated into the base O/S
- ❖ **BACKUP /ENCRYPT**
  - ❖ Encryption increases CPU utilization ! Duh? You'd be surprised....
- ❖ Roll your own encryption functions

43 3E07 OpenVMS Information Desk 43




## BACKUP Performance?

- ❖ Focus on **\*total\*** **restore & recovery** performance...
  - ❖ Locate media, transport media, mount it, etc
  - ❖ Zero TPS when the system is down

However...if you do care about performance...


44 3E07 OpenVMS Information Desk 44



## BACKUP

- ❖ Enabling media compaction increases throughput
- ❖ **SET RMS...**
  - ❖ `/BLOCK_COUNT = 127 (or 124)`
  - ❖ `/BUFFER_COUNT = 4 (?)`
  - ❖ `/EXTENDED_QUANTITY = 65535 (or 65532)`
- ❖ Compression

45 3E07 OpenVMS Information Desk 45



## Online Indexed File Backup

- ❖ **CONVERT /SHARE**
  - ❖ Record copy of indexed file
  - ❖ Uncorrupted output file
- ❖ Run prior to online VMS backup for things like SYSUAF, NETUAF, RIGHTSLIST, etc.
- ❖ Potential issue: discoordinated updates among files

46 3E07 OpenVMS Information Desk 46

## SORTing

- ❖ HYPERSORT
  - ❖ Multi-threaded
  
- ❖ Spread work files among disks/controllers/adaptors
  - ❖ Apart from input/output disks
  - ❖ No problem to have input and output on same disk
  
- ❖ Specification files are very powerful

47 3E07 OpenVMS Information Desk 47

## Sort 100,000,000 Records

Method	Processor	Time (HH:MM:SS)
Sort32 CPU	rx8620 1.6 16p	~29:00
	GS1280 16p	~25:00
Sort32 Elapsed	rx8620 1.6 16p	~29:00
	GS1280 16p	~25:00
HyperSort CPU	rx8620 1.6 16p	~12:00
	GS1280 16p	~11:00
HyperSort Elapsed	rx8620 1.6 16p	~10:00
	GS1280 16p	~8:00

- ❖ 100 bytes each
- ❖ 19,531,250 blocks
- ❖ 3 work files
- ❖ ~618,000 IO Sort32
- ❖ ~922,000 IO HyperSort
- ❖ No XFC file caching of input, output or work
- ❖ HyperSort Elapsed < CPU

48 3E07 OpenVMS Information Desk 48



## FLT - Alignment Fault Tracing

- ❖ Ideal is no alignment faults at all!
  - ❖ Poor code & unaligned data structures do exist
- ❖ **Faults on I64 vastly slower than Alpha & impact on all processes on system**
- ❖ Alignment fault summary...
  - ❖ SDA> FLT START TRACE
  - ❖ SDA> FLT SHOW TRACE /SUMMARY
  - ❖ flt\_summary.txt
- ❖ Alignment fault trace...
  - ❖ SDA> FLT START TRACE [/CALL]
  - ❖ SDA> FLT SHOW TRACE
  - ❖ flt\_trace.txt

49 3E07 OpenVMS Information Desk 49

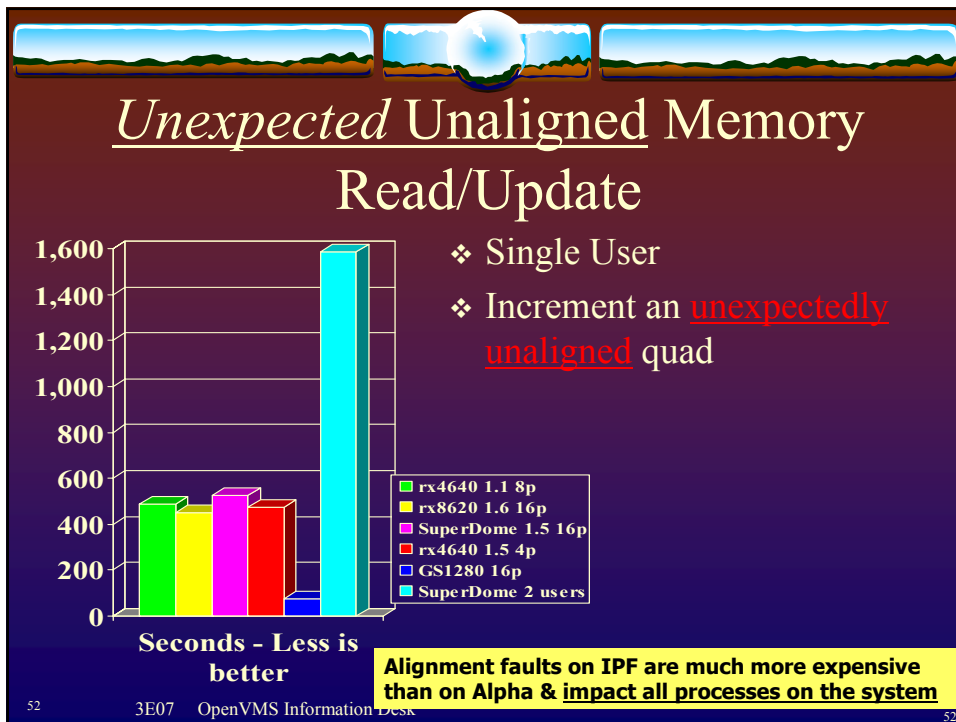
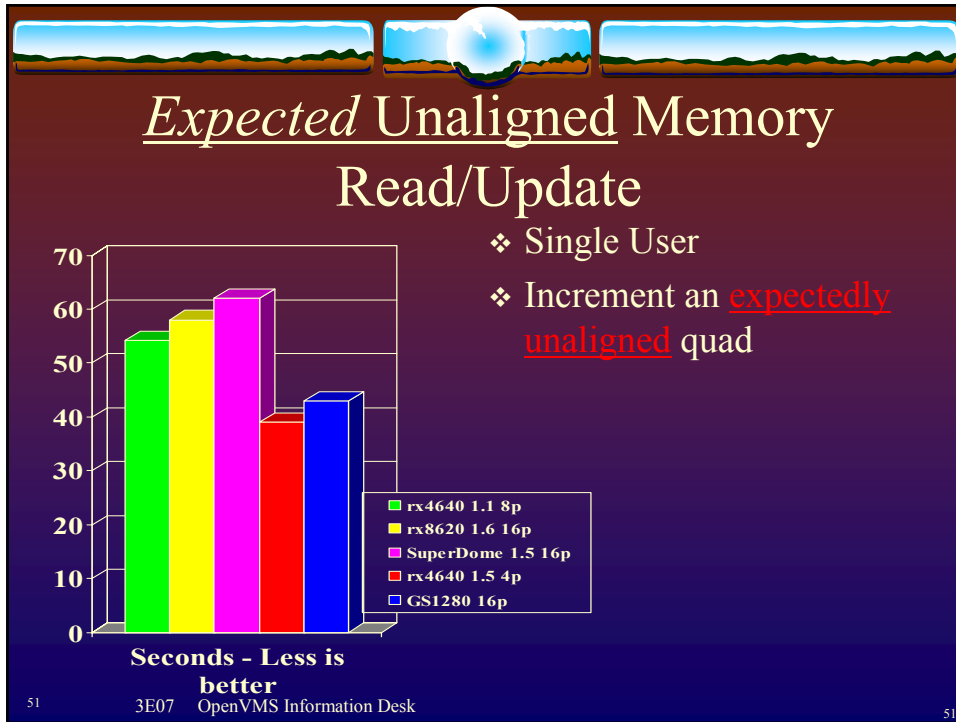
## Random Memory Read/Update Performance Comparison

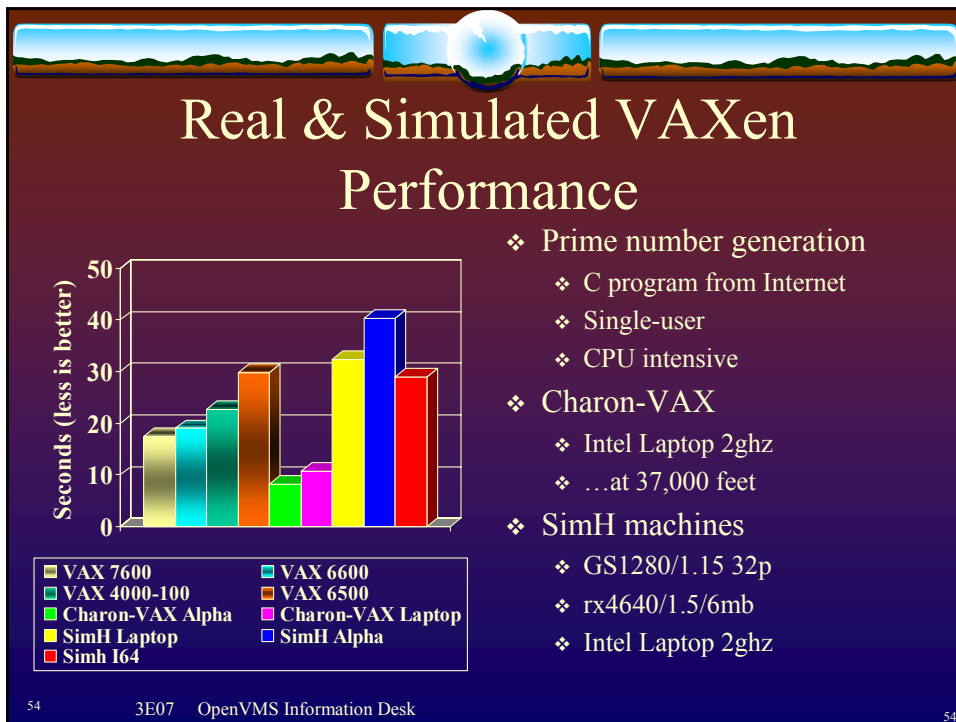
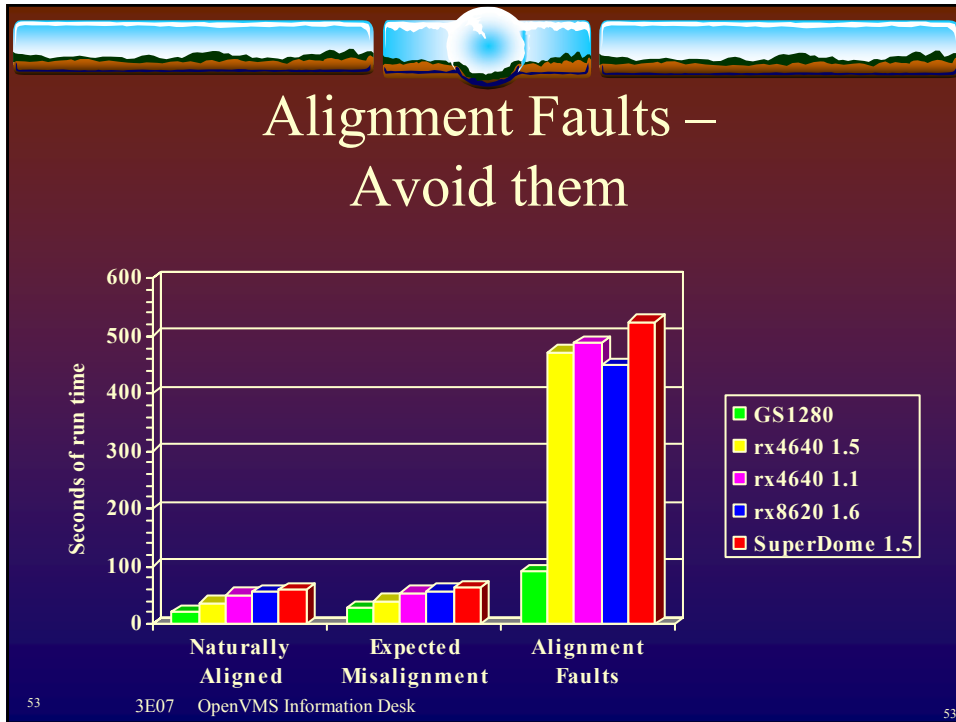
Configuration	Performance (Seconds)
rx4640 1.1 8p	50
rx8620 1.5 16p	55
SuperDome 1.6 16p	60
rx4640 1.5 4p	38
GS1280 16p	42

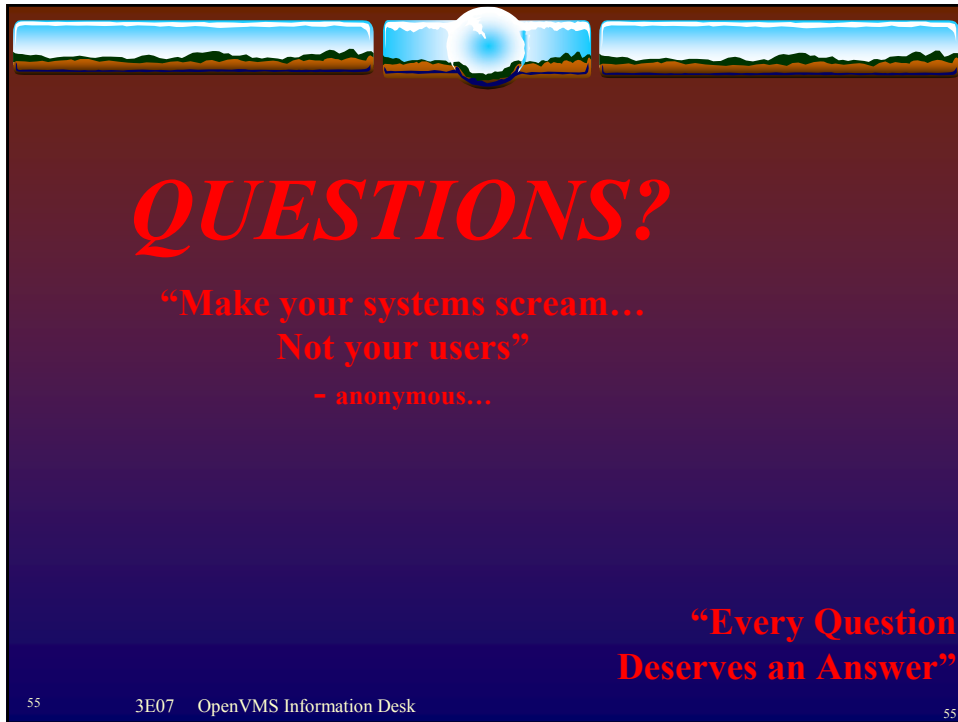
- ❖ Single User
- ❖ 1Gb global section
- ❖ 100,000,000 Loops
- ❖ Increment a random quad

**Seconds - Less is better**

50 3E07 OpenVMS Information Desk 50







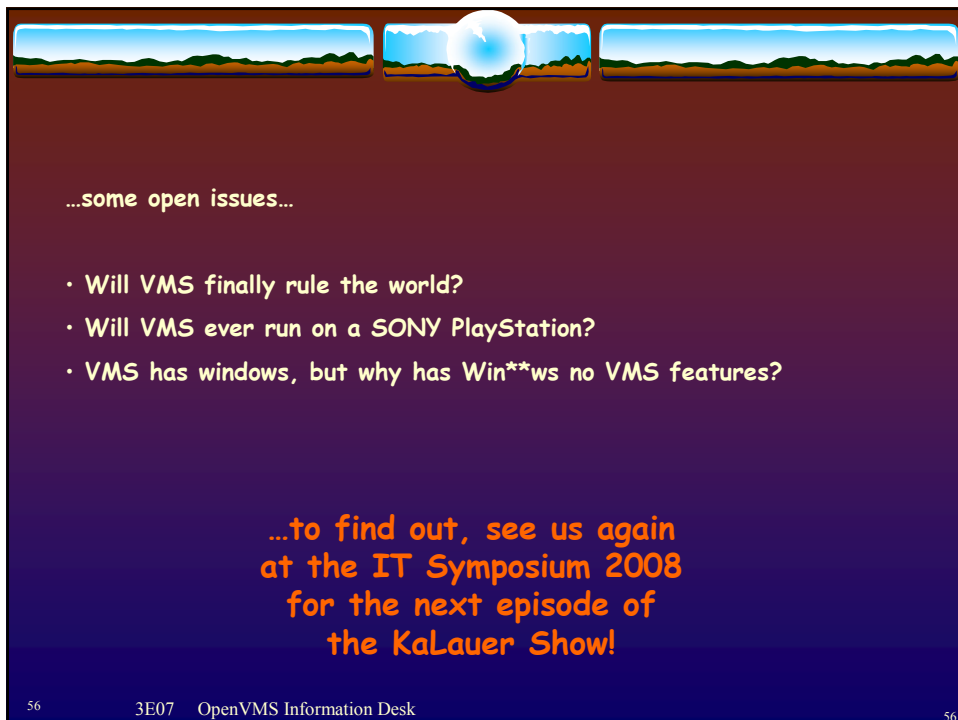
**QUESTIONS?**

**“Make your systems scream...  
Not your users”**

**- anonymous...**

**“Every Question  
Deserves an Answer”**

55 3E07 OpenVMS Information Desk 55



**...some open issues...**

- Will VMS finally rule the world?
- Will VMS ever run on a SONY PlayStation?
- VMS has windows, but why has Win\*\*ws no VMS features?

**...to find out, see us again  
at the IT Symposium 2008  
for the next episode of  
the KaLauer Show!**

56 3E07 OpenVMS Information Desk 56