Providing a Production System under VM

T.Y. Johnston

Stanford Linear Accelerator Center
Stanford University, Stanford, California 94305

## Introduction

At SHARE 58 in Los Angeles I talked about the batch monitor which we run.  Today I would like  to go beyond the batch  monitor and talk about  the environment we have created for the running of our  production work.   The batch monitor is one of the corner stones of this system,  but  primarily I would like to address the other functions which make this a quite complete production system.

As one thinks  about a service one  needs to think about  the separate groups involved with the service.  These are:

1.   The users of the  service.   How does the service help  them in doing their work?

2.   The providers of the service.  What tools have been given to the operators?

3.   The managers of the service.  How is the service managed and what tools are provided for the managers  to measure the quality and quantity  of the service?

Each of these three groups has a  different viewpoint of what the service is. Each also needs a different set of tools for interacting with the service.  This paper will explore  the tools and environment  from the viewpoints of  the user, the operations staff,  and facility management.   Attention will be paid to each of these groups ability to adapt their environment to fit their needs.   VM provides a very  adaptable environment for accomplishing this with  the concepts of nested EXECs and service virtual machines,  and  these concepts have been key to our success.

◇◇◇◇◇◇◇◇◇◇◇◇◇◇◇◇◇◇◇◇◇◇◇

## The User Viewpoint

The user clearly needs the greatest flexibility in establishing his/her environment. The functional needs of various users vary widely, the sophistication of the users also vary widely, and the cultures of various groups of users are also disparate. We have spent a major portion of our effort in trying to provide a framework for addressing these problems. One of the goals was to try and make the batch environment as close as possible to the interactive one. If a person has spent time developing a program interactively, then it should not be a difficult transition to take that program and run it in batch. This not only applies to programs, but more importantly to the environment within which the program runs. A significant weakness of VM is the inability to save a programs current environment in a machine processable form. We have taken some steps to assist the user in this area:

1.   QDISKID - returns the USERID and VADR for accessed disks.

2.   QFILEID - returns a complete definition of outstanding filedefs.

With these two commands a user is able to query some of the status of his virtual machine, save it, modify it, and restore it later in the same or a different virtual machine.

Although it is important to be able to move as much as possible of a users interactive environment from interactive to batch, batch is different. It is particularly different when one starts talking about production running where one does many things that are not done interactively. Some examples are:

1.   Go through other large aggregates of data (Such as 10 to 20 tapes of data).

2.   Consume major amounts of CPU time (minutes to hours).

To address these and related problems we have provided the following functions.

1.   SETUP and MOUNT commands to provide the user the ability to process tapes in a reasonable manner.

2.   BATCH SUBMIT, QUERY, CHANGE, CANCEL, and PURGE commands to allow the user to control his/her work. There are over 20 different options available on the SUBMIT command including an option which allows nesting options up to a total depth of 10. Most options that are specifiable on the SUBMIT command can be modified by the CHANGE command. The QUERY command uses a shared segment to make the responses to queries fast and inexpensive.

3.   JOBSTOP - allows a long running job to be shutdown gracefully either by the operator or the submittor.

4.   IOCOUNT - provides any user a count of I/O by virtual device for any VM.

5.   CMSINFO - returns to the caller information about the environment in which the command was issued.

6.   Multiple Unit Record Support - provides users multiple virtual readers, printers, and punches.   Primarily used in batch where frequently a program will have a listing on one printer and plots on another.

7.   CMS Abend Recovery - provides the ability for necessary cleanup to occur in case of such failures as exceeding batch limits, program failures (such as program checks), and other phenomena such as reading blank tapes.

8.   Hold after print - provides an extra option on the SPOOL and CHANGE command to allow a print file to not be purged after printing.   Reduce need for rerun caused by lost or poor quality output.

9.   JOIN - provides the ability to make a single spool file out of multiple spool files.

10.  Fortran command interface - provides the ability from a FORTRAN program to issue any CP and many CMS commands.

11.  Disk Space Manager - provides allocation and management of short term (usually 24 hour) and long term minidisks.   Allows group "czars" to manage the space for their group.   Communicates with DIRMAINT program product to provide user space allocation.


## The Operator Viewpoint

    The running of production in addition to interactive works adds significant demands to the operator.   In addition to having to be concerned about the interactive users and their modems, switches, controllers, et al, the operator also has to be concerned about keeping the production work flowing through the system.   Most of the routine functions such as scheduled production schedule changes are handled by a service VM.   However, the operator needs additional tools for monitoring and controlling the many facets of the production system. Some of these functions are:

1.   Operator EXECs and commands provide the ability for the operator to manage the facility.   Some of the functions included are:

     a.   Ability to place tape volumes in hold.   Batch jobs requesting such volumes will be placed in hold and automatically released when the volume is again available.

     b.   Dynamically updated system status display.   Displays various indicators of overall system usage including work queues.   Concentrates on displaying unusual conditions.

     c.   Batch EXECs to start, stop, drain, and redefine individual batch workers and the complete batch system.

2.   Multiple Console Support - provides the ability to route messages to

operator work stations associated with specific devices or operations functions.

3.    Tape monitor - provides on a 3278 mod 4 using XMENU a dynamically updated display of status of all 21 tape drives and also displays mount messages.

4.    Service Machine Monitor - Alerts operations when a malfunction of one of the service virtual machines occurs. Keeps track of the status of tape drives so that devices that should not be in service are not inadvertently put back in service across a system restart. Monitors the load on the system and reduces the number of active batch workers during peak load times, so that the interactive users will not be adversely affected.


## Facility Management Viewpoint

Facility management is concerned with several aspects of the production system. Are we providing a quality service that is responsive to the user's needs? Are there bottlenecks in the system can than be aleviated? Are the various groups who provide maintenance performing up to expectations?

To manage these various areas we depend on two primary systems.

1.    A problem tracking and management system. Associated with this system is a weekly status meeting with the vendors to discuss problems, plan solutions, and follow up on actions.

2.    The Accounting system. This system provides a variety of functions for the installation to monitor the use of the facility and for users to monitor their own activity. The system integrates the SLAC produced accounting records with the standard IBM records and provides accounting, reporting, and charging. Most of the reporting is oriented towards online examination of the data. Most of the accounting system is written using SAS. Some of the functions are:

   a.    Logging of calls to specific functions to understand usage patterns.

   b.    Batch accounting records for turnaround and utilization reporting.

   c.    Tape utilization records. These are used for charging, error analysis (bad tape drives and or volumes), and usage reporting and management. The data are available online (for some time) for both user and facility review. Some of the items available are:

      1.    Exception reports both daily and monthly indicating jobs that waited a long time before tape drive(s) could be allocated to the jobs to track whether we have sufficient tape drives.

      2.    Exception reports showing setup requests that took the operators

longer than expected.  This coupled with a graph showing tape mounts by hour by time of day  can tell us whether we have people goofing off or inadequate staffing.

3.  A daily  report by  tape drive  showing each  tape that  exceeded reliability boundaries (1 or more permanent errors or more than 5 temporary errors).   This report also  shows how many  tapes and I/Os occurred on  the drive before the  incident occurred.   This helps in determining whether it is a media or drive problem.

d.  Printer Utilization records.    These are used for  charging and usage reporting and management.

## Summary

Although the SLAC production environment provides  a rich environment,  it still is missing some important production facilities.  Areas where further extensions are planned are:

1.  Job networking (Similiar to JES3 DJC networks)  – although simple user networks are now  possible the user needs  the ability to define  more complex networks.

2.  Resource Control – Multiple users may want access to a single real resource in batch, and it is necessary to control the serialization of the resource.

3.  From a facility management viewpoint the biggest single area to clean up is the operator interface.   The programmed operator support of VMSP release 2 should be a step in the right direction.   Since we are a 3081 we must wait for HPO3 before we can start using the functions.  We believe that with the functions that we have  in our system and those available  with SP2 that we should be  able to  go a long  way towards spotting  faulty tape  drives as problems develop  rather than  relying on  either angry  users or  somebody reading a LOGREC  report hours later to tell  us that a tape  drive is malfunctioning.

This has been a brief  look at a  production system.   Such a system is viewed by many people  in a variety of ways.   The operator wants a  work station that does not get in his/her way in performing his/her tasks;   the user wants a system with function,  but where simple tasks are not difficult;  and facility management want  to be  able to  understand and  manage this  facility.   Our work coupled with VM concepts such as nested  execs and service virtual machines have made this a reality.  Each of the major classes of users has that user's view of the system tailored by EXECs.   Some are  facility supplied and some are written by the user or a  peer,  but it all gets down to the  ability to have a tailored environment which supports the users work rather than getting in the way.

## Appendix A

## A MYTH EXPLODED

The myth has persisted for some time that VM is "expensive". Certainly, the running of guest operating systems can have significant costs, but I believe that the following statistics demonstrate that VM can provide a very effective and efficient production environment.

| Description | Amount | Percent |
|---|---|---|
| Total CPU hours January 1983 (TTIME) 2 CPUs and percent of hours in month | 1,389 | 93 |
| Total CPU hours by users (TTime excluding overhead account service VMs) and percent of hours in month | 1,325 | 89 |
| Net User cycles (VTIME excluding VTIME for overhead account service VMs) and percent of hours in month | 1,154 | 77 |
| Net Batch cycles (Vtime for batch work) and percent of hours in month | 1,053 | 71 |
| Total Batch CPU Time (TTIME) and percent of total user time (TTIME) | 1,122 | 85 |
| User TTIME percent of total cycles | | 95 |
| User VTIME percent of total cycles | | 83 |
| Batch percent of user VTIME | | 91 |
| Batch percent of total CPU (Batch VTIME divided by total TTIME) | | 77 |
| Batch jobs per month | 21,846 | |
| Total tape mounts for the month | 15,009 | |
| Tape mounts per average weekday | 550 | |
| Interactive Sessions | 29,717 | |
| Man months of Interactive connect time | 209 | |
| Weekday average maximum number connected users | 167 | |

We have found VM to be an efficient vehicle for the delivery of computing to our users. The major reason for lost cycles during the month is the paging and other I/O contention in the late morning and mid afternoon of weekdays.