Intermediate MPE Performance Disk & Memory

HP World 2001 – Chicago, Illinois - Booth 931

IT CONSULTING CONSORTIUM
Your Total Technology Solution

Craig L. Solomon

### The IT Consulting Consortium

Info@craigs.com - 541-924-0736 - http://www.craigs.com

# 5 Components
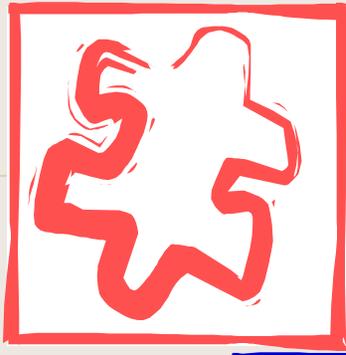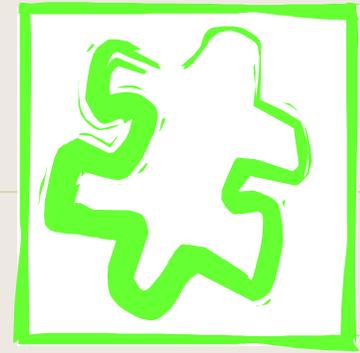
DISK

MEMORY
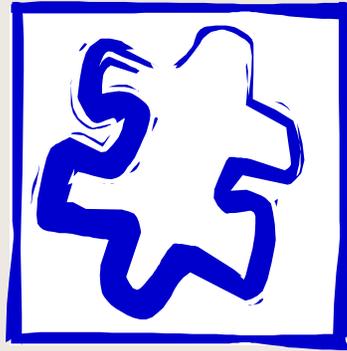
DATA

NETWORK
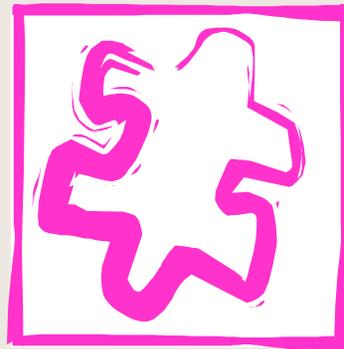
CPU

We Will Concentrate On…
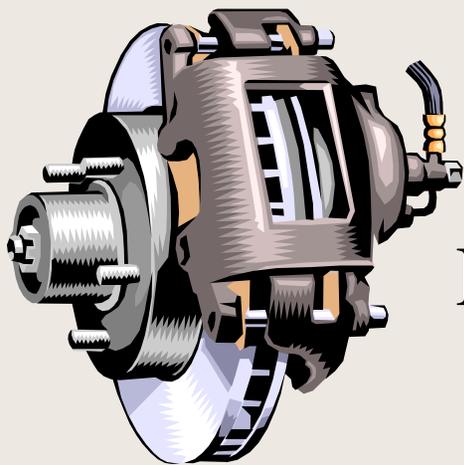
DISK
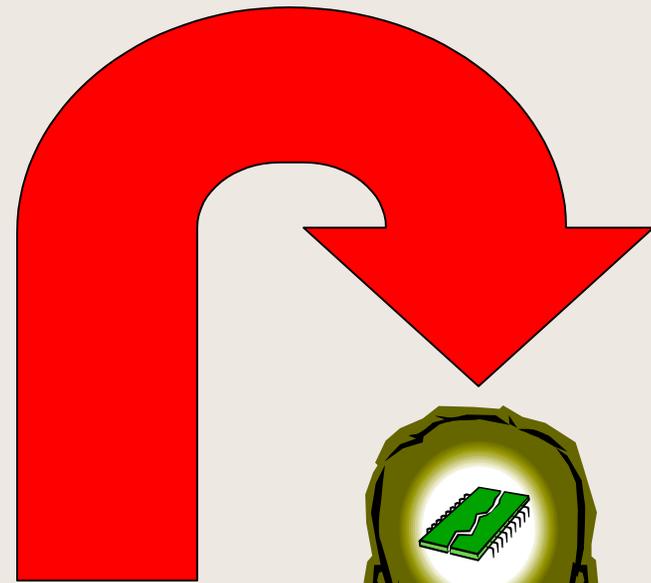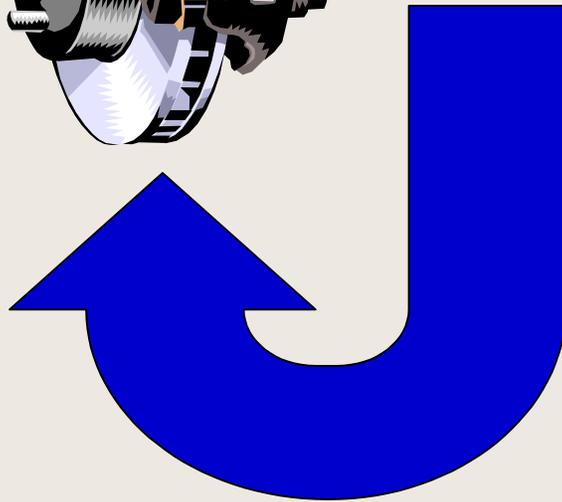
MEMORY

Disk Speed In Review

- 7200 RPM = 160 MB Second T/R = 7.4 MS Seek

- 10,000 RMP = 160 MB Second T/R = 5.2 MS Seek

- 15,000 RMP = 320 MB Second T/R = 3.5 MS Seek

•**Seek Time Defined:** Seek time is an average of how long a drive takes to move the read/write heads to a particular track on the disc. It includes controller overhead but does not include drive latency.

•**Disk I/O:** is by far the slowest of all major performance resources due to the pure physical and mechanical devices

Trickle Down…

B
U
S

Channel

| D I S C | D I S C | D I S C | D I S C |
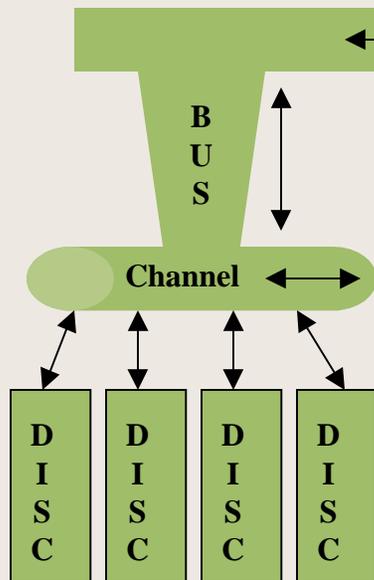
No more then
8 FW-SCSI
per card.

System Device Speed

SCSI Disk Drives   2 MB/sec

LAN (10 Megabit)   1 MB/sec

LAN (100 Megabit)   10 MB/sec

System Channel Speed

FW-SCSI   5 MB/sec

What Disks To Buy

9 GB

18 GB

72 GB
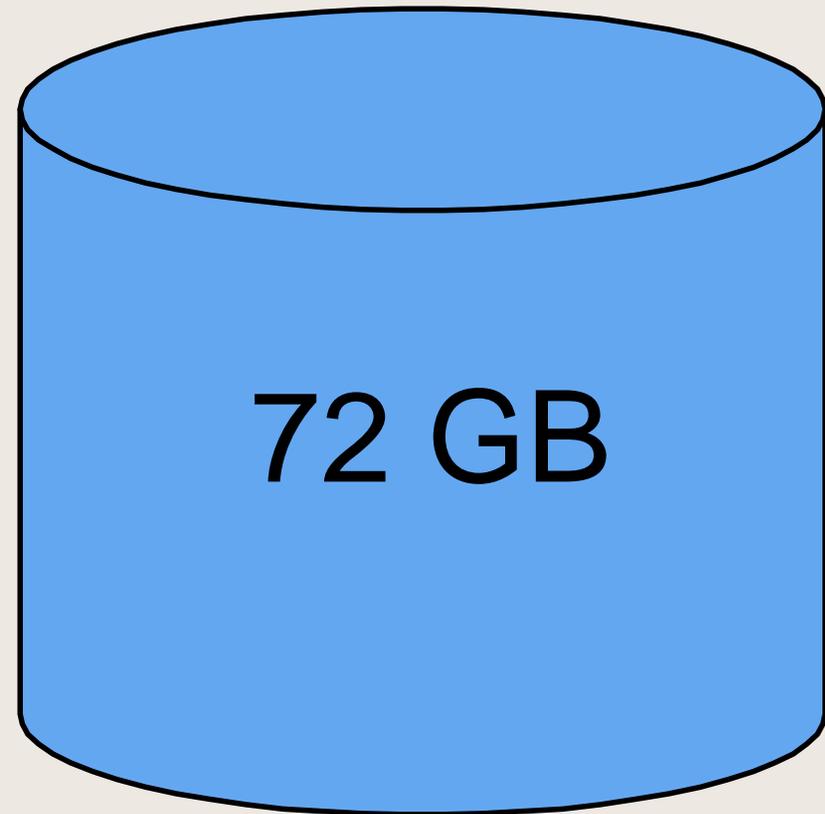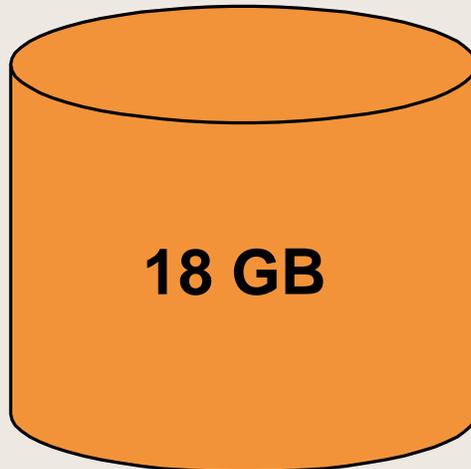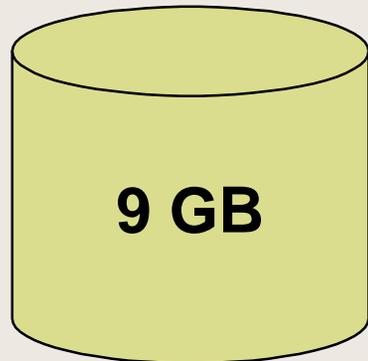
MPE uses the drive
with most available space first.

Volume Set = PROD

Master

9 GB

Member
4 GB

Member
4 GB
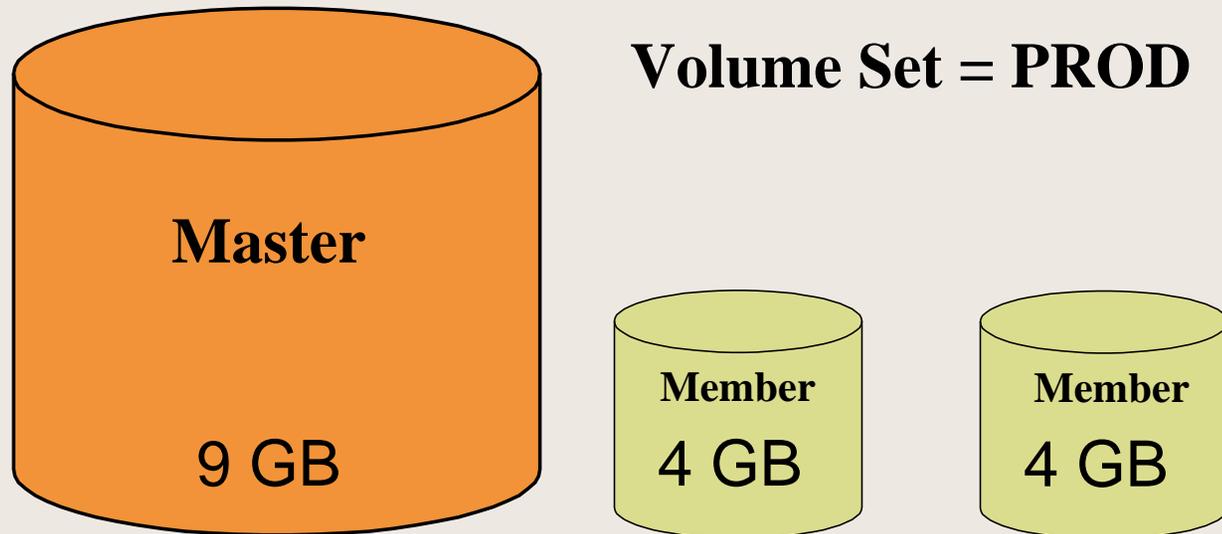
Is 50% of 9 GB the same as 50% of 4 GB?

Master	Member	Member
9 GB	4 GB	4 GB

Volume Set = PROD
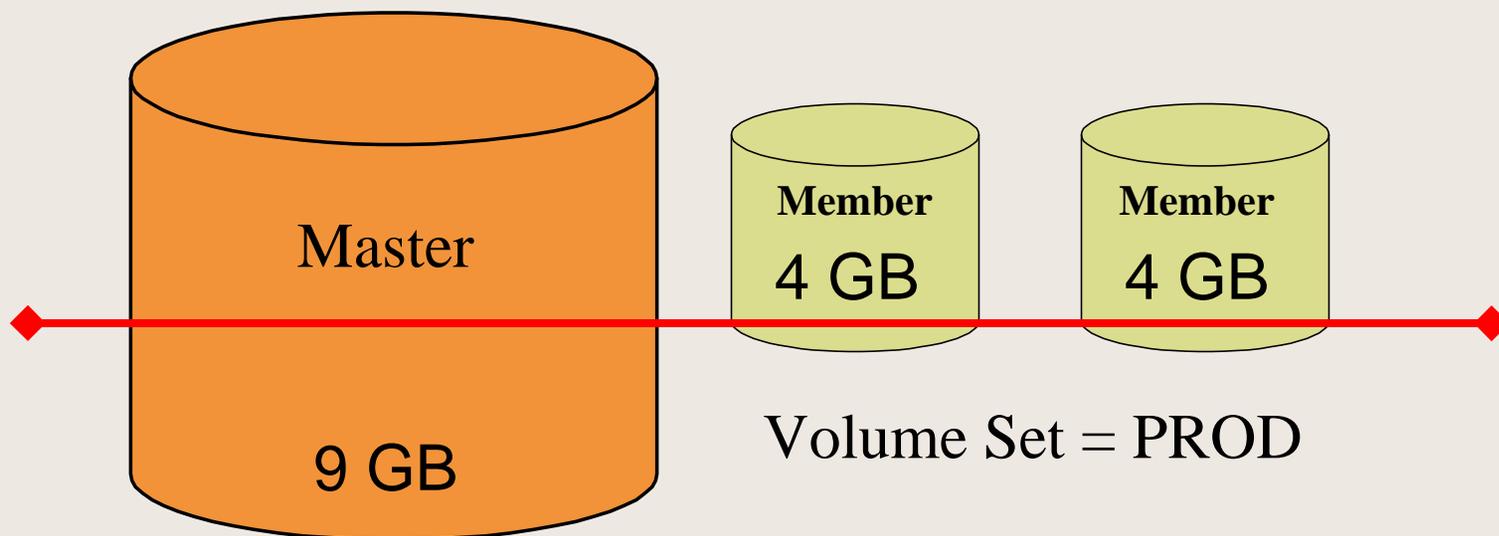
First 4 GB (approx) will be written to only the first drive (VS Master) of PROD.

What happens to Data & Locality?

Disk Space Management

- Try to keep free disk space at greater than 20%

- Double the estimated disk space requirements for new applications

- Trend disk space usage with a tool like Excel.

Transaction "XM" Manager

User Transaction

MEMORY

**Prod Volume
Set Master**

XM
LOG

The XM Reaches A Check
Point When 1/2 Full ! Then
Flushes To Disk !

Member
1

Member
2

Member
3

Member
4

Member
5
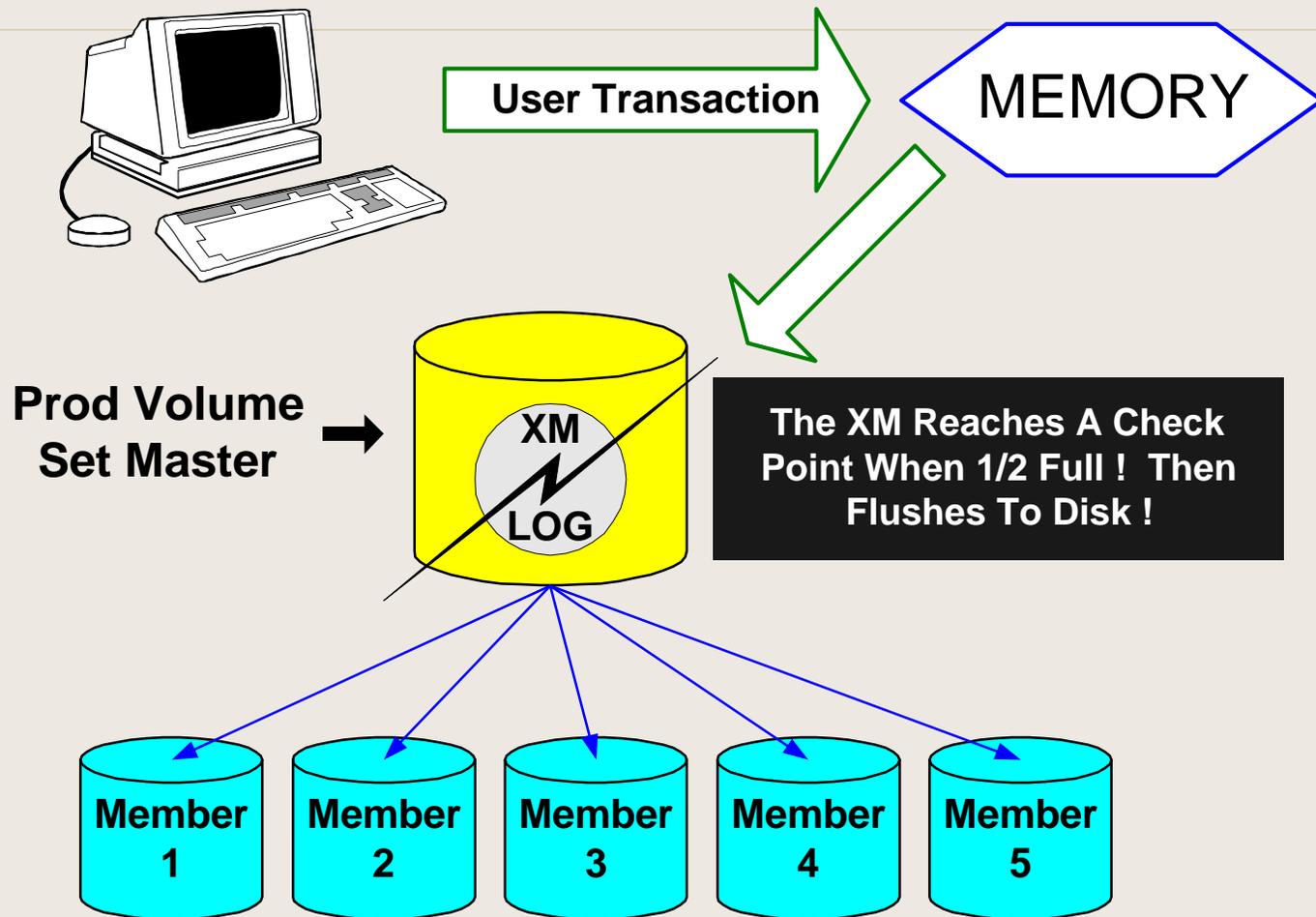
XM Checkpoint Continued.

- Now HP allows **you** to modify the Transaction Manager Log File.

The questions are…

- Should You Make The Log File Larger?
- Should You Make the Log File Smaller?

XM Checkpoint Info!

Detour

XM Checkpoint Information included at the end of this presentation!

Extents

- MPE/iX writes files to disk (normally) in multiple parts. These are called **extents**.

- When a disk is heavily fragmented, then MPE must break the file into smaller and significantly more extents.

- This increases Disk I/O and hinders performance.
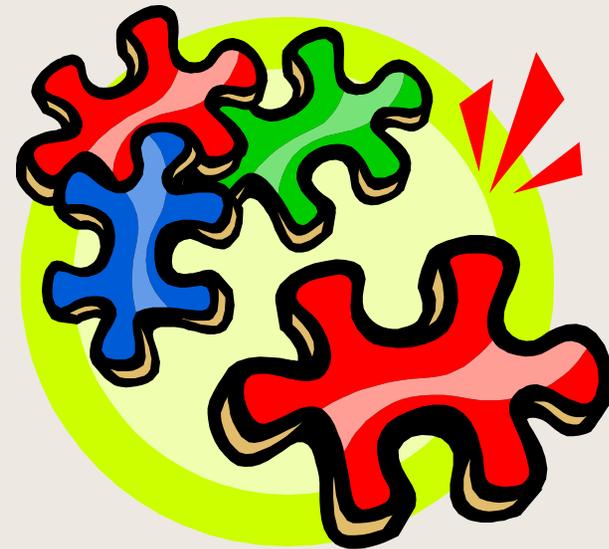
Extents Sample

Note: maximum number of extents in any file seen:

12,964, for file FILENAME.DATA.ONLINE ⬅ **WOW**

Fragmentation by fileset:

Frag%	#Pages	#MBs	#Extents	#Files
11%	6,979	27	216	47

files examined: 7,234,755

Space savable by TRIM: 83,214 pages; 325.0 MBs

files on/partially-on ldev 65 = 572

What's worse yet…

- MPE/iX cannot write to chunks smaller than 1000 sectors.

- Too much wasted space on the disks, too much fragmentation, too many extents = too many headaches!
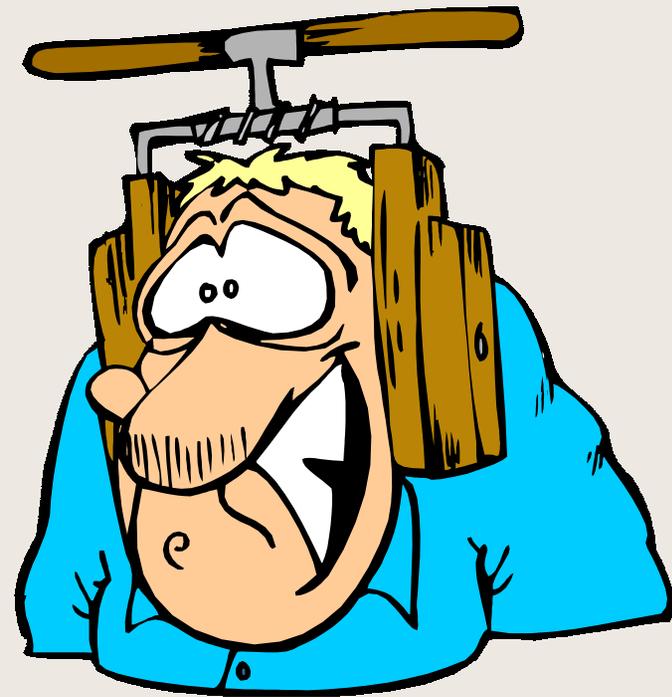
Unusable Disk Space

Unusable

```
                     * * * Disc Space Utilization * * *
      ----Used---   ---Avail---   --Max %---   --------Contiguous Free Space--------
Dev  Trans  Perm   Trans  Perm   Trans Perm    Largest   <100 <1000  <10K <100K 100K+

  1 1859K 2938K   3532K 3309K     75    75    2648464   8560 62640   240K  188K 3032K
  2  848K 3939K   3542K 3542K    100   100    2330288  32320  203K   387K  589K 2330K
  3  723K 4034K   3572K 3572K    100   100    2180240  26704  255K   366K  149K 2774K
  4  722K 4084K   3523K 3523K    100   100    2559072  41088  327K   322K  273K 2559K
  5     0 6203K   2127K 2127K    100   100    1891504    208  3680   6160     0 2117K
 12     0 6215K   2115K 2115K    100   100    2103072   1232  1648   9088     0 2103K
 13     0 6284K   2046K 2046K    100   100    2039776   2560  3072   1104     0 2039K
 14     0 6306K   2024K 2024K    100   100    2016832     16  2000   5200     0 2016K
 30     0   11M   5775K 5775K    100   100     304416  58720  528K  1726K 2527K  934K
 40     0   11M   5431K 5431K    100   100     380416  18528  285K  1844K 2185K 1097K
 41     0   11M   5387K 5387K    100   100     304336  19168  355K  2075K 2178K  758K
 42     0   12M   5213K 5213K    100   100     485216  26352  448K  2350K 1174K 1213K
 43     0   11M   5306K 5306K    100   100     381680  24112  395K  2179K 1940K  768K
 44     0   12M   5180K 5180K    100   100     384608  33024  470K  2129K 1661K  885K
 50     0   11M   5743K 5743K    100   100    4975392  32080  109K   564K 61696 4975K
 51     0   11M   5725K 5725K    100   100    4903504  37968 95040   607K 81616 4903K
```

What Do We Do?

So, let us make life easy; let's buy all the same drives and reload the system.

Life is now wonderful... Right? It is supposed to be, isn't it?

Oh NO! Now what?

What About The Data?

- Reloads solve DISK fragmentation.

- Reloads do nothing for file fragmentation or data locality.

FREE SPACE

DATASET "B"   DATASET "A"
DATASET "A"   DATASET "B"
DATASET "B"   DATASET "A"
DATASET "C"   DATASET "C"
    DATASET "B"
DATASET "C"   DATASET "A"   DATASET "C"
DATASET "B"   DATASET "A"
DATASET "C"   DATASET "C"
DATASET "C"
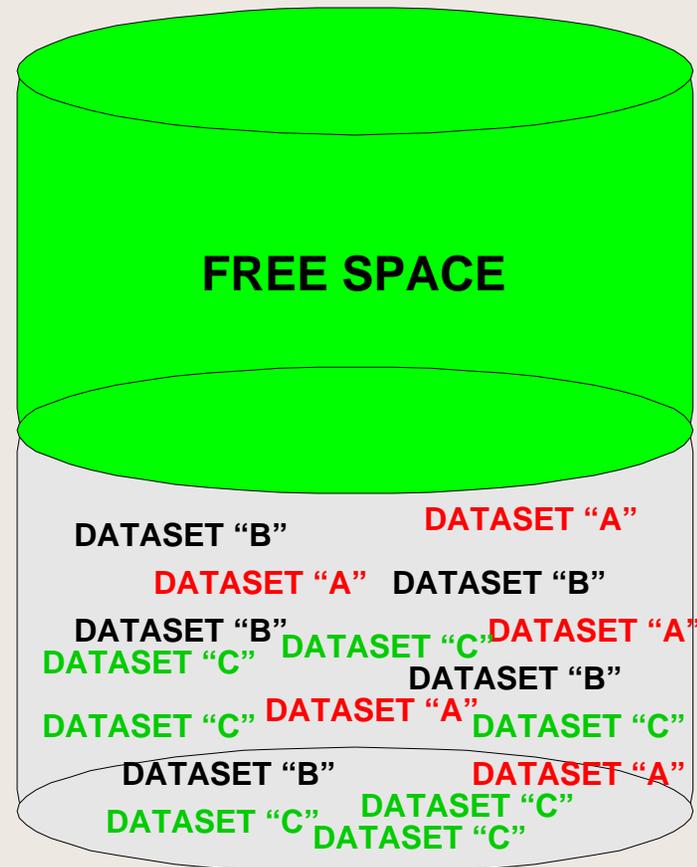
What About The Data?

- What if database "A" has 4 data sets?

- What if you have 4 disk drives on PROD, where database "A" resides?

- What if most of 3 of your 4 datasets fall on a single LDEV of your PROD volume set?

What About Data Locality?

FREE SPACE

FREE SPACE

FREE SPACE

FREE SPACE

Assorted Other Files

Assorted Other Files

DATASET "2"

DATASET "1"

DATASET "3"

DATASET "4"

Assorted Other Files

Assorted Other Files

IT Consulting Consortium – www.craigs.com - 541-924-0736

I/O = Off Balance!

I/O = Balanced!

Disk Pulse Points

Indicator	Good	Avge	Bad	Comments
MEMORY				
Memory Manager	< 4	4 - 8	> 8	CPU % Managing Memory
Page Fault	< 20	20 - 60	> 60	Varies Based On System
Clock Cycles Per Hour	< 10	10 - 15	> 15	Solid Memory Indicator
Disk				
Pause (CPU Pause)	< 5	5 - 15	> 15	Data Locality & I/O Efficiency
Disk Queue Length	< .5	.5 - 1.0	> 1.0	Firm Indicator Of I/O Efficiency
Read Hit %	> 95	90 - 95	< 90	Indicator Of Data Locality

Above based on an Small System A-Class / 9x7

Disk I/O Queue Length

I/O Queue Length Should Not Exceed .5
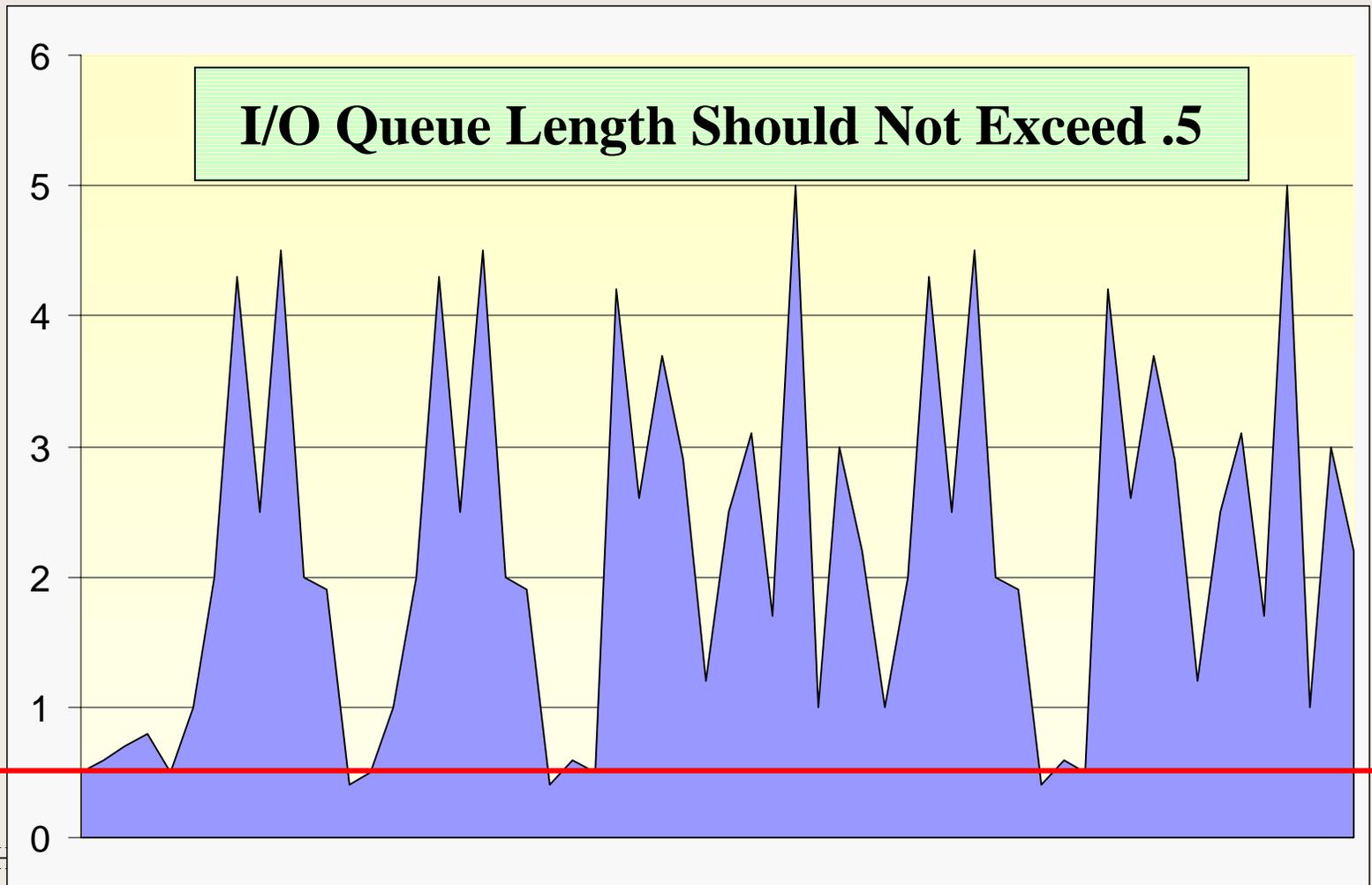
I/O Queue Length By Drive
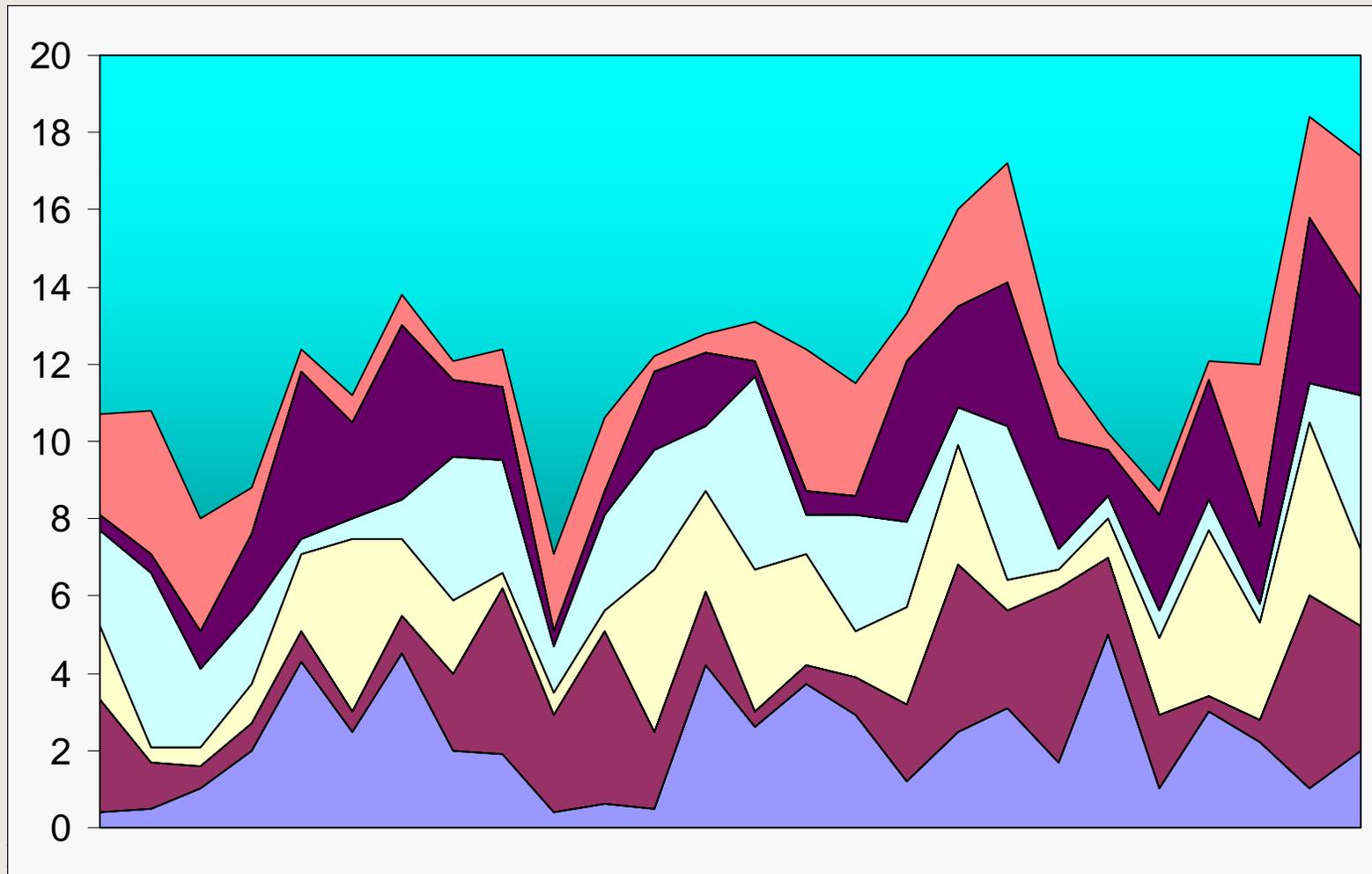
Disk Maintenance Tips
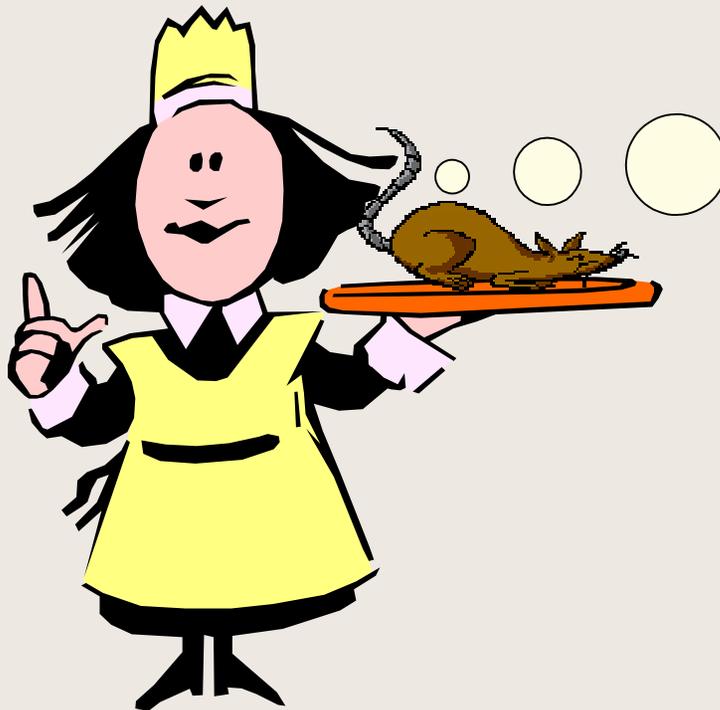
Disk Tips @

The End Of The

Presentation!

Memory

Memory Management

- Memory Management is actually a CPU measurement.

- Page Faults, up until recently were referred to as a indication of system death. However with larger systems and larger installed memory page faults are expected as long as they are within system limits.

- Processes rely on memory to access ready requests. Memory is infinitely faster then disk. Performance goal = avoid DISK I/O!

- New Memory Management on version 6.5 and 7.0 Express 1, including larger memory support.

Memory – Page Fault

- When a chunk of data is missing from memory, the memory manager will issue a page fault.

- Then a swap will occur pulling the desired page into memory.

A Page Fault is a reference page of memory that is absent.
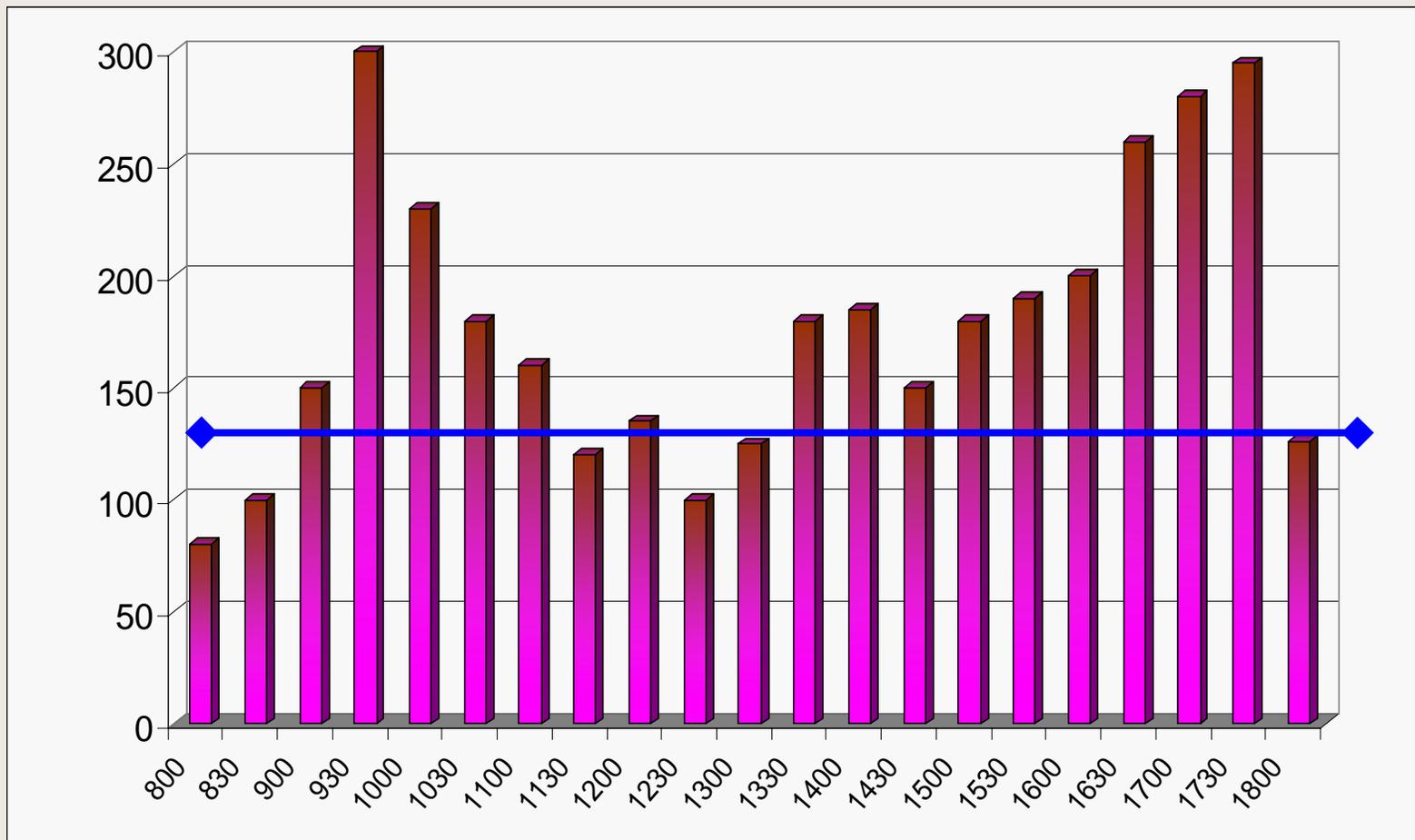
Read Hit Rate

- Read hit rate is the percentage of time that disk read requests were satisfied in main memory. This is primarily a function of Data Locality. A shortage of main memory can also negatively affect this number.

- Poor memory pre-fetch also could account for an overactive Memory Management Reading.

Read Hit Rate
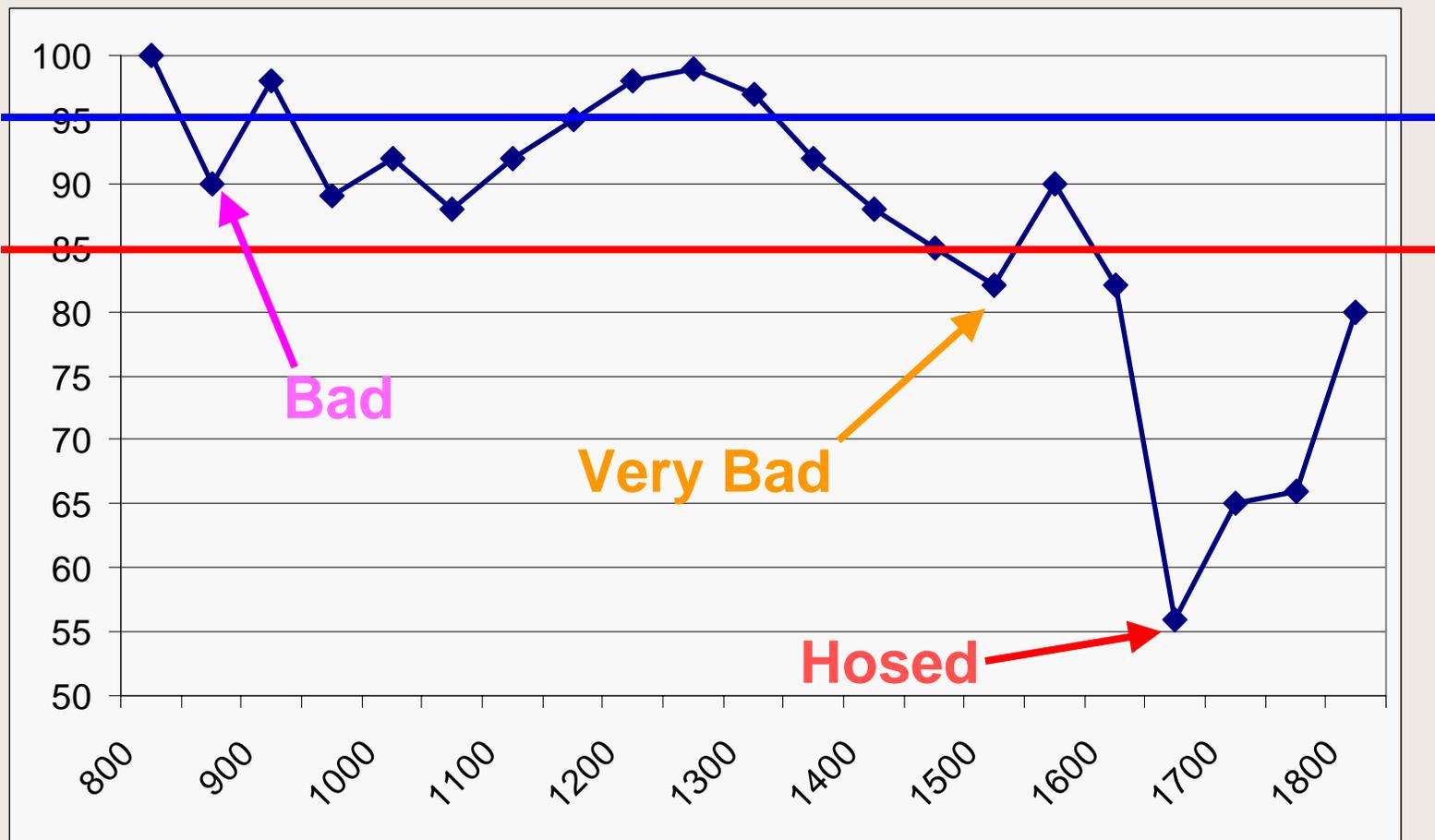
Memory Management Overview

•Kicks out infrequently accessed pages to make room for urgent requests. Data gets written to virtual memory (ts). Code simply gets overlaid.

•Allocates permanent memory to Operating system.

•Use CPU to do its duties, thus the more CPU spent on this activity the more likely memory pressure exists.

•Keeps track of the busiest pages of memory. These pages get to stay longer.

•Dirty pages are modified and swapped back to disk when required.

Memory Manager

If Memory Manager exceeds 4% there is a good chance that there can be a Memory shortage. This can also be a sign of data issues as well as assisting in causing a CPU bottleneck
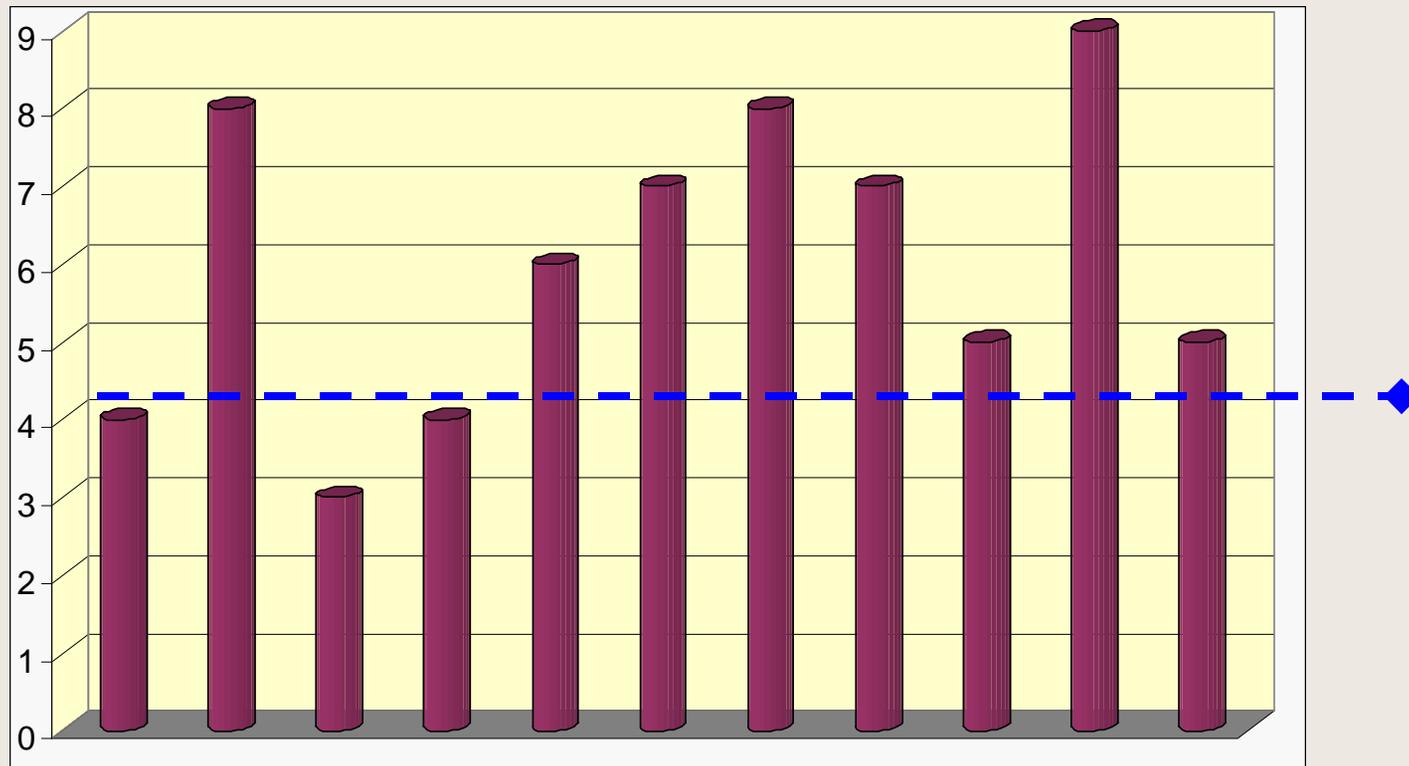
Sequential Reads

- 1st Read - (3) Pages 12KB I/O

- 2nd Read - (6) Pages 24KB I/O

- 3rd Read - (12) Pages 48KB I/O

- 4th Read - (24) Pages 96KB I/O

More Memory Info @ the end of this presentation!

Memory Pulse Points

Indicator	Good	Avge	Bad	Comments
MEMORY				
Memory Manager	< 4	4 - 8	> 8	CPU % Managing Memory
Page Fault	< 20	20 - 60	> 60	Varies Based On System
Clock Cycles Per Hour	< 10	10 - 15	> 15	Solid Memory Indicator
Disk				
Pause (CPU Pause)	< 5	5 - 15	> 15	Data Locality & I/O Efficiency
Disk Queue Length	< .5	.5 - 1.0	> 1.0	Firm Indicator Of I/O Efficiency
Read Hit %	> 95	90 - 95	< 90	Indicator Of Data Locality

Above based on an Small System A-Class / 9x7

Memory Tips

- **<u>Main Memory Rules</u>**
  - 128 MB per CPU
    - 1=128   2=256
    - 3=384   4=512
    - 5=640   6=768
    - 4=896   8=1024
  - 45 MB for Operating System
  - 1 to 2MB per User
  - 2 to 4 MB per Batch Job
  - Plus Application Requirements

Thanks For Coming
HP World 2001

IT Consulting Consortium - Booth 931

Disk Space Management Tips

- Archive and purge extraneous system log files, such as LOGXXX.PUB.SYS. Only 3 to 6 months worth of log files need to be keep on the system.

- Re-size any databases that are used for archiving data on a regular basis..

- Look for files that have not been accessed recently. Scan a full system backup for files that are not part of active applications..

- Check for purge candidates among files having an "SD" after the size. These are self describing files created by a save of records found in QUERY.

- Backup and purge all unused groups and accounts.

- Perform a ALTFILE;TRIM in MPEX to recover any unused disk space after the EOF file marker.

- Purge files that have a single letter followed by a series of numbers. These are generally temporary files created by an update / edit of files via Quad or Editor.

Memory Pulse Points Defined

- **MEMORY PULSE POINTS**
  - CPU Memory %
    - The % of time the CPU spends cycling through main memory.
  - Page Fault Rate/sec
    - The number of times per second that memory page faulting occurred.
  - Clock Cycles per hour
    - The number of times memory must complete a full cycle while looking for free space in memory.
  - Swaps/Launch (Ratio)
    - The ratio of the number of swaps to launches. This means for every time a process comes up to bat with the CPU, necessary segments were not present in main memory. Therefore the process could not run.
    - Library Fault (%)

XM Checkpoint Info Page 1

•A new enhancement to the XM in 6.5 is to increase the userlog. The current userlog size is 64 MB per volume set. Currently, for each transaction, the information that can be contained in the userlog is limited to 4 MB. Due to the increase in the transaction size for each process (to 32 MB), the total userlog file should also be increased. The userlog resides on the master volume of a volume set. User can use **VOLUTIL** utility to increase the size of the userlog (the default value is 64 MB) with the following,

•From **VOLUTIL, ALTUSERLOG <volset-name> <new size>**This command accepts the volume set name and the new size of XM userlog for it in sectors. It can not be less than 264256 sectors corresponding to 64 MB. One caution is that ldev 1 has limited reserved space and one can not expand the XM userlog beyond that.

•The actual expansion of the XM userlog happens when user boots the system next time. This is because the XM log files can not be expanded/truncated while in use.

XM Checkpoint Info Page 2

- During bootup, just after recovery of the userlog, the XM userlog is expanded and a message is written onto the console.

- One can also truncate the userlog after expansion if there are any disk space or performance issues. The volutil interface is same for this operation too.

- One can backdate the OS to an earlier version with the userlog in the expanded state.

- The expansion of the userlog may fail due to one or more reasons. In that case, the system reports it through a warning message on the console.

- If it fails due to lack of disk space, make sure that there is sufficient contiguous space on the master that would fit the need for the increment in the userlog size. Then reboot the system so as to enable it to complete the expansion of userlog. You must ensure that memory beyond 256 MB is available on the system, before increasing the userlog through VOLUTIL.

Courtesy MPE/iX Communicator