Active Memory Expansion
for AIX 6 & 7

Nigel Griffiths
IBM Power Systems

Presentation Version 7

Active Memory Expansion Announcement Confusion

1. Marketing thought:
   - "Expansion" sounded better than "Compression"
2. It does not "just compress memory pages"
   - It's a lot cleverer than that!
3. "AME" also used for AIX Management Edition
   - So ActMemExp was used ….. "doh!"
   - Now AME = Active Memory Expansion

Active Memory Expansion Pre-Reqs:

POWER7 based machine
AIX 6.1 TL04 SP2+

Also note:

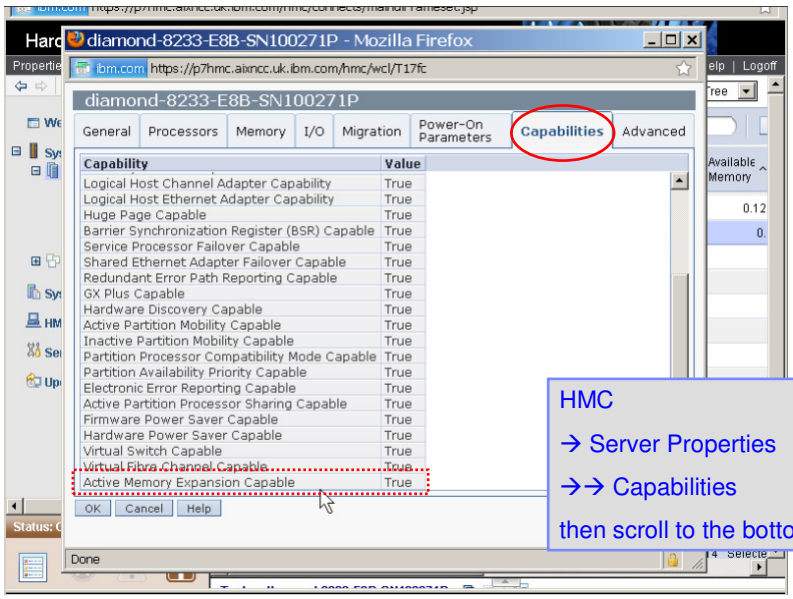
- Transparent to all applications
- Not IVM - Activation key via the HMC
  - But configured at LPAR level
- AME will switch off AIX 64KB page support
  - Can be enabled but tests showed it was slower

Permanent Enablement - Chargeable

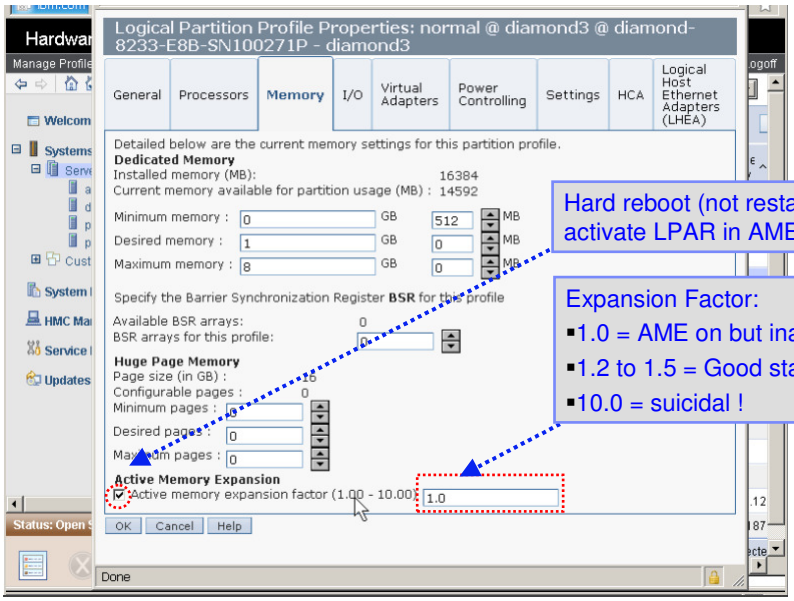
- One feature per server
  - No matter how many partitions (LPARs) use it
  - Permanent enablement → new server or via MES order
  - Enablement "VET" code applied to the VPD anchor card
  - Once enabled: no mechanism to move it to a different server
- Power 750 & Power 755
  - #4792 AME Enablement Feature
- Power 770 & Power 780
  - #4791 AME Enablement Feature
- It is a processor feature !!!
- One-time, 60-day Trial - No charge
  - Request via Capacity on Demand Web page www.ibm.com/systems/power/hardware/cod/

How do we switch AME on?

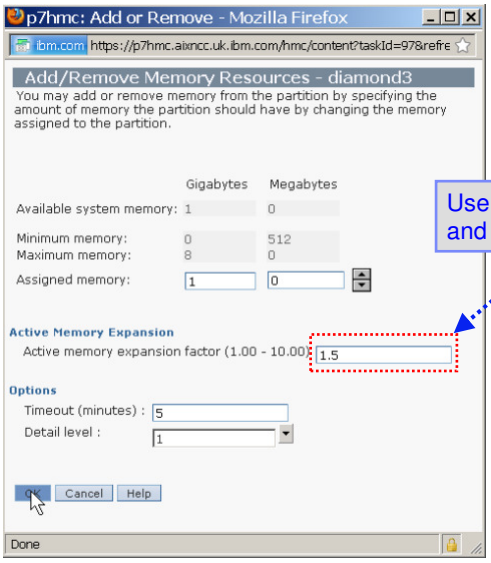
Is the machine AME Capable?



Activate on AME on the LPAR profile



Dynamically changing the Expansion Factor



How does AME work?

AME Conceptual Model

Memory Pages

AME Conceptual Model

Memory Pages

RAM Disk

Not actually a RAM disk
but similar concept

AME Conceptual Model

Memory Pages

RAM Disk

Page in “on demand”

Page out LRU pages

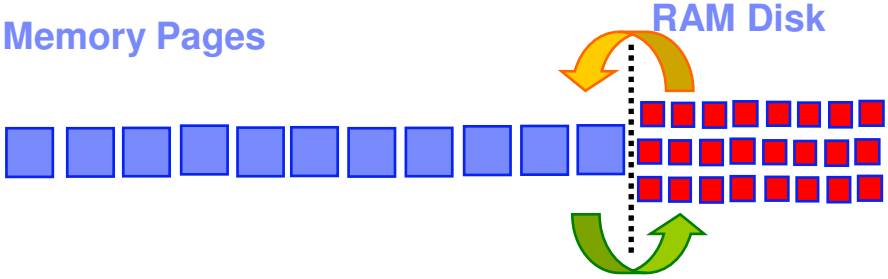
Use it like a very fast
paging device

LRU = Least Recently Used = oldest unused

AME Conceptual Model

Memory Pages

Compressed
RAM Disk



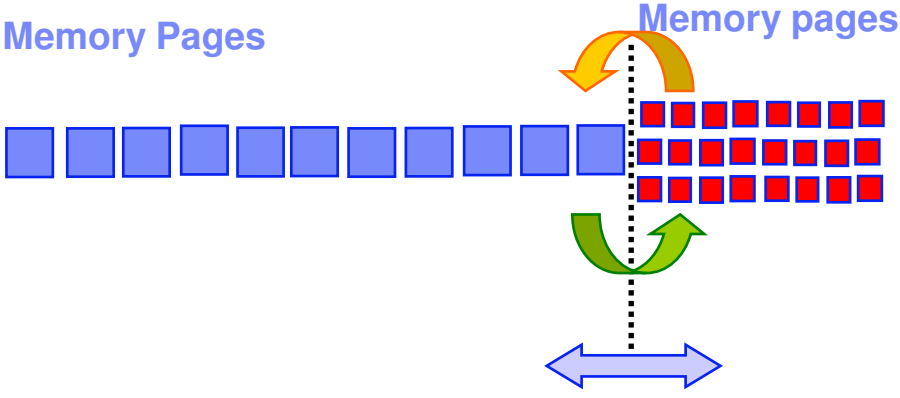
Now while paging,
shrink the memory pages
so many more pages fit

15 true memory →11+24=35 so Expansion Factor=15:35 = 2.33

AME Conceptual Model

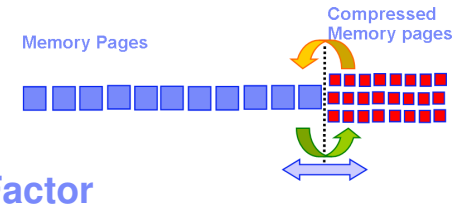
Memory Pages

Compressed
Memory pages



Dynamically adjusted depending
on compression ratio & target

AME Practicalities

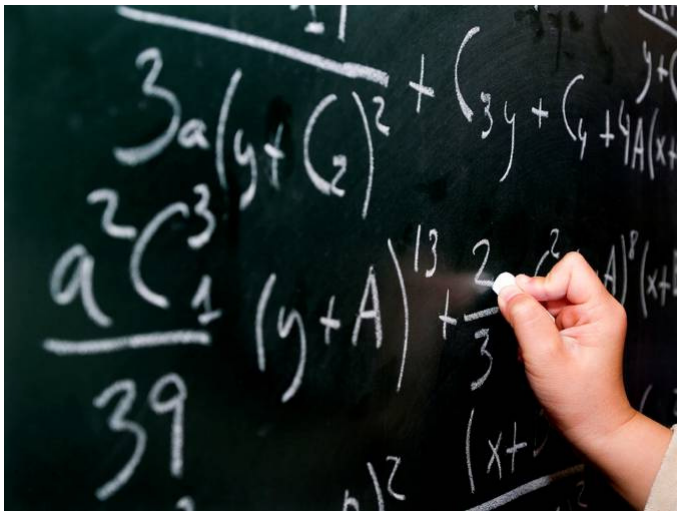


Lower Expansion Factor
= Compress once little used pages
= Near zero CPU cycles

Higher Expansion Factor
= More compression
= More CPU cycles

Balance more RAM verses more CPU cycles

Technical Details



Bad Compression Targets

- AIX Kernel
  - Not a AME target

- Filesystem cache, code or memory mapped files
  - Best to page out to filesystems
  - Performance tools → "numperm"

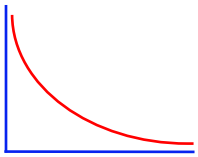
- Pinned Memory
  - Pinned = never page out (AME is like paging)
  - Performance tools → "pinned pages"
- So what can AME compress?

Good Compression Targets

- Mostly private pages within programs
  - Data
  - Heap
  - Stack
  - Not the code

Excellent Compression Targets

Data that compresses well
- Data only used on program initialisation
- Pages allocated but unused = full of zeros/blanks
- Pages with lots of repeat data like database records



Access Pattern
- Some **hot** pages, some **warm**, some **freezing**
- All pages equally used (HPC) – not so good

How can I work that out?

Do I have
→ Good or bad compression ratio?
→ Friendly or hostile access pattern?

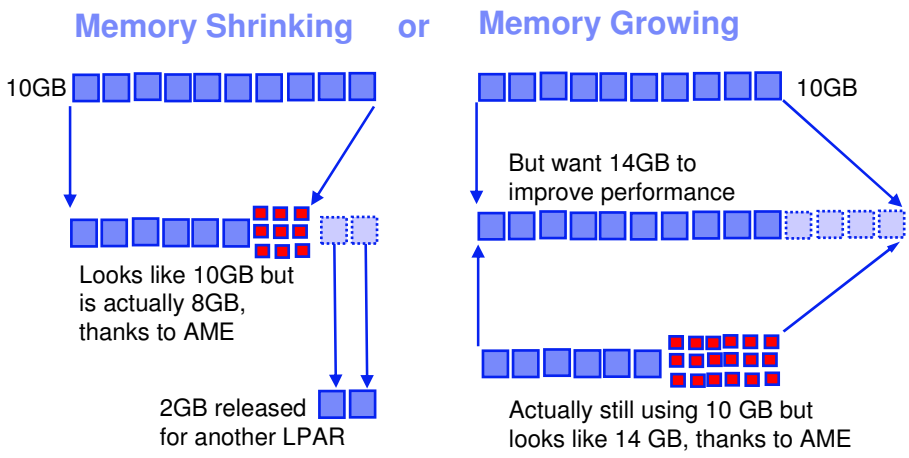
Normally, you can't !!

until now ….

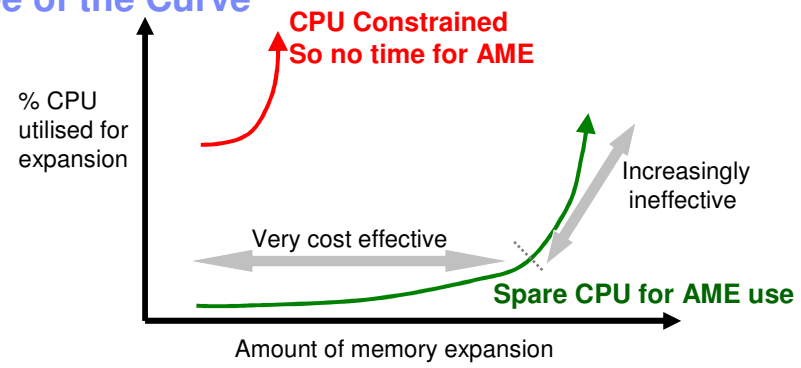
Planning for Active Memory Expansion

- A new AIX command: **amepat**
  - Active Memory Expansion Performance Analysis Tool
  - Or someone called Patrick/Patricia – you decide!
- Scans actual memory use
  - Determines compression ratio & CPU requirement
- With AME on or AME off
  - AIX 6.1 TL04 SP2+ also works on POWER4/5/6/7

What is your Plan?



Knee of the Curve



- Busy processor cores don't have resources to spare for AME
- The Expansion Factor "knee" depends on the compressibility of memory

amepat command - Don't worry as its easy

```
# amepat -?
Usage: amepat [-u minucomp_poolsize] [-m min_mem_savings]
            {[-t tgt_expmem_size] | [-a]} [-n num_entries] [-P recfile]
            {[-e startexpfactor[:stopexpfactor[:incexpfactor]]] |
                { [-c max_cpu_overhead%] | [-C max_cpu_overhead] }}
            [-v] { [Duration] | [interval [samples]] }

     amepat -R recfile { [Duration] | [Interval [Samples]] }

     amepat -N [{-P|-R} recfile] [-v] {Duration]|[Interval [Samples]]}

-m min_mem_savings     Unit is MB
-c max_cpu_overhead%   Unit is percentage
-C max_cpu_overhead    Unit is in number of Physical Processors
-u minucomp_poolsize   Unit is MB
-t tgt_expmem_size     Unit is MB
Duration               Unit is minutes
interval               Unit is minutes

Note: -N flag will turn off Active Memory Expansion Modeling.
      All options except -P, -R will be disabled when -N is used.
```

amepat - Basics are Easy

Run & report mode gets frustrating so …

Capture your busy hour for the whole hour

- amepat –R ame.out 60 [60 minutes]

Then try various reports

- Shrink memory:
  - amepat –P ame.out
- Expand memory:
  - amepat –P ame.out –t 4096 [target memory size in MB]

Warning ….

- Small micro-partition example
  - Less than a whole CPU & only 1 GB memory
  - Easier to generate workload to use all memory
- Typically, LPARs are much larger
  - Rule of Thumb: 8 -16 GB per CPU or higher
- Large memory LPARs will give AME more scope

amepat output – Machine Summary

```
# amepat
Date/Time of invocation        : -
Total Monitored time           : NA
Total Samples Collected        : NA
```

Machine Summary

```
System Configuration:
---------------------
Partition Name                 : diamond3
Processor Implementation Mode  : POWER7
Number Of Logical CPUs         : 16
Processor Entitled Capacity    : 0.80
Processor Max. Capacity        : 4.00
True Memory                    : 1.00 GB
SMT Threads                    : 4
Shared Processor Mode          : Enabled-Uncapped
Active Memory Sharing          : Disabled
Active Memory Expansion        : Enabled
Target Expanded Memory Size    : 1.00 GB
Target Memory Expansion factor : 1.00
```

1 GB = very small for my test case

Factor=1 → No Compression

Memory Summary

```
System Resource Statistics:                Current
---------------------------            ----------------
CPU Util (Phys. Processors)                 0.01 [  0%]
Virtual Memory Size (MB)                     790 [ 77%]
True Memory In-Use (MB)                      985 [ 96%]
Pinned Memory (MB)                           371 [ 36%]
File Cache Size (MB)                         179 [ 17%]
Available Memory (MB)                        184 [ 18%]
. . .
. . .
. . .
```

Not compressed by AME

amepat –P ame.out

1GB with ½ memory unused

Nothing Compressed as there is no need

```
-> REMOVED CONFIG DETAILS ABOVE HERE
. . .
AME Statistics:                            Current
---------------                        ----------------
AME CPU Usage (Phy. Proc Units)             0.00 [  0%]
Compressed Memory (MB)                         0 [  0%]
Compression Ratio                           2.28

Active Memory Expansion Modeled Statistics:
-------------------------------------------
Modeled Expanded Memory Size   :   1.00 GB
Average Compression Ratio      :   2.28
```

Good compression achievable

```
Expansion    Modeled True      Modeled              CPU Usage
Factor       Memory Size       Memory Gain          Estimate
---------    -------------     ------------------   ------------
   1.00        1.00 GB          0.00 KB [  0%]     0.00 [  0%]
   1.14      896.00 MB        128.00 MB [ 14%]     0.00 [  0%]
   1.33      768.00 MB        256.00 MB [ 33%]     0.00 [  0%]
```

Various combinations with increasing factor with decreasing RAM

No CPU use

```
Active Memory Expansion Recommendation:
---------------------------------------
The recommended AME configuration for this workload is to configure
the LPAR with a memory size of 768.00 MB and to configure a memory
expansion factor of 1.33.  This will result in a memory gain of 33%.
With this configuration, the estimated CPU usage due to AME is
approximately 0.00 physical processors, and the estimated overall peak
CPU resource required for the LPAR is 0.02 physical processors.
```

AME thinks remove 0.25 GB of unused RAM is OK

Note: it does no try below 512 MB – that is just too small.

amepat –P ame.out

1GB with 99% used

Nothing Compressed as there is no need

```
-> REMOVED CONFIG DETAILS ABOVE HERE
. . .
AME Statistics:                              Current
---------------                          ----------------
AME CPU Usage (Phy. Proc Units)             0.00 [  0%]
Compressed Memory (MB)                         0 [  0%]
Compression Ratio                           2.02

Active Memory Expansion Modeled Statistics:
-------------------------------------------
Modeled Expanded Memory Size    :   1.00 GB
Average Compression Ratio       :   2.02

Expansion     Modeled True      Modeled              CPU Usage
Factor        Memory Size       Memory Gain          Estimate
---------     -------------     ------------------   ------------
    1.00        1.00 GB           0.00 KB [  0%]     0.00 [  0%]
    1.14      896.00 MB         128.00 MB [ 14%]     0.34 [  8%]
    1.33      768.00 MB         256.00 MB [ 33%]     0.72 [ 18%]

Active Memory Expansion Recommendation:
---------------------------------------
The recommended AME configuration for this workload is to configure
the LPAR with a memory size of 896.00 MB and to configure a memory
expansion factor of 1.14.  This will result in a memory gain of 14%.
With this configuration, the estimated CPU usage due to AME is
approximately 0.34 physical processors, and the estimated overall peak
CPU resource required for the LPAR is 1.18 physical processors.
```

Good compression achievable

Various combinations
with increasing factor
with decreasing RAM
with rising CPU use

AME thinks remove 0.125 GB of RAM is OK with 0.34 CPU used

amepat –P ame.out -t 1536

How can I get to 1.5 GB?

Compression found for current memory content

```
-> REMOVED CONFIG DETAILS ABOVE HERE
. . .
AME Statistics:                              Current
---------------                          ----------------
AME CPU Usage (Phy. Proc Units)             0.02 [  1%]
Compressed Memory (MB)                        65 [  4%]
Compression Ratio                           2.04

Active Memory Expansion Modeled Statistics:
-------------------------------------------
Modeled Expanded Memory Size    :   1.50 GB
Average Compression Ratio       :   2.04

Expansion     Modeled True      Modeled              CPU Usage
Factor        Memory Size       Memory Gain          Estimate
---------     -------------     ------------------   ------------
    1.00        1.50 GB           0.00 KB [  0%]     0.00 [  0%]
    1.09        1.38 GB         128.00 MB [  9%]     0.00 [  0%]
    1.20        1.25 GB         256.00 MB [ 20%]     0.00 [  0%]
    1.33        1.12 GB         384.00 MB [ 33%]     0.13 [  3%]
    1.50        1.00 GB         512.00 MB [ 50%]     0.28 [  7%]

Active Memory Expansion Recommendation:
---------------------------------------
The recommended AME configuration for this workload is to configure
the LPAR with a memory size of 1.00 GB and to configure a memory
expansion factor of 1.50.  This will result in a memory gain of 50%.
With this configuration, the estimated CPU usage due to AME is
approximately 0.28 physical processors, and the estimated overall peak
CPU resource required for the LPAR is 0.85 physical processors.
```

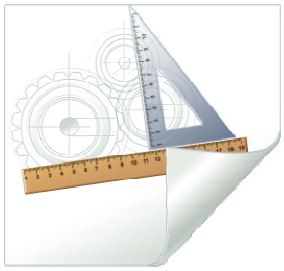
Estimating for 1.5GB

Various combinations
to get to 1.5GB
with decreasing RAM
with increasing CPU

AME thinks 0.28 CPU for 0.5 GB RAM is a good trade-off = last combination

You have to make up your own mind!

Monitoring Active Memory Expansion in use



lparstat -i

RAM= True Memory

Target
Expansion Factor
AME memory

```
# lparstat -i |pg
Node Name                                  : diamond3
Partition Name                             : diamond3 AME
Partition Number                           : 3
Type                                       : Shared-SMT-4
Mode                                       : Uncapped
Entitled Capacity                          : 0.80
Partition Group-ID                         : 32771
Shared Pool ID                             : 0
Online Virtual CPUs                        : 4
Maximum Virtual CPUs                       : 8
Minimum Virtual CPUs                       : 1
Online Memory                              : 1024 MB
Maximum Memory                             : 8192 MB
Minimum Memory                             : 512 MB
Variable Capacity Weight                   : 128
Minimum Capacity                           : 0.10
Maximum Capacity                           : 8.00
Capacity Increment                         : 0.01
Maximum Physical CPUs in system            : 8
Active Physical CPUs in system             : 8
Active CPUs in Pool                        : 6
Shared Physical CPUs in system             : 6
Maximum Capacity of Pool                   : 600
Entitled Capacity of Pool                  : 500
Unallocated Capacity                       : 0.00
Physical CPU Percentage                    : 20.00%
Unallocated Weight                         : 0
Memory Mode                                : Dedicated-Expanded
Total I/O Memory Entitlement               : -
Variable Memory Capacity Weight            : -
Memory Pool ID                             : -
Physical Memory in the Pool                : -
Hypervisor Page Size                       : -
Unallocated Variable Memory Capacity Weight: -
Unallocated I/O Memory entitlement         : -
Memory Group ID of LPAR                    : -
Desired Virtual CPUs                       : 4
Desired Memory                             : 1024 MB
Desired Variable Capacity Weight           : 128
Desired Capacity                           : 0.80
Target Memory Expansion Factor             : 1.50
Target Memory Expansion Size               : 1536 MB
#
```

vmstat –c 1

```
^C# vmstat -c 1

System configuration: lcpu=16 mem=1536MB tmem=1024MB ent=0.80 mmode=dedicated-E

kthr                      memory                              page              faults                cpu
------- ------------------------------------------- ------------------ ------------------ -----------------------
 r   b        avm        fre        csz        cfr        dxm    ci    co    pi    po    in    sy    cs us sy id wa    pc    ec
 1   1     358101     124511      18565       5692          0  5162  5139     0     0    44    49   513 69 10 21  0  1.55 193.2
 1   0     358101     124557      18565       5750          0  6151  6143     5     0   148   150   664 62 16 21  1  1.51 188.3
 1   0     358102     124499      18565       5681          0 11911 12009     0     0    68    41   918 64 14 22  0  1.45 181.8
 1   1     358102     124740      18565       5857          0  7606  7220     0     0     1    31   457 65 14 21  0  1.46 182.3
 4   0     358102     124800      18565       5877          0  9169  9145     0     0    45    41   633 61 18 20  1  1.52 189.7
 1   0     358102     124791      18565       5857          0  6191  6250     0     0    22    32   459 66 13 20  1  1.44 180.1
 1   0     358102     124810      18565       5894          0  4569  4478     0     0    20    32   426 71  8 20  1  1.42 177.0
 2   0     358102     124777      18565       5786          0  3384  3577     1     0    46    41   410 70  9 21  0  1.48 184.8
 2   0     358102     124752      18565       5833          0  3322  3219     0     0    34    45   409 73  6 21  0  1.44 179.8
 2   0     358102     124416      18565       5695          0  2564  2823     0     0   165   102   613 72  8 18  2  1.43 179.0
 1   0     358101     124479      18565       5646          0  2576  2706     0     0    13    31   309 69  8 23  0  1.50 186.9
 1   0     358101     124538      18565       5774          0  3723  3479     0     0    72    90   451 72  7 21  1  1.45 181.0
 1   0     358101     124546      18565       5763          0  4135  4156     0     0    26    34   437 71  8 21  0  1.46 182.2
 1   0     358101     124519      18565       5750          0  4331  4324     0     0   105    61   593 65 13 22  1  1.49 186.1
 1   1     358101     124530      18565       5703          0  8336  8393     0     0    67    54   803 60 15 24  0  1.50 187.4
 1   0     358101     124699      18565       5773          0  6073  5997     0     0    28    32   551 66 13 20  1  1.45 181.5
 1   1     358111     124717      18565       5827          0  5412  5272     0     0    67   216   566 66 12 22  0  1.48 184.8
 1   0     358111     124726      18565       5812          0  3594  3652     0     0    18    32   413 67  9 24  0  1.51 189.0
 1   0     358111     124818      18565       5839          0  4273  4228     0     0    11    36   426 59 10 30  0  1.58 197.8
 1   0     358111     124631      18565       5787          0  4898  4989     0     0   116    61   679 61 10 29  1  1.56 194.8
```

mem = apparent Memory
tmem = True Memory

CI = Compressed Page In
CO = Compressed Page Out

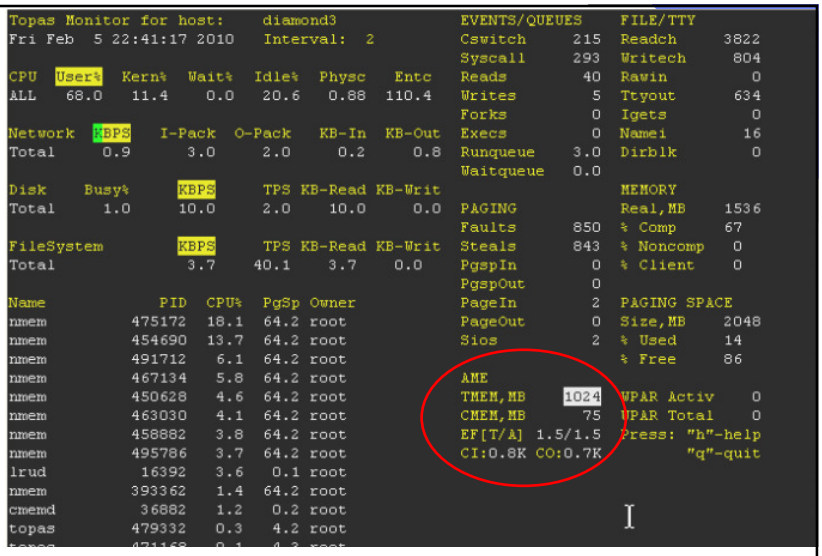
lparstat –c 1

%user	%sys	%wait	%idle	physc	%entc	lbusy	vcsw	phint	%xcpu	dxm
83.2	0.8	0.0	16.0	1.02	126.9	10.3	1545	1	0.0	0
93.6	0.6	0.0	5.7	1.02	127.1	12.3	1542	0	0.0	0
93.8	0.6	0.0	5.5	1.01	126.8	12.3	1527	6	0.0	0
93.8	0.7	0.0	5.5	1.01	126.8	12.4	1554	6	0.0	0
94.0	0.5	0.0	5.5	1.01	126.0	12.6	1543	5	0.0	0
93.9	0.6	0.0	5.5	1.01	126.6	12.4	1563	5	0.0	0
82.0	12.4	0.0	5.6	1.02	126.9	12.5	1576	0	11.3	0
81.0	13.0	0.0	6.1	1.02	127.3	12.4	1611	2	11.4	0
76.7	11.2	0.0	12.2	1.00	124.8	10.9	1551	3	14.4	0
77.5	8.9	0.0	13.6	1.03	128.4	10.9	1554	5	11.7	0
83.7	4.3	0.0	12.0	1.01	126.8	11.8	1559	9	5.3	0
82.9	5.4	0.0	11.7	1.01	126.7	11.1	1533	6	4.6	0
81.9	7.3	0.0	10.8	0.90	112.1	10.0	1561	4	7.7	0
72.9	9.2	0.0	17.9	0.98	123.1	9.5	1568	2	13.0	0
73.1	12.9	0.0	14.0	0.97	121.1	10.3	1573	8	18.8	0
76.5	9.4	1.6	12.6	0.96	119.9	10.1	1616	7	12.6	0
79.8	3.4	0.1	16.7	0.97	121.8	9.8	1578	0	3.6	0
80.6	1.2	0.1	18.2	0.98	122.1	9.5	1601	6	0.3	0
80.1	0.6	0.0	19.3	0.97	121.8	9.1	1546	5	0.0	0
81.2	1.0	0.0	17.8	0.82	103.1	8.1	1535	1	0.1	0
80.5	0.7	0.0	18.8	0.86	107.3	8.4	1550	1	0.0	0

%xcpu percentage of CPU time used in eXpansion!

Note: %user + %sys + %wait + %idle still = 100%

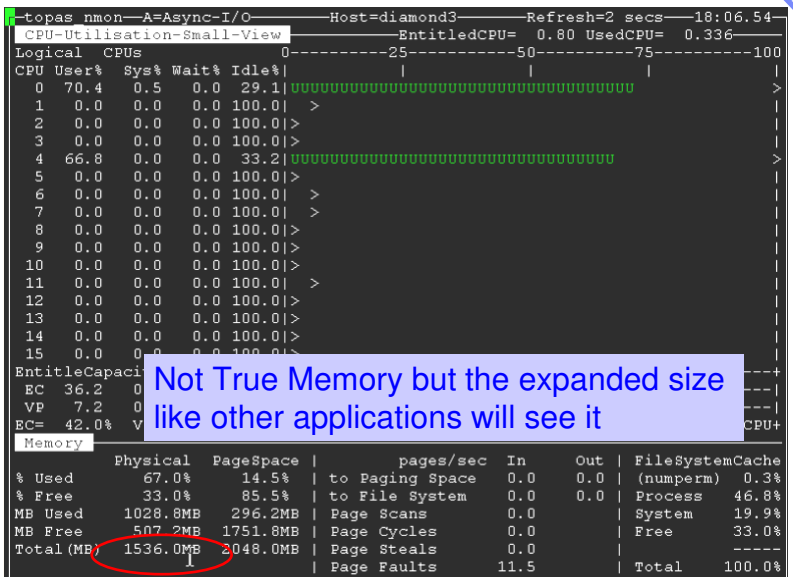
topas



TMEM = True Memory
CMEM = Compressed Memory
CI = Compressed Page In
CO = Compressed Page Out

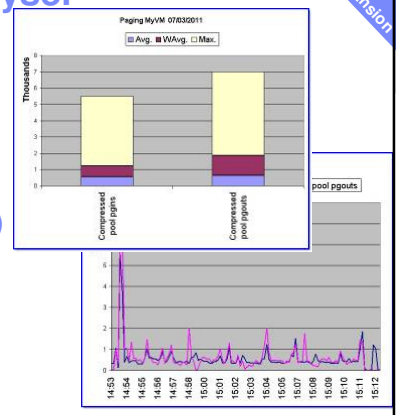
EF = Expansion Factor
T = target
A = Actual

topas_nmon not aware of AME (yet!)



Not True Memory but the expanded size like other applications will see it

nmon capture to file then Analyser



- MEM tab
  - Size of the Compressed pool (MB)
  - Size of true memory (MB)
  - Expanded memory size (MB),
  - Size of the Uncompressed pool (MB)
- MEMNEW tab
  - Compressed Pool%
- PAGE tab
  - Paging rates (pages per second) but this time very quickly in & out of the compressed memory area
  - Compressed pool pgins - other tools like topas call this CI
  - Compressed pool pgouts - other tools like topas call this CO

Topas AME - Very small change

```
Topas Monitor for host:green3                EVENTS/QUEUES    FILE/TTY
Thu Oct  6 17:48:40 2011   Interval:2        Cswitch      290 Readch     3703
                                             Syscall      219 Writech     287
CPU      User% Kern% Wait% Idle%   Physc  Entc% Reads         29 Rawin          0
Total    99.8   0.2   0.0   0.0    2.00 199.91 Writes         0 Ttyout        287
                                             Forks          0 Igets           0
Network     BPS  I-Pkts  O-Pkts    B-In   B-Out Execs          0 Namei           7
Total    1.20K   10.00    1.50   836.0   392.0 Runqueue    9.00 Dirblk          0
                                             Waitqueue    0.0
Disk    Busy%      BPS      TPS  B-Read  B-Writ                   MEMORY
Total     0.0        0        0       0       0 PAGING           Real,MB    2048
                                             Faults         0 % Comp        53
FileSystem          BPS      TPS  B-Read  B-Writ Steals         0 % Noncomp     40
Total             3.62K    28.50   3.62K       0 PgspIn         0 % Client      40
                                             PgspOut        0
WLM-Class (Active)       CPU%    Mem%  Blk-I/O% PageIn         0 PAGING SPACE
System                      0      65         0 PageOut        0 Size,MB       512
Shared                      0       5         0 Sios           0 % Used          1
                                                               % Free         99
Name           PID  CPU%  PgSp Class
ncpu       7536784 25.0  148K wp06              AME
ncpu       9044056  0.0  144K wp06              TMEM        2.00G PAR Activ     2
ncpu       8388632  0.0  108K wp06              CMEM      512.00M PAR Total     2
ncpu       9306134  0.0  136K wp06              EF[T/A]    1.0/1.0 Press: "h"-help
                                                CI:     0 CO:     0        "q"-quit
```

- Units have moved
  - TMEM, MB 2000 → TMEM 2.00G

svmon

- Good luck with that one !

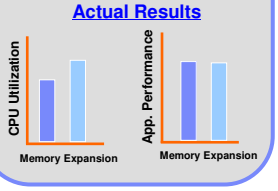
AME Deployment Steps

① Planning Tool

A. amepat part of AIX 6.1 TL4
B. Calculates data compressibility & estimates CPU overhead due to AME
C. Provides initial recommendations

② 60-Day Trial

A. One-time, temporarily enablement
B. Config LPAR based on planning tool
C. Use AIX tools to monitor AME environment
D. Tune based on actual results



③ Deploy into Production

A. Permanently enable AME
B. Deploy workload into production
C. Continue to monitor workload using AIX performance tools

Active Memory Expansion

AME Wiki page & AME Forum

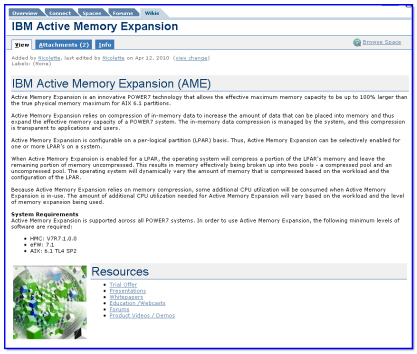
AME Public Wiki (on the AIX wiki)

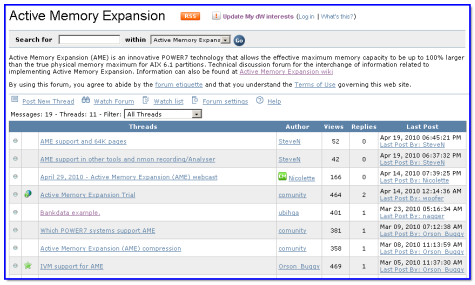
http://www.ibm.com/developerworks/wikis/display/WikiPtype/IBM+Active+Memory+Expansion

URL for trial, presentation, Forum, whitepapers, manual page, Perf Tune Guide, movies

AME Forum

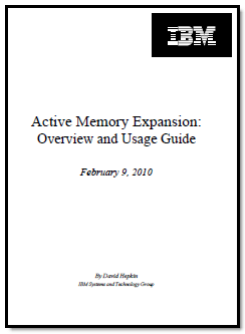
http://www.ibm.com/developerworks/forums/forum.jspa?forumID=2179

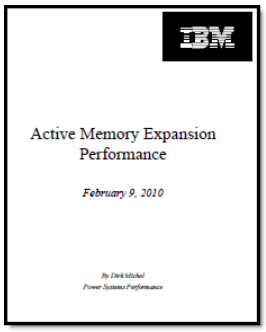




Active Memory Expansion

Docs

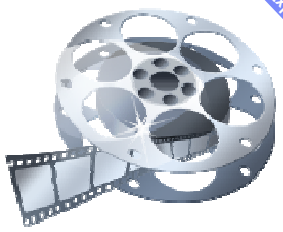




1) AME Overview & Usage Guide by David Hepkin 25 pages
2) AME Performance by Dirk Michel 18 pages
From
http://www.ibm.com/systems/power/resources/index.html
Then click on Whitepapers

AIX Commands Infocenter → amepat, topas, vmstat, lparstat

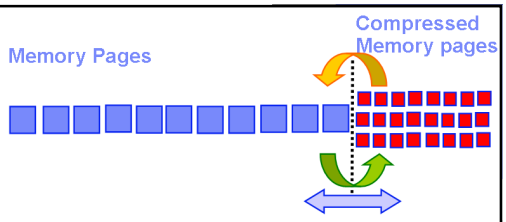
AME – at the Movies



http://tinyurl.com/AIXmovies

Look for Movie 79

Summary & Questions



1. POWER7 & AIX 6 TL04 SP2+
2. Activation Key by machine
3. 60 day Trial available
4. Set at LPAR level & dynamic Expansion Factor
5. Planning: **amepat**
6. Monitor: vmstat -c, lparstat –c, topas

http://www.ibm.com/developerworks/wikis/display/WikiPtype/ … IBM+Active+Memory+Expansion