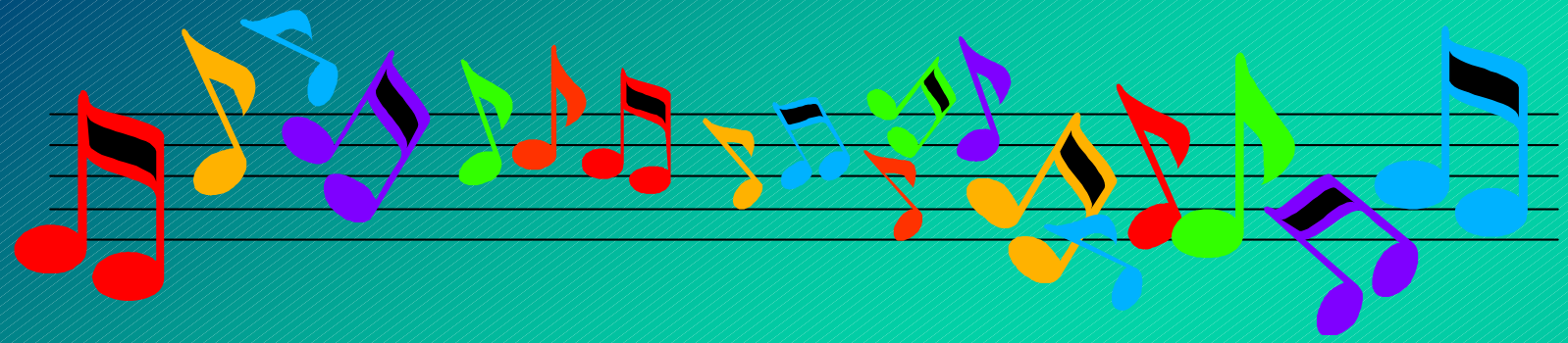
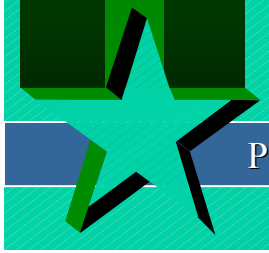


HP-UX Kernel Tuning
A Primer

Bill Hassell

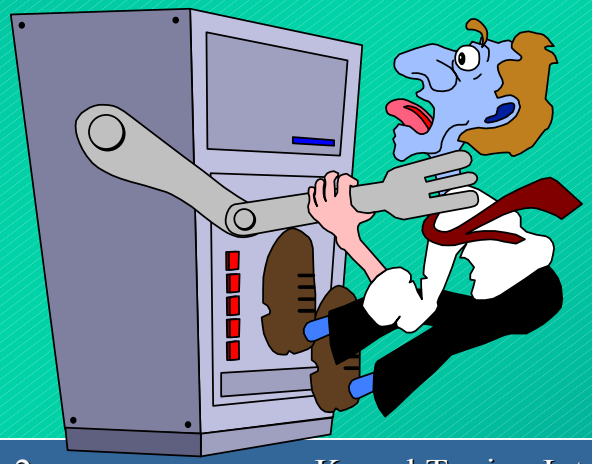
HP-UX Consultant

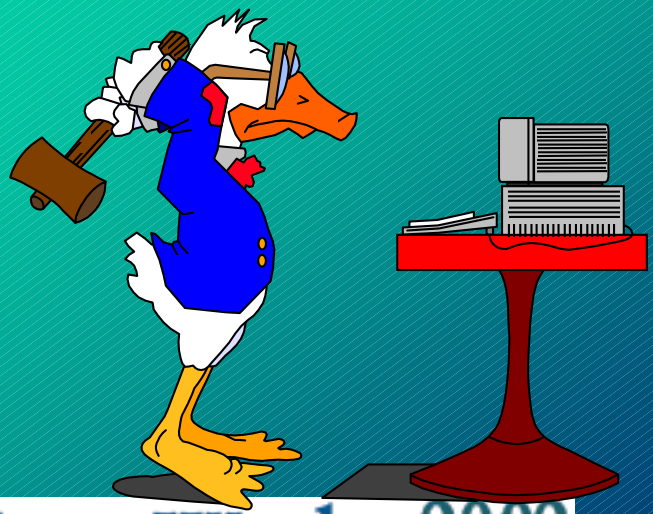


InterWorks 2002
THE HP TECHNICAL TRAINING CONFERENCE

Introduction

- What is kernel ‘tuning’?
- Parameters and their effects
- What tuning cannot do

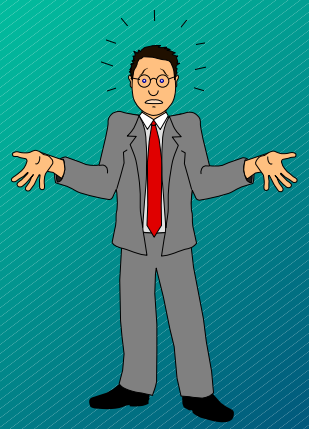




InterWorks 2002
THE HP TECHNICAL TRAINING CONFERENCE

Vocabulary

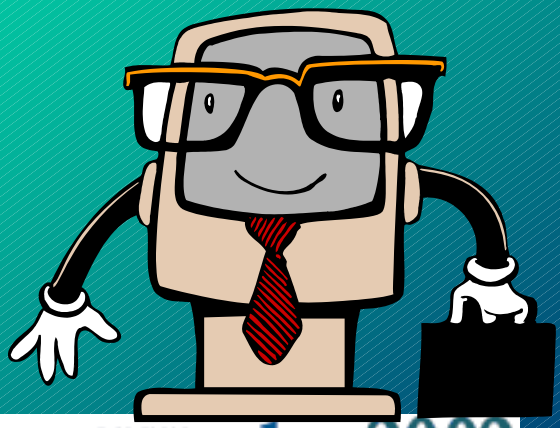
- Kernel file
  - » /stand/system (and /stand/build/system)
  - » Tunable parameters
  - » sysdef
- Measurement tools
  - » sar, vmstat, Glance



InterWorks 2002
THE HP TECHNICAL TRAINING CONFERENCE

Parameter changing

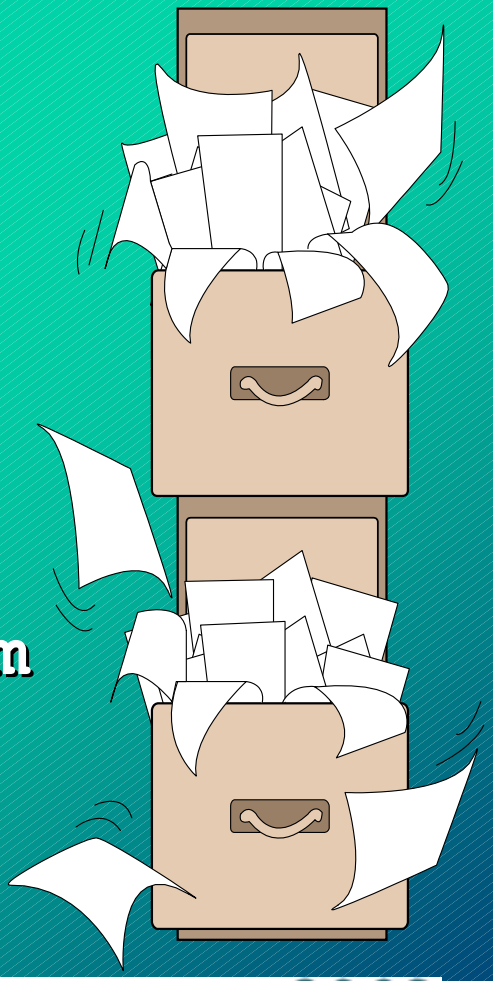
- SAM
  - » easiest to use
  - » built-in documentation
  - » some bounds checking and interaction tests
- Manual method
  - » manual edit
  - » no boundary checks



InterWorks 2002
THE HP TECHNICAL TRAINING CONFERENCE

Filesystem Parameters

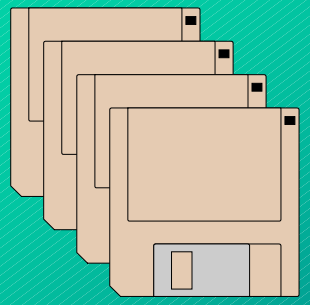
- nfile, nflocks
- ninode, vx_ncsize, ncdnode
- bufpages, nbuf, dbc_max_pct, dbc_min_pct
- maxfiles, maxfiles_lim
- fs_async, default_disk_ir
- disksort_seconds



InterWorks 2002
THE HP TECHNICAL TRAINING CONFERENCE

nfile - max file opens

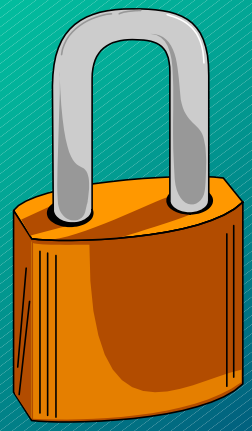
- Maximum number of file opens
  - » open counted for same files or different
  - » min 3 per process
    - `stdin`
    - `stdout`
    - `stderr`
- `file: table is full`



InterWorks 2002
THE HP TECHNICAL TRAINING CONFERENCE

nflocks - max file locks

- Maximum number of open file locks
- Highly application-dependent
  - » one file may have several locks
  - » Databases may need hundreds

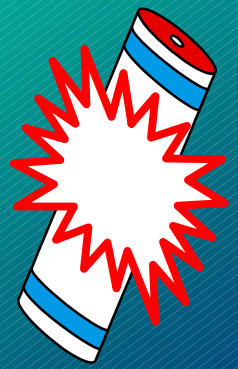


InterWorks 2002
THE HP TECHNICAL TRAINING CONFERENCE

ninode - kernel inode cache

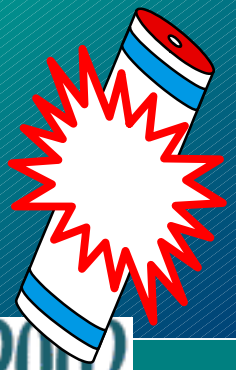
- In-core cache of ***unique*** current and recent HFS inodes locations
- Speeds reopens, multi-process file access
- Keep small for typical databases, larger for development and NFS (HFS only)
- ninode < nfile

  `inode: table is full`
- Formula may be way too large (vxfs)



ncdnode

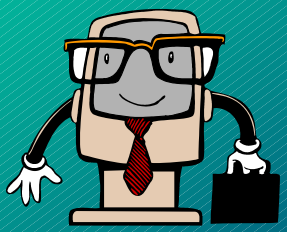
- In-core cache of ***unique*** current and recent CDROM inodes
- Speeds reopens, multi-process file access
- Generally very small (one user for CDROM)
- Copy CD's to hard disk for multi-user access



InterWorks 2002
THE HP TECHNICAL TRAINING CONFERENCE

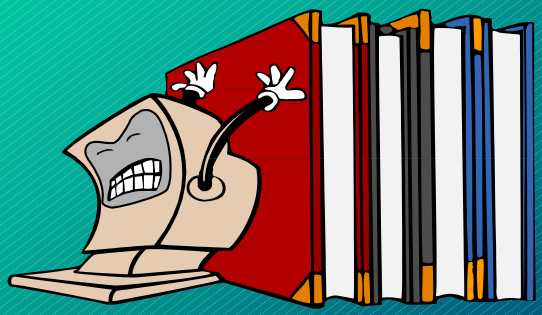
bufpages, nbuf, dbc_max_pct, dbc_min_pct - buffer cache

- Buffercache similar to DOS SMARTDRV
- used only with file access, not raw
- nbuf non-zero not recommended
- fix with bufpages
- range 200 Mb to 2 Gb (rd/wt ratio)
- Dynamic Buffer Cache min/max %



maxfiles, maxfiles_lim

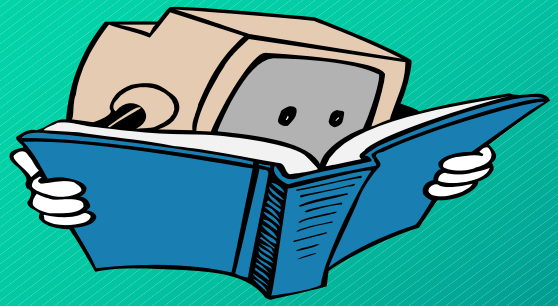
- Single process file open limits
- defaults to 60
- soft limit, extend with `setrlimit`(2) up to `maxfiles_lim`
- POSIX shell: ulimit -f



InterWorks 2002
THE HP TECHNICAL TRAINING CONFERENCE

fs_async - sync policy

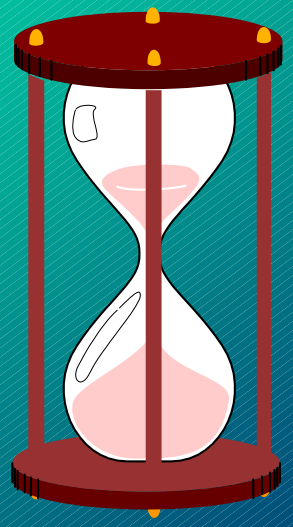
- Async filesystem writes (inc. metadata)
- 0 = safe for panics, powerfails
- 1 = risky (filesystem corruption is likely) but can improve writes (only) by 30%



InterWorks 2002
THE HP TECHNICAL TRAINING CONFERENCE

default_disk_ir - immediate reporting

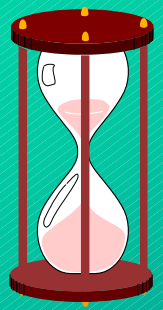
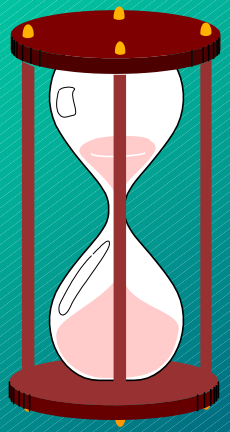
- Disk Immediate Reporting - no wait for write to complete
- Similar to **`fs_async`** except applies to all disk writes including raw
- Kernel param or scsictl by disk



InterWorks 2002
THE HP TECHNICAL TRAINING CONFERENCE

disksort_seconds

- HP-UX gives priority to serial rd/wt
- Intense serial I/O slows random I/O
- Value to wait before changing priority

InterWorks 2002
THE HP TECHNICAL TRAINING CONFERENCE

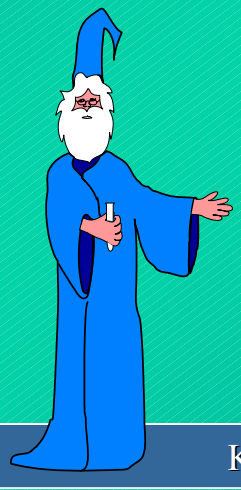
Process Parameters

- nproc, maxuprc
- maxdsiz, maxdsiz_64bit
- maxssiz, maxssiz_64bit
- maxtsiz, maxtsiz_64bit

InterWorks 2002
THE HP TECHNICAL TRAINING CONFERENCE

nproc, maxuprc - processes

- **`nproc`** = max number of processes
  - » **`proc: table is full`**
- **`maxuprc`** = max processes per UID
  - » collective UID processes, not per login

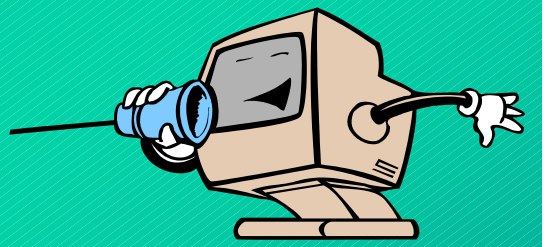


InterWorks 2002
THE HP TECHNICAL TRAINING CONFERENCE

maxdsiz - data segment

- Maximum data segment size
- Both default to 64 megs (typically too small)
- Set to apx. 900 megs or 1750 megs
- no kernel size penalty (just a fence for runaways)
- maxdsiz_64bit max is 4 Tb
- MAGIC notes (/usr/share/doc)

maxssiz - stack size



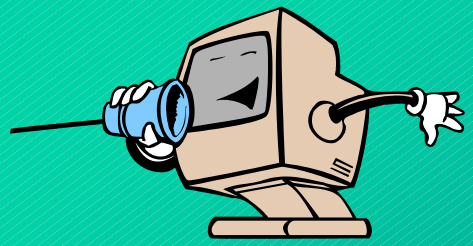
- Maximum stack size
- runaway recursion can exceed
- 79 megs is a 10.20 kernel limit, 200 megs for 32 bit 11.0 kernels and 64bit 11.0 allows up to 1 Gb
- FORTRAN arrays passed as data can easily exceed - use COMMON

maxtsiz - text size

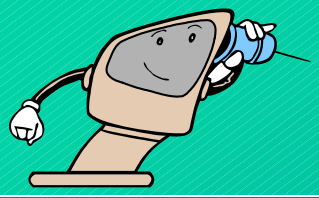
- Unchanging executable instructions
- Directly related to file size
- Seldom needs changing unless a process has an exceptionally large number of instructions
- 10.20:
  - » 64 Mb default, 4 Gb max
- 11.x:
  - » 64 Mb default, 2 Gb max (32 bit), 4 Tb max (64 bit)

InterWorks 2002
THE HP TECHNICAL TRAINING CONFERENCE

Network Parameters

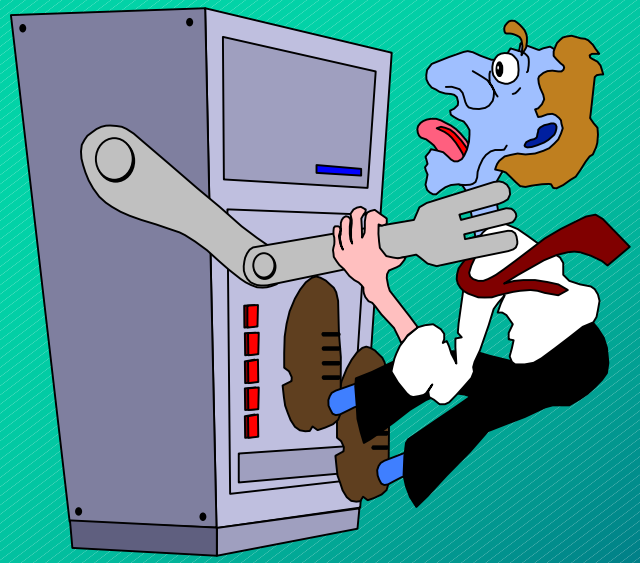


- Networking
  - **`nni`**
    - number of network interfaces
    - multiple LAN cards
    - SLIP/CSLIP/PPP interfaces
  - **`nettune`** (10.20), **`ndd`** (11.0)
    - Setup in /etc/rc.config.d



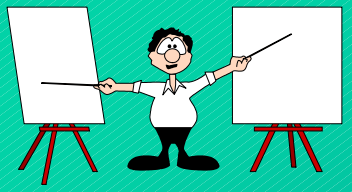
Virtual Memory (swap)

- `maxswapchunks`
- `swchunk`
- `nswapdev, nswapfs`
- `swapmem_on`



InterWorks 2002
THE HP TECHNICAL TRAINING CONFERENCE

maxswapchunks, swchunk - Swap Space



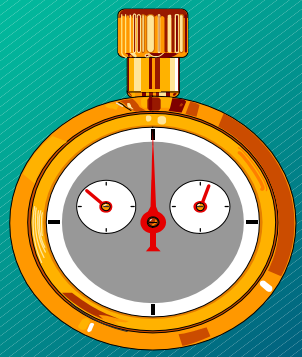
- Maximum addressing limits for swap
- Usable swap is defined as:
  - maxswapchunks * swchunk * dev_bsize where: swchunk=2048 and dev_bsize=1024
- Leave swchunk to default
- Formula simplifies to:
  - maxswapchunks=DESIRED-SWAP / 2097152 or
  - Maxswapchunks=DESIRED-SWAP / 2megs

nswapdev, nswapfs

- nswapdev - max number of swap devices
- nswapfs - max number of filesystems that will be used for swap
  - » filesystem swap performance
  - » one way assignment policy (not returned)

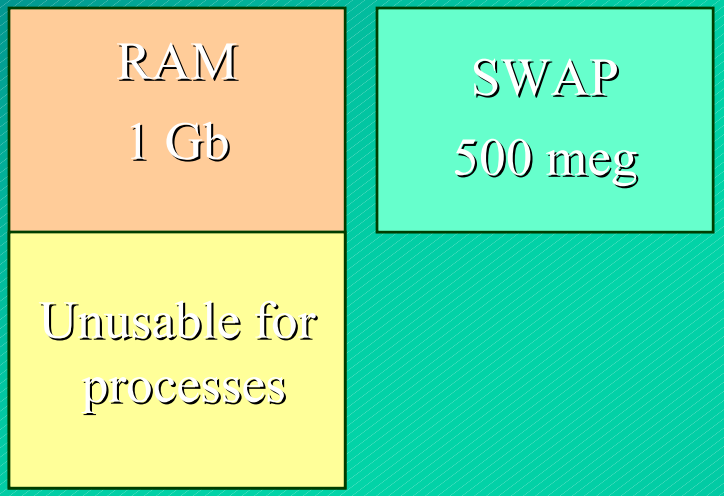
swapmem_on

- HP-UX normally needs 1:1 RAM:swap
- `swapmem_on` creates an overallocation policy, typically 75% of RAM
- Can be used in both low RAM and high RAM systems



InterWorks 2002
THE HP TECHNICAL TRAINING CONFERENCE

swapmem_on

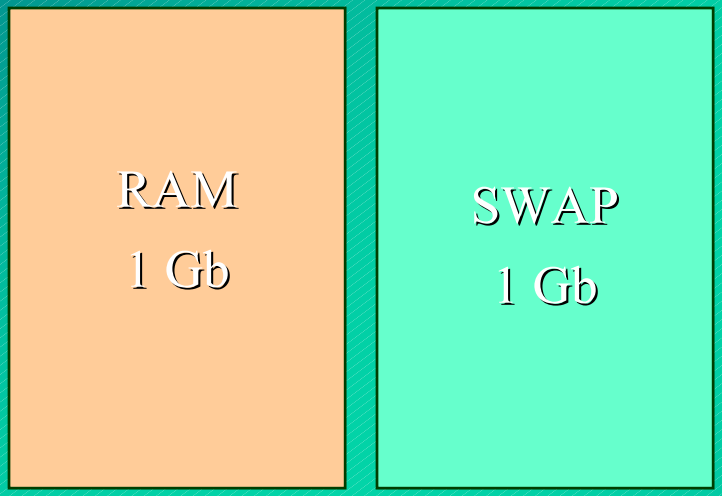


swapmem_on=0

In this example, only 500 megs is usable for processes since Virtual Memory is only 500 megs.

InterWorks 2002
THE HP TECHNICAL TRAINING CONFERENCE

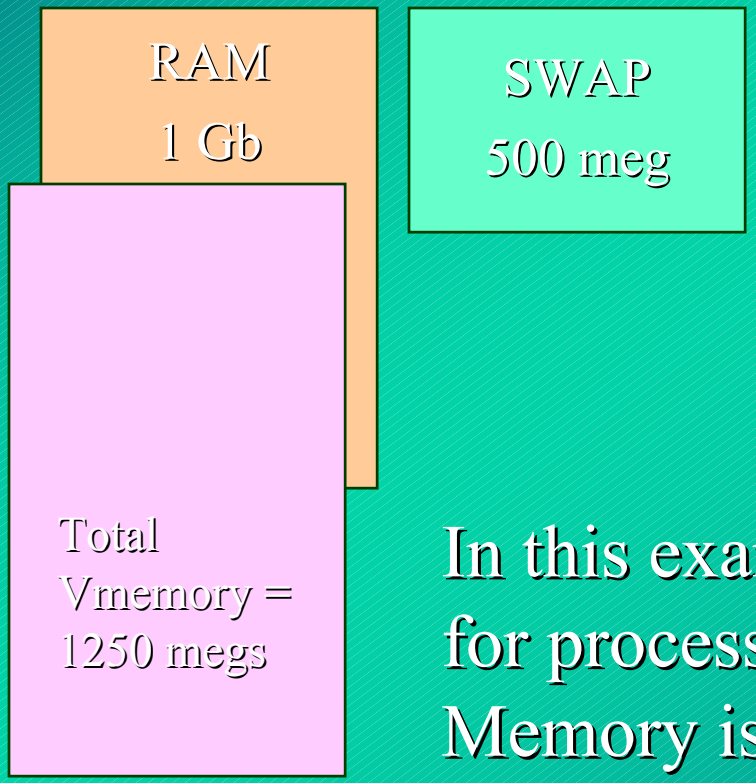
swapmem_on



swapmem_on=0

In this example, 1 Gb is usable for processes since Virtual Memory is 1 Gb too...but no paging needed

swapmem_on



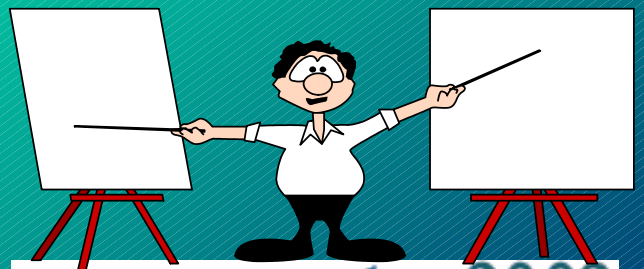
swapmem_on=1

In this example, 1.25 Gb is usable for processes since Virtual Memory is 75% RAM + swap (750 + 500)

InterWorks 2002
THE HP TECHNICAL TRAINING CONFERENCE

Miscellaneous Parameters

- Miscellaneous
  - **`timezone`**, **`dst`** (0,1,2,3 policy)
  - **`npty`** (SAM vs. manual)
  - **`timeslice`**
- Non-parameters:
  - **`maxusers`** (pseudo/formula parameter)



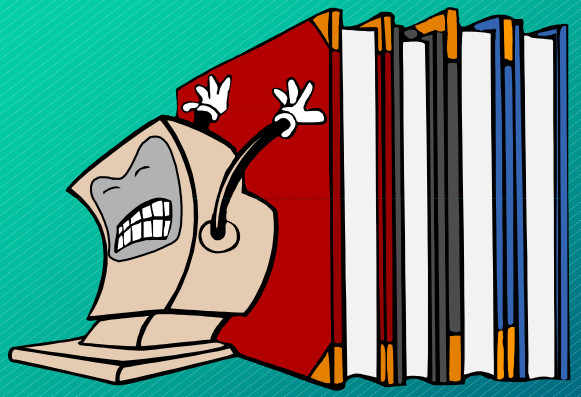
InterWorks 2002
THE HP TECHNICAL TRAINING CONFERENCE

Some last tips

- Web help:

`docs.hp.com//hpux/onlinedocs/os/KCparams.OverviewAll.html`

- SysAdmin Courses
- SAM help (10.xx+)
- Adjust carefully
  - » Small changes
  - » Few changes per sysgen



InterWorks 2002
THE HP TECHNICAL TRAINING CONFERENCE