Performance Key Metrics: March 29, 2011 Version – Bill Bitner bitnerb@us.ibm.com

While there are 1000s of VM Performance Metrics, the following are some of the most important ones to consider. I've included possible places to find these metrics in various sources. Note that many fields can be found in multiple places. Additionally, there are rules of thumb, which should be viewed only as guidelines, not law:

- **LPAR Overhead:** (Monitor: D0/R17 field SYTCUM_LCUMGTM, D0/R16 fields SYTCUP_LCUCACTM & SYTCUP_LCUCLPTM; Toolkit: FCX126 LPAR, some on FCX202 LPARLOG). There are two aspects of this, the LPAR management time associated with a particular partition, and the LPAR management time that cannot be charged to any one logical partition. This should generally be less than 5% of the Physical IFLs (CEC in an all-IFL configuration) for general LPAR management overhead, and then less than 5% of the z/VM partition IFLs .
- **LPAR Suspend Time:** (Monitor: D0/R2 formula - 100%-(SYTPRP_PFXUTIME+SYTPRP_PFXTMSYS+SYTPRP_PFXTOTWT; Toolkit: FCX126 LPAR or FCX202 LPARLOG). Suspend time is an approximation of how long the logical processor or partition was in queue to be dispatched at the LPAR level. RoT: 5% Suspend time is yellow line, 10% is red line for concern.
- **Total Processor Utilization:** (Monitor: D0/R2; Toolkit: FCX100 CPU; OMEGAMON: System workspace under headings of Percent CPU). This is the processor utilization from the VM perspective and includes CP, VM System, and Virtual CPU time. It is often beneficial to break this down into the three components:
  - **System Time:** This is the processor time used by the VM control program for system functions that are not directly related to anyone virtual machine. This should be less than 10% of the total processor time for the z/VM LPAR.
  - **CP Processor Time:** This is the processor time used by the VM control program in support of individual virtual machines.
  - **Virtual Processor Time: (Emulation Time):** This is processor time consumed by the virtual machine and the applications within it.
- **Total to Virtual Ratio:** (Monitor: D0/R2; Toolkit: FCX239 PROCSUM and FCX112 USER for virtual machines; OMEGAMON: System Workspace). The ratio of total processor time to virtual processor time is often used as an indicator of z/VM efficiency or overhead. The closer to 1.0, the better the z/VM efficiency. RoT: Should explore causes of a ratio over 1.30.
- It can be very valuable to convert the above to seconds of processor time and normalize it to the transactions where possible. Very valuable for capacity planning and problem determination.

- **Total Real Memory:** (Monitor: D1/R7; Toolkit: FCX103 STORAGE, FCX253 STORLOG, FCX143 PAGELOG; OMEGAMON: Storage Workspace) Track the total real central storage and expanded storage for a system.
- **Pageable Real Memory:** (Monitor: D0/R3 SYTRSG_RSAPGABL, SYTRSG_RSALGFRM; Toolkit: FCX253 STORLOG; OMEGAMON: Storage Workspace). The memory left over when CP nucleus, and non-pageable control blocks are taken out of the total real memory.

- **Paging I/O to DASD:** (Monitor: D0/R1; Toolkit: FCX143 PAGELOG) SSCHs used for paging to DASD.
- **Paging I/O from DASD:** (Monitor: D0/R1; Toolkit: FCX143 PAGELOG) SSCHs used for paging from DASD
- **PGINs:** (Monitor: D0/R5; Toolkit: FCX143 PAGELOG) Page rate of operations from expanded storage to central storage
- **PGOUTs:** (Monitor: D0/R5; Toolkit: FCX143 PAGELOG) Page rate of operations from central storage to expanded storage
- **Migration Rate:** (Monitor: D3/R9; Toolkit: FCX143 PAGELOG). Rate of page movement from expanded storage out to DASD.
- **Dispatch List Size:** (Monitor: D0/R10 SYTSCG_SRMCDSIP; Toolkit: FCX145 SCHEDLOG) Average number of virtual processors in the dispatch list (a significant difference here tends to indicate variations in workload).
- **Eligible List Size:** (Monitor: D0/R10 sum of SYTSCG_SRMCxELG; Toolkit: FCX145 SCHEDLOG) Average number of virtual processors in the eligible list (would expect this to be zero).
- **Total Virtual I/O Rate:** (Monitor: D4/R3; Toolkit: FCX100 CPU, FCX112 USER; OMEGAMON: Workload Space) Rate of traditional I/Os per second issued by the guests on the system.
- **Total Real I/O Rate:** (Monitor: D0/R1 SYTSYP_PLSCTSS; Toolkit: FCX100 CPU) Rate of traditional real I/Os per second issued to real devices.
- **Real DASD I/O Rate:** (Monitor: D6/R3; Toolkit: FCX183 DASDLOG; OMEGAMON: DASD Workspace). Rate of traditional real I/Os per second to real DASD devices.
- **Avg Pending Time for DASD:** (Monitor: D6/R3; Toolkit: FCX183 DASDLOG & FCX108 DEVICE; OMEGAMON: DASD Workspace) Average pending time for real DASD I/Os. RoT: Should be less than 1 millisecond.
- **Avg Connect Time for DASD:** (Monitor: D6/R3; Toolkit: FCX183 DASDLOG & FCX108 DEVICE; OMEGAMON: DASD Workspace) Average connect time for real DASD I/Os.
- **Avg Disconnect Time for DASD:** (Monitor: D6/R3; Toolkit: FCX183 DASDLOG & FCX108 DEVICE; OMEGAMON: DASD Workspace) Average disconnect time for real DASD I/Os.
- **Avg Service Time for DASD:** (Monitor: D6/R3; Toolkit: FCX183 DASDLOG & FCX108 DEVICE; OMEGAMON: DASD Workspace) Average service time for real DASD devices (sum of the pending, connect, and disconnect times).

- **Avg Response Time for DASD:** (Monitor: D6/R3; Toolkit: FCX183 DASDLOG & FCX108 DEVICE; OMEGAMON: DASD Workspace) Average response time for real DASD devices (Service Time + Queuing time).
- **SIE Instruction Rate:** (Monitor: D0/R2 SYTPRP_PFXRUNCP; Toolkit: FCX239 PROCSUM). The rate of the SIE instruction used by z/VM. This is measure of times we cross the boundary between the VM control program and the hardware running the guests.
- **Processor Spin Time:** (Monitor: D0/R2 SYTPRP_PFXSPINT; Toolkit: FCX239 PROCSUM). The percentage of time processors spend spinning on formal spin locks. RoT: Should be less than 10%.
- **Privileged Instruction Rate:** (Monitor: D0/R2 SYTSYP_PLSPRIVS; Toolkit: FCX100 CPU). The rate at which z/VM needs to simulated privileged instructions (a subset of the SIE Rate).
- **Diagnose Rate:** (Monitor: D0/R2 SYTSYP_PLSDIAGT; Toolkit: FCX100 CPU) The Rate at which z/VM processes diagnose instructions issued by guests (another subset of the SIE Rate).
- **Pct Allocation of Page DASD Space:** (Toolkit: FCX146 AUXLOG, FCX109 DEV CPOWN; OMEGAMON: Storage Workspace). The guideline of paging DASD is to keep it not much higher than 50%.
- **Block Read Size:** (Toolkit: FCX109 DEV CPOWN, FCX103 STORAGE) This goes in conjunction with Paging DASD space. I like to see this around 14 pages per Paging read I/O, but it is workload dependent. The key is when you see this drop significantly (and you're doing a fair amount of paging), then check allocation of VM page space. Chances are you have fragmented page space and VM can no longer form sizeable block sets.
- **Available List Empty:** (Monitor: D0/R4 SYTRSP_PLSALEMP, SYTRSP_PLSALEMG, SYTRSP_PLSALECL, SYTRSP_PLSALECG; Toolkit: FCX254 AVAILLOG). This is the count of times the available list was completely empty. Toolkit combines the count for both singles and contiquous frame available lists. If a non-zero value appears here, one should investigate how to relieve memory constraints.
- **Master Processor Utilization:** (Toolkit: FCX100 CPU, FCX144 PROCLOG). The utilization of z/VM's master processor. Important to watch for n-way limitations.
- **Various values for key virtual machines can be important:** (Toolkit: FCX112 USER, FCX257 UQDIO; OMEGAMON: Workload Workspace)**:**
  - **Virtual Processor Time:** Amount of processor time spent for virtual machine running Linux and its applications.
  - **Total Processor Time:** Total processor time includes both Virtual and CP processor time.
  - **TV Ratio:** The Total to Virtual Processor time ratio. A fair measure of the overhead of z/VM. The closer the 1.0 the smaller the percentage of processor time in z/VM control program.
  - **Virtual I/O Rate:** Rate of traditional I/Os performed by a guest.
  - **QDIO Rate:** Rate of QDIO operations (I/O to network and FCP devices) performed by a guest.

- o **Memory Distribution:** (Monitor: D4/R3; Toolkit: FCX113 UPAGE; OMEGAMON: Workload Workspace)
    - **Resident:** Amount of memory of the virtual machine residing in real memory of z/VM
    - **Expanded Storage:** Amount of memory of the virtual machine that has been paged out to expanded storage.
    - **DASD Slots:** Amount of memory of the virtual machine that has been paged out to DASD (note: there are a few rare cases where pages may be duplicated).
    - **Guest Page Reads:** Page rate from DASD to central storage
    - **Guest Page Writes:** Page rate from central storage to DASD
    - **Guest PGIN Rate:** Page rate from expanded storage to central storage.
    - **Guest PGOUT rate:** Page rate from central storage to expanded storage.
    - **Guest Migrate rate:** Page rate of migration from expanded to DASD.
    - **Normalized Values:** Again many of the data points above can be very helpful when normalized to transaction rates.
- **High-Frequency State Sampling:** (Monitor: D4/R4; Toolkit: FCX114 USTAT). These metrics show up in the Perfkit USTAT report. Remember the percentages are of "active", so if the guest is relatively inactive, who cares. The key states are:
    - o **Running (Run):** percentage of active samples where the guest was in SIE executing instructions for kernel & application work.
    - o **Simulation Wait (Sim):** percentage of active samples where the guest is in the process (i.e. CP is executing instructions on its behalf) or waiting for CP to simulate a privileged instruction of some type.
    - o **CPU Wait (CPU):** percentage of active samples where the guest virtual processor(s) are set to run, but are waiting to be dispatched. One exception to note is that this state can also result from operations such as diagnose x'44' where in a virtual n-way configuration, one virtual processor is trying to yield to other virtual processors in the same virtual machine.
    - o **SIO (SIO):** percentage of active samples where the virtual machine was waiting for synchronous I/O of some sort to complete.
    - o **Page Wait (Pag):** percentage of active samples where the virtual machine is waiting for synchronous page faults to be handled.
    - o **X-SVM:** ignore these SVM related counters for now, really only meaningful in CMS environments.
    - o **Console Function Wait (CF):** percentage of active samples that the virtual machine was in console function mode. CF aligns with the architecture model that certain functions serialize the machine (virtual machine in this case). The most common cause is CP commands that modify the virtual machine configuration. Many CP commands start in CF, but break free later.
    - o **Test Idle (Tst Idl):** Test Idle Wait is really an extension of idle time. z/VM uses the concept of test idle. Think of it as a sort of grace period

when a virtual machine goes idle (300 milliseconds). If the guest remains idle for this period, then it is dropped from the dispatch list to the dormant list.

- o **Asynchronous I/O (Asynch I/O):** I/O that does not stop the guest, but that is outstanding I/O.
- o **Asynchronous Page (Asynch Pag):** Linux (and VSE) can handle asynchronous page faults. Counts here indicate we took a page fault for one Linux task and are waiting for completion and there are not other tasks currently available for Linux to dispatch.
- o **Loading User (Ldg):** The guest is not currently delayed, but has been categorized as a "loading" user. That is it seems to be bringing in significant chunks of its working set.
- o **Other (Oth):** The other stuff, if you get consistent values of substance here, call me.
- o **Limit List (Lim Lst):** Percent of active samples where the guest was delayed in the limit list due to SHARE LIMITHARD.