IBM Power Systems Technical University

October 18–22, 2010 — Las Vegas, NV

Session Title: Introduction to System Monitoring

Session ID: SM09

Speaker Name: Grover Davidson

grover@us.ibm.com

Authorized IBM Training

Objectives

- Why systems should be monitored and what are the advantages
- AIX Logging mechanisms
  - alog, system error log, syslog, sulog, utmp/wtmp, .history file, tunables
- Performance Monitoring
  - vmstat, iostat, netstat, nfsstat, lsps, svmon, sar, topas
- Monitoring for WLM
- VIO Monitoring
- Monitoring packages
- Top Monitoring

System Monitoring and Advantages

- Successful monitoring involves:
  - Periodically obtaining performance-related information from the operating system
  - Storing the information for future use in problem diagnosis
  - Displaying the information for the benefit of the system administrator
  - Detecting situations that require additional data collection or responding to directions from the system administrator to collect such data, or both
  - Tracking changes made to the system and applications
- Advantages to continuously monitoring system performance.
  - Sometimes detect underlying problems before they have an adverse effect
  - Detect problems that affect a user's productivity
  - Collect data when a problem occurs for the first time
  - Allow you to establish a baseline for comparison

Logging Mechanisms: alog

- Check for console/boot messages
- Default mechanism used by AIX for logging console messages
- Active even during the boot phase
- To determine available logs:

```
#  alog -L
boot
bosinst
nim
console
cfg
lvmcfg
lvmt
dumpsymp
```

- Determine location of log file (boot log in this case):

```
# alog -L -t boot
#file:size:verbosity
/var/adm/ras/bootlog:131072:1
```

- 'alog –f /var/adm/ras/bootlog –o' will display log contents

Logging Mechanisms: alog

- alog –o –t cfg can be run to display any captured config log data
- Config logging is enabled by default and the related tunable is in ODM in SWservAt. The environment variable CFGLOG is used to select log types.

```
# odmget -q "attribute like cfg_*" SWservAt

SWservAt:
        attribute = "cfg_logname"
        deflt = "/var/adm/ras/cfglog"
        value = "/var/adm/ras/cfglog"
SWservAt:
        attribute = "cfg_logsize"
        deflt = "1048576"
        value = "1048576"
SWservAt:
        attribute = "cfg_logverb"
        deflt = "1"
        value = "1"
```

- New option with alog is the lvm logging option
  - alog –t lvmcfg –o
  - alog –t lvmt -o
- Alog command works with log files specified in configuration file or can be user specified
- The file, size, verbosity of each defined Logtype can be changed using odmadd command

Logging Mechanisms: system error log

- AIX default mechanism for reporting errors and notifications
- Error information is written to the /dev/error special file
- errdemon constantly checks this file for new entries and when new data is written, compares the label sent by the kernel or application code to the contents of the Error Record Template Repository
- Error logging is initialized when the kernel is loaded. The errdemon is started by **rc.boot** script during system initialization and is automatically stopped by the **shutdown** script during system shutdown
- Records hardware/software messages
- The diag command uses error log to diagnose PERM hardware problems and ignores all hardware INFO type of errors
- Default length of time that hardware error entries remain in the error log is 90 days and other errors is 30 days
- Wraps and fixed size – managed with errdemon command
- Consider using error notification for important events

Logging Mechanisms: system error log

- errclear command runs from root's crontab and clears any SW and errlogger message more than 30 days old and HW messages more than 90 days old
- /usr/lib/errdemon –l lists the current settings
- Error logging can be customized using the errdemon command
- Error log file location, size and error log device driver's internal buffer can be changed.
- errpt –t –F Report=0 lists all events for which reporting is currently disabled
- Use 'errpt' to see condensed log or 'errpt –a' for complete details of each entry

```
# errpt
A63BEB70   0828144907 P S SYSPROC        SOFTWARE PROGRAM ABNORMALLY TERMINATED
A2205861   0828094807 P S SYSPROC        Excessive interrupt disablement time
A63BEB70   0827160007 P S SYSPROC        SOFTWARE PROGRAM ABNORMALLY TERMINATED
A63BEB70   0827091307 P S SYSPROC        SOFTWARE PROGRAM ABNORMALLY TERMINATED
09890235   0803070707 P H hdisk11        ARRAY DRIVE FAILURE
3074FEB7   0802172507 T H fscsi0         ADAPTER ERROR
3074FEB7   0802172507 T H fscsi0         ADAPTER ERROR
```

Logging Mechanisms: system error log

- Reading an error report:
  - **LABEL -** Predefined name for the event.
  - **ID -** Numerical identifier for the event.
  - **Date/Time -** Date and time of the event.
  - **Sequence Number -** Unique number for the event.
  - **Machine ID**- identification number of your system processor unit.
  - **Node ID -** Mnemonic name of your system.
  - **Class -** General source of the error. The possible error classes are:
    - **H -** Hardware. (When you receive a hardware error, refer to your system operator guide for information about performing diagnostics on the problem device or other piece of equipment. The diagnostics program tests the device and analyzes the error log entries related to it to determine the state of the device.)
    - **S -** Software.
    - **O -** Informational messages.
    - **U -** Undetermined (for example, a network).

Logging Mechanisms: system error log

- Reading an error report:
  - **Type -** Severity of the error that has occurred. The following types of errors are possible:
    - **PEND -** The loss of availability of a device or component is imminent.
    - **PERF -** The performance of the device or component has degraded to below an acceptable level.
    - **PERM -** Condition that could not be recovered from. Error types with this value are usually the most severe errors and are more likely to mean that you have a defective hardware device or software module. Error types other than PERM usually do not indicate a defect, but they are recorded so that they can be analyzed by the diagnostics programs.
    - **TEMP -** Condition that was recovered from after a number of unsuccessful attempts. This error type is also used to record informational entries, such as data transfer statistics for DASD devices.
    - **UNKN -** It is not possible to determine the severity of the error.
    - **INFO -** The error log entry is informational and was not the result of an err

Logging Mechanisms: system error log

```
---------------------------------------------------------------------------
LABEL:          FCP_ARRAY_ERR14
IDENTIFIER:     09890235

Date/Time:      Fri Aug  3 07:07:36 2007
Sequence Number: 51630
Machine Id:     00CADD5D4C00
Node Id:        HOSTA
Class:          H
Type:           PERM
Resource Name:  hdisk11
Resource Class: disk
Resource Type:  array
Location:       U7879.001.DQDLYTR-P1-C3-T1-W200A00A0B80FCC28-L5000000000000

Description
ARRAY DRIVE FAILURE

Probable Causes
ARRAY DASD DEVICE
ARRAY DASD MEDIA

Failure Causes
ARRAY DASD MEDIA

        Recommended Actions
        PERFORM PROBLEM DETERMINATION PROCEDURES

Detail Data
SENSE DATA
0600 0308 0000 FF00 0000 0004 0000 0000 0000 0000 0000 0000 0000 7000 0600
0000 0098 0000 0000 3F80 4700 0000 0000 0000 0000 0000 000D 0000 0000 0000 0000
0008 0500 0000 0000 0000 0000 0000 0000 0000 3154 3332 3439 3639 3134 2020 2020
2020 0623 0500 0005 0100 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000
0005 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000
0000 0000 0338 D05D 3038 3033 3037 2F30 3534 3830 3100 0000 0000 0000 0000 0000
0000 0000 4A99 9000 F205 3402 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000
0000 0000
```

Logging Mechanisms: system error log

- Works by error data being saved in nvram and then retrieved by errdemon process
- Some messages are informational – 'error logging started' for example
- 'temporary' errors that occur repeatedly may not be temporary and should be explained/resolved
- errlogger command allows the system administrator to record messages in the error log
- ras_logger command provides a way to log any error from the command line and can be used to test newly created error templates and provides a way to log an error from a shell script

Alerting with errnotify

- Tied into the error log system
- User specifies these conditions and actions in an Error Notification object.
- Each time an error is logged, the **error notification** daemon determines if the error log entry matches the selection criteria of any of the Error Notification objects
- Error Notification object class is located in the **/etc/objrepos/errnotify** file
- Error Notification objects are added to the object class by using Object Data Manager (ODM) commands
- Only processes running with the root user authority can add objects to the Error Notification object class
- Error notification will fail if new processes cannot be created
- Experiment and test to make certain it works as desired using errlogger to inject errors

Logging Mechanisms: syslog

- syslog is started in /etc/rc.tcpip by default on AIX
- Entries controlled by /etc/syslog.conf – highly configurable
- Error log messages and syslog messages can be listed in a single report by specifying errlog as the destination in the syslog.conf file
- Each message is one line and messages longer than 900 bytes may be truncated
- If rotate option is specified, then the log files are limited by **size** or **time or both** and will be rotated once the log file reaches either the size limit or time limit or whichever is earlier
- If compress option is specified then the rotated log files that are not in use will be compressed and log file names are generated with ".Z" extension
- 'lssrc –l –s syslogd' shows current configuration:

```
Subsystem         Group            PID          Status
 syslogd          ras              18894        active
```

- Used by numerous demons to log information
- Applications can use libc API to write to syslog
- 'logger' is command line interface for syslog
- Log files MUST exist when syslogd starts – use 'touch /tmp/syslog.out' in this example
- Can be sent to remote hosts for security purposes using /etc/syslog.conf

Logging Mechanisms: sulog

- Default mechanism to track all attempts to change user ID with su command
- Flat text file readable only by root: /var/adm/sulog

```
SU 03/06 14:28 + pts/1 root-test
SU 03/11 12:37 + pts/5 root-test
SU 03/13 23:40 - pts/6 root-test
SU 03/13 23:41 + pts/6 root-test
SU 03/14 16:05 + pts/6 root-test
```

- One line per su attempt, +/- symbol indicates success or failure to su
- Records original user and target user
- Should be reviewed regularly
- Must be cleared manually

Logging Mechanisms: utmp

- Part of accounting system – contains user accounting records
- *utmpd* daemon monitors the */etc/utmp* file for validity of the user process entries at regular intervals
- Formatted with 'who –a /etc/utmp' (old=not active in the last 24 hours):

```
   -         system boot Jul 17 07:16
   -         run-level 2 Jul 17 07:16
srcmstr   -      .        Jul 17 07:16     0:23
LOGIN     - tty0          Jul 17 07:16     old
LOGIN     - tty1          Jul 17 07:16     old
rcafs     -      .        Jul 17 07:18     0:23
```

- Can be useful to resolve inittab issues (who –d /etc/utmp)

```
   .            .        Jul 17 07:16     0:20     18370 id=rcnfs   term=0 exit=0
   .            .        Jul 17 07:18     0:20     29438 id=rcdcecl term=0 exit=80
   .            .        Jul 17 07:18     0:20     31038 id=nimclie term=0 exit=11
   .            .        Jul 17 07:18     0:20      6250 id=piobe   term=0 exit=0
```

- Similar to wtmp file
- Records:
  - exit status of inittab entries
  - reboots
  - Logins and logouts

Logging Mechanisms: wtmp

- Part of the accounting system – contains connect-time accounting records
- wtmp will grow unless either processed by accounting or cleared with:

```
cp /dev/null /var/adm/wtmp
```

- Other ways to clear wtmp may result in a variety of problems
- Formatted with 'last' command which shows login information back to the beginning of the wtmp file

```
root     pts/0        xxx.yyy.ibm.com      Feb 09 05:16 - 05:16  (00:00)
root     pts/0        xxx.yyy.ibm.com      Feb 09 05:16 - 05:16  (00:00)
root     pts/0        xxx.yyy.ibm.com      Feb 08 05:16 - 05:16  (00:00)
root     pts/0        xxx.yyy.ibm.com      Feb 08 05:16 - 05:16  (00:00)

wtmp begins      Jan 22 16:39
```

- Check when unauthorized access is suspected
- Information about terminations due to rebooting and sessions that are still continuing is included if applicable

Logging Mechanisms: history file

- Default log for all commands a user runs
- Flat text file
- Located in their home directory
- May contain evidence of what has happened to explain unexpected events/behavior
- Can be used to recover previously executed commands
- Determined by the HISTFILE environment variable
- If unset, uses $HOME/.sh_history
- Limited size controlled by HISTSIZE environment variable – defaults:
  - Root user – last 512 commands
  - Non-root – last 128 commands
- Used by ksh – other shells may use other files

Logging Mechanisms: Tuning files

- Introduced in AIX 5.2
- Stored in /etc/tunables directory
- lastboot.log contains log of tunable changes/errors from the last time the system was booted
- nextboot file shows tunables that will be applied at the next time the system is booted
- Some tunables stored in ODM SWservAt and are zapped into kernel during bosboot
- Changed with vmo/ioo/schedo/no/nfso using ‘-p’ or ‘-r’ flags.

Performance Monitoring with vmstat

- Displays cpu and virtual memory statistics
- Lightweight – very little overhead
- First report displays cumulative activity since the last system boot. The second report shows activity for the first 5-second interval.
- Quick summary of
  - Total active virtual memory used by all the processes in the system
  - Number of real memory page frames on the free list
  - Paging space activity
  - LRU activity
- Adding a timestamp by using the ‘-t’ option makes it easy to find data at a specific time of day
- If -@ flag is specified, report consists of system and WPAR configuration

Performance Monitoring with vmstat

- File IO statistics can be included (vmstat –I -l):

```
# vmstat -I -l

System configuration: lcpu=4 mem=8191MB ent=0.50

kthr memory              page                    faults          cpu                 large-page
----- ---------- ------------------------------ ----------- ---------------------      ----------
 r  b avm    fre          re pi  po  fr    sr  cy  in sy  cs us sy id wa pc    ec      alp  flp
 1 1 347391 1787037 0  0   0   0     0   0   8  36 98  0 99  0  0   0.06 11.4 0    0
```

Performance Monitoring with iostat

- Reports statistics for
  - CPU
  - Asynchronous input/output (AIO)
  - TTY devices
  - Adapters
  - Disk CD-ROMs
  - Tapes
  - Filesystems
- Generates 4 utilization reports
  - TTY and CPU
  - Disk/tape
  - File system
  - System and adapter throughput
- %tm_act can be misleading
  - 100% active means there is 1 OR MORE commands in the disk queue each time it was sampled.
  - Does not measure queue full conditions (use iostat –Dl)
- iowait is a special form of *idle* time and should be considered as idle time
- Can help identify hotspots, IO patterns and other IO related issues

Performance Monitoring with iostat

- 'iostat –D' monitors physical disk IO timing and queuing to the physical disk itself:

```
--------------------------------------------------------------------------------
hdisk0          xfer:  %tm_act      bps      tps      bread      bwrtn
                        10.2    134.9K    24.8        0.0      134.9K
                read:      rps  avgserv  minserv  maxserv   timeouts      fails
                         0.0      0.0      6.7     13.6           0          0
               write:      wps  avgserv  minserv  maxserv   timeouts      fails
                        24.8      9.9      2.0     39.2           0          0
               queue:  avgtime  mintime  maxtime  avgwqsz    avgsqsz     sqfull
                       487.9      0.0    945.7      2.1        0.1       19.4
hdisk2          xfer:  %tm_act      bps      tps      bread      bwrtn
                         0.0      0.0      0.0        0.0        0.0
                read:      rps  avgserv  minserv  maxserv   timeouts      fails
                         0.0      0.0      0.0      0.0           0          0
               write:      wps  avgserv  minserv  maxserv   timeouts      fails
                         0.0      0.0     11.1     11.1           0          0
               queue:  avgtime  mintime  maxtime  avgwqsz    avgsqsz     sqfull
                         0.0      0.0      0.0      0.0        0.0        0.0
```

Performance Monitoring with netstat

- Useful in determining the number of sent and received packets
- Displays information regarding traffic on the configured network interfaces, like:
  - The address of any protocol control blocks associated with the sockets and the state of all sockets
  - The number of packets received, transmitted, and dropped in the communications subsystem
  - Cumulative statistics per interface
  - Routes and their status
- 'netstat –v' dumps adapter card statistics – output varies on adapter type

```
ETHERNET STATISTICS (ent0) :
Device Type: 2-Port 10/100/1000 Base-TX PCI-X Adapter (14108902)
Hardware Address: 00:00:00:00:00:aa
Elapsed Time: 156 days 0 hours 56 minutes 11 seconds
Transmit Statistics:                          Receive Statistics:
--------------------                          -------------------
Packets: 1327178089                           Packets: 2692469327
Bytes: 438185132440                           Bytes: 3569019395354
Interrupts: 0                                 Interrupts: 976134763
Transmit Errors: 0                            Receive Errors: 0
Packets Dropped: 0                            Packets Dropped: 0
                                              Bad Packets: 0

Max Packets on S/W Transmit Queue: 51
S/W Transmit Queue Overflow: 0
Current S/W+H/W Transmit Queue Length: 1

Broadcast Packets: 14436                      Broadcast Packets: 69525941
Multicast Packets: 0                          Multicast Packets: 638
No Carrier Sense: 0                           CRC Errors: 0
DMA Underrun: 0                               DMA Overrun: 0
Lost CTS Errors: 0                            Alignment Errors: 0
```

Performance Monitoring with netstat

- 'netstat –s' shows per-protocol statistics including ip/tcp/udp/icmp/igmp/ip6:

```
icmp:
        395551 calls to icmp_error
        0 errors not generated because old message was icmp
igmp:
        638 messages received
        0 messages received with too few bytes
tcp:
        1202513342 packets sent
                170687741 data packets (3965307727 bytes)
                343196 data packets (485968042 bytes) retransmitted
udp:
        151102808 datagrams received
        0 incomplete headers
        0 bad data length fields
        0 bad checksums
        395551 dropped due to no socket
        404818 broadcast/multicast datagrams dropped due to no socket
        0 socket buffer overflows
```

- Check for tcp retransmissions to indicate dropped packets – target of less than 0.5% retransmitted reasonable
- Check for udp socket overflows – add udp receive space or demons to process udp packets

Performance Monitoring with nfsstat

- Provides NFS client/server/rpc statistics (nfsstat –csnr):

```
Client nfs:
calls       badcalls    clgets      cltoomany
13055660    3           0           0
Version 2: (12167953 calls)
null        getattr     setattr     root        lookup      readlink    read
0 0%        5641402 46% 1410498 11% 0 0%        1559777 12% 0 0%        315112 2%
wrcache     write       create      remove      rename      link        symlink
0 0%        3110987 25% 16963 0%    1213 0%     16485 0%    0 0%        0 0%
mkdir       rmdir       readdir     statfs
0 0%        0 0%        94416 0%    1100 0%
Version 3: (887797 calls)
null        getattr     setattr     lookup      access      readlink    read
0 0%        228374 25%  782 0%      320955 36%  155638 17%  0 0%        48510 5%
write       create      mkdir       symlink     mknod       remove      rmdir
38925 4%    600 0%      1 0%        0 0%        0 0%        41641 4%    0 0%
rename      link        readdir     readdir+    fsstat      fsinfo      pathconf
1 0%        0 0%        17 0%       48093 5%    3741 0%     5 0%        0 0%
commit
514 0%
```

Performance Monitoring with nfsstat

- Can provide per-mount point options including timer values used retransmission with udp (nfsstat –m):

  /pages/cat from HOSTA/m:

  Flags: vers=3,**proto=tcp**,auth=unix,soft,link,symlink,rsize=32768,wsize=32768,retrans=5

  All: srtt=0 (0ms), dev=0 (0ms), cur=0 (0ms)

  /user5 from HOSTB:/ltest

  Flags: vers=2,**proto=udp**,auth=unix,soft,dynamic,rsize=8192,wsize=8192,retrans=5

  Lookups: srtt=7 (17ms), dev=3 (15ms), cur=2 (40ms)

  Reads: srtt=3 (7ms), dev=3 (15ms), cur=1 (20ms)

  All: srtt=7 (17ms), dev=3 (15ms), cur=2 (40ms)

- Can be used to:
  - Determine if NFS is dropping packets (udp only)
  - Display dynamic timers for retransmission (udp only)
  - Analyze the traffic patterns (read/write/directory ops, etc)

Performance Monitoring with sar

- System Activity Recorder – part of accounting
- Can collect statistics on almost anything (no network information)
- Reports either system-wide CPU statistics or for each individual processor
- Requires configuring sadc – sar data collector
- Normally configured in adm users crontab
- Records data in binary format
- Binary data is formatted with sar or sa2 command
- See your favorite admin book for more details – too extensive to cover here

Performance Monitoring with svmon

- Captures and analyzes a snapshot of virtual memory
- Displays information about the current state of memory, though the displayed information does not constitute a true snapshot of memory since command runs at user level with interrupts disabled
- 'svmon –G' displays global summary:

```
# svmon -G

                size       inuse        free         pin         virtual
memory       2097136      379316     1787468      241082      347087
pg space      131072        2669

              work       pers      clnt       other
pin           103316         0         0      137766
in use        347087         0     32229

PageSize  PoolSize      inuse       pgsp        pin      virtual
s    4 KB         -     235220       2669     167082       202991
m   64 KB         -       9006          0       4625         9006
```

- Statistics reported are expressed in terms of pages

Performance Monitoring with lsps

- Monitors paging space
- 'lsps –a' shows actual paging space assigned and in use

```
# lsps -a
Page Space Physical Volume Volume Group Size %Used Active Auto Type Chksum
hd6        hdisk0          rootvg      1024MB 1    yes    yes    lv  0
```

- 'lsps –s' shows paging space in use and reserved

```
# lsps -s
Total Paging Space   Percent Used
      1024MB                 1%
```

- For NFS paging spaces, the PV and VG name will be replaced by host name of NFS server and path name of the file that is used for paging
- May show different % used if deferred paging space allocation policy is not in use
- Running out of paging space can prevent new commands from being forked
- If AVM (vmstat command) exceeds real memory, paging cannot be stopped

Performance Monitoring with topas

- Displays
  - Overall system statistics
  - List of busiest processes (-p)
  - WLM statistics (-W)
  - List of hot physical disks (-D)
  - Logical partition display (-L)
  - File system (-F)
  - LVM (-V)
  - Cross-Partition View (AIX® 5.3 with 5300-03 and higher)
- Default output consists of 2 fixed parts and a variable section
- Disks and network adapters added after starting **topas** or any other SPMI consumer will not be reflected in **topas**

Monitoring for WLM

- wlmstat
  - Shows per class resource utilization statistics
  - Displays the contents of WLM data structures retrieved from the kernel
  - Provides a per-second view of WLM activity hence it is not suited for the long-term analysis
- wlmmon, wlmperf
  - Provide graphical views of Workload Manager (WLM) resource activities by class
  - Provide reports of WLM activity over much longer time periods, with minimal system impact
  - wlmmon generates reports only for the latest 24-hour period and has no usage options
  - wlmperf can generate reports from trend recordings made by the PTX daemons for periods covering minutes, hours, days, weeks, or months

Monitoring for WLM

- xmwlm
  - Introduced in perfagent.tools 5.3.0.50
  - Records performance data (binary format) by default in /etc/perf/xmwlm.YYMMDD (year month day) by default
  - Samples data every 15 seconds and averages the values every 5 minutes
  - Save 2-7 days of data depending on exact level of fileset
  - Data is formatted using /usr/bin/topasout and supports several formats including nmon analyzer, CSV (Comma Separated Values) and text.
  - May not be detailed enough to resolve performance issues

Virtual I/O Server Performance Monitoring Tools

- Pre-installed Tools
  - topas –C and topas –cecdisp
  - iostat
  - lparstat
  - mpstat
- External Monitoring – via daemons
  - Performance Toolbox (PTX)
    - Daemon runs on each AIX LPAR
    - Can be used to build a “monitor” to capture dynamically certain statistics
    - The capture and saving of the data to the files can be automated
    - The “monitor” can be replayed
    - Available only on AIX
    - PTX GUI is X windows based
  - IBM Tivoli Monitoring System Edition for System p (ITMSESP)
    - ITM SE for System p V6.1 enables the monitoring of multiple System p servers
    - Provides graphical views of the virtualization environment to ensure complete monitoring and quick time to value
    - Best practice solutions include predefined threshold
    - Provides explanation of alert and recommends potential actions to take to resolve the issue
    - Users can visualize the monitoring data in the Tivoli Enterprise Portal

Virtual I/O Server Performance Monitoring Tools

- External Monitoring – via daemons
  - Data from xmtopas (installed by default in VIOS) via SNMP
  - LPAR2RRD – CPU Cross Partition Graphs from HMC data with RRDTool
    - Produces historical CPU utilization of LPARs and shared CPU usage
    - All LPARs (AIX, VIOS,i5OS, Linux on POWER) and all CPU stats are included
    - Data is extracted from HMC via ssh and loaded into the RRDTool database
    - Gives CPU Cross Partition Graphs based on 60 secs CPU utilization averages provided by HMC
    - Collects complete physical and logical configuration of all managed systems and their lpars and all changes in their state and configuration
- Nmon
  - Available in VIOS 2.1
  - Built into topas command
  - As padmin user – topas and then hit “~” to go into nmon mode
  - As root user, nmon command can be used directly which starts topas in binary mode but in nmon mode

Virtual I/O Server Performance Monitoring Tools

- Stealth Tools – workable but not supported by AIX Support
  - Ganglia
    - Open source for monitoring clusters
    - Data is typically stored in rrdtool database
    - Hundreds of machines and LPARs and history data over hours, days, weeks and months can be viewed
  - Lparmon
    - Simple graphical tool that shows what is going on in the machine

Monitoring Packages

- topas – supplied with AIX – supports recording and uses nmon analyzer to generate reports – see /usr/lpp/perfagent/README.perfagent.tools
- AIX Resource Monitoring and Control (RMC) http://publib.boulder.ibm.com/infocenter/pseries/v5r3/index.jsp?topic=/com.ibm.help.rsct.doc/rsct_books/rsct_admin_guide/bl5adm1112.html
- Nmon – Recording performance data – community supported - http://www.ibm.com/developerworks/aix/library/au-analyze_aix/index.html
- Ganglia - http://ganglia.sourceforge.net/ and http://www-941.ibm.com/collaboration/wiki/display/WikiPtype/ganglia
- Also see the *AIX Performance Management Guide* for more details about interpreting the output of monitoring commands

Other Monitoring Tools

- lparstat
  - Reports logical partition (LPAR) related information and statistics since boot time
  - The WPAR configured ID and the WPAR key by running the lparstat command with the -W flag
  - This command also displays processor information that might be helpful for licensing
- lvmstat
  - The **lvmstat** command generates reports that can be used to change logical volume configuration to better balance the input/output load between physical disks.
  - By default, the statistics collection is not enabled in the system. You must use the **-e** flag to enable this feature for the logical volume or volume group in question.
  - Useful to detect whether certain areas or partitions of a logical volume are accessed more frequently than others
- fcstat
  - The fcstat command displays the statistics gathered by the specified Fibre Channel device driver. It collects the statistics using the following procedure:
    - Opens the message catalog of fcstat and checks the parameter list.
    - Accesses the ODM database for information relating to the selected adapter.
    - Accesses the ODM database for information relating to ports of the selected adapter.
    - Opens and accesses adapter statistics.
    - Reports statistics and exits

Top Monitoring

- Simplifies performance analysis of large server configurations
- Focuses on elements consuming most system resources
- Sorted into lists referred to as *top-lists.*
- Helps identifying and diagnosing issues by focusing on constrained resources.
- Top framework records a defined number of performance metrics, by resource, at all times for all systems and consists of
  - User centered distributed top resource client
  - Always-on agent data collection and recording
  - Tabular report summaries
  - Near real-time response for active monitoring
  - Playback function
  - Common recording format allows support by existing trend analysis client (**jazizo**)

Top Monitoring

- *Top* agent is added to /etc/inittab on installation so that it is enabled by default
- *Top* agent can be disabled by commenting out the *xmtrend* entry in the file
- *Top* agent uses /**usr**/**lpp**/**perfagent**/**jtopas.cf** configuration file
- *Top* data is recorded into the /**etc**/**perf**/**Top**/ directory by default
- *Top* data recordings can be viewed by jtopas client or jazizo trend analysis tool
- *jtopas* is a Java-based system monitoring tool that has
  - Console to view summary of overall system
  - Separate consoles to focus on particular subsystem

Summary

- System monitoring is critical for
  - Reliability
  - Availability
  - Security
- Monitoring decreases response time by identifying the issue more quickly
- Knowing what is in the logs can significantly improve resolution time
- Unplanned failures are less likely to cause unscheduled downtime

Trademarks

The following are trademarks of the International Business Machines Corporation in the United States, other countries, or both.

Not all common law marks used by IBM are listed on this page. Failure of a mark to appear does not mean that IBM does not use the mark nor does it mean that the product is not actively marketed or is not significant within its relevant market.

Those trademarks followed by ® are registered trademarks of IBM in the United States; all others are trademarks or common law marks of IBM in the United States.

For a complete list of IBM Trademarks, see www.ibm.com/legal/copytrade.shtml:

*, AS/400®, e business(logo)®, DBE, ESCO, eServer, FICON, IBM®, IBM (logo)®, iSeries®, MVS, OS/390®, pSeries®, RS/6000®, S/30, VM/ESA®, VSE/ESA, WebSphere®, xSeries®, z/OS®, zSeries®, z/VM®, System i, System i5, System p, System p5, System x, System z, System z9®, BladeCenter®

The following are trademarks or registered trademarks of other companies.

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.
Cell Broadband Engine is a trademark of Sony Computer Entertainment, Inc. in the United States, other countries, or both and is used under license therefrom.
Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.
Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.
Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.
UNIX is a registered trademark of The Open Group in the United States and other countries.
Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.
ITIL is a registered trademark, and a registered community trademark of the Office of Government Commerce, and is registered in the U.S. Patent and Trademark Office.
IT Infrastructure Library is a registered trademark of the Central Computer and Telecommunications Agency, which is now part of the Office of Government Commerce.

* All other products may be trademarks or registered trademarks of their respective companies.

Notes:
Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.
IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.
All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.
This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.
All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.
Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.
Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.