Saltzer

To: Distribution

From: T. H. Van Vleck

Date: June 17, 1976

Subject: New Storage System Long Range Plans (revised)

This document reports changes since MTR-119.

OVERVIEW

The following table shows the major phases of the implementation of the new storage system.

Phase		Date
I	Command Level One user at command level	May 75
II	Prototype Running Several users	June 75
III	Design Review Error recovery, backup, mount/demount	Oct 75
IV	Installable System Run mini-service at CISL	Nov 75
V	Initial Installation at MIT No mount/demount	Feb 76
VI	Follow-up Installation at MIT Operational enhancements	March 76
VII	MR 4.0 Installed at MIT With mount/demount	May 76
VIII	Release MR 4.0	June 76
IX	Further Enhancements Administrative improvements	-

CHANGES SINCE LAST REPORT

Since MTR-119 was published, the following events have occurred:

1. Systems 28-1 through 28-8 have been installed at MIT and Phoenix, completing the capabilities required for release 4.0.

2. Operational experience led to the design and installation of several important reliability and performance enhancements.

3. Operational and user documentation was updated. This includes:

   MPM Commands
   MPM Subroutines
   MPM Reference Guide
   MPM SWG
   Multics Operator's Handbook
   Info segments
   SRB for 4.0

WORK COMPLETED

1. Statement of Problem.

MTB-017, November 1973. The Storage Problem

2. Design.

MTB-055, April 1974. Overview
MTB-060, May 1974. Paging Analysis
MTB-065, April 1974. New Bound on Performance
MTB-095, June 1974. Removable Disk Packs
MTB-110, August 1974. Implementation of New Storage System
MTB-167, February 1975. Disk Usage
MTB-203, June 1975. Attributes of Backup System
MTB-206, June 1975. SAVE and RESTOR
MTB-213, July 1975. Disk Definition
MTB-220, September 1975. Salvager
MTB-221, September 1975. Salvager Implementation
MTB-229, October 1975. Demountable Logical Volumes
MTB-233, November 1975. Backup
MTB-237, November 1975. Overview
MTB-238, November 1975. Initial MIT Installation
MTB-239, November 1975. Error Recovery
MTB-243, December 1975. User Ring Changes
MTB-246, December 1975. List Command
MTB-260, February 1976. Interim Mount
MTB-261, March 1976. Master Directory Control
MTB-266, March 1976. Removing DC from the Security Kernel

3. Task Schedules.

MTR-068, October 1974. Implementation Plans
MTR-081, March 1975. Long Range Plans
MTR-084, April 1975. Task List Phases I and II
MTR-095, September 1975. Long Range Plans
MTR-112, December 1975. Long Range Plans
MTR-119, March 1976. Long Range Plans
MTR-120, March 1976. Discussion of MTB-250
MTR-121, March 1976. Discussion of MTB-233
MTR-122, March 1976. Discussion of MTB-261

32a. Directory Control Checking.

This task adds to directory control new code for maintenance of the new directory fields. (See also task 32b.)

Installed in system 28-8.

37. Master Directory Operations.

This task adds ring-1 support for operations on master directories. User calls are create, delete and list. The

create_dir and delete_dir commands are modified for this case.

The ring-1 programs use the volume registration data base, access control segments, and the Master Directory Control Segments (MDCSs). Administrative commands to manage these data bases are provided. For the volume librarian, we need register, unregister, modify, and list.

Installed in system 28-4.

38. Ring 1 Volume Mount Module (interim).

When a logical volume is to be mounted the operator must type several commands to tell the system that the volumes are online. These commands check the volume label against the registration and when the volume is completely mounted call the hardcore to cause the logical volume to be accepted for paging.

Installed in system 28-4.

39. User Request to Connect Logical Volume (interim).

User requests to connect to a logical volume will be passed through ring 1. If the user process is permitted to connect to the logical volume, and the logical volume is already mounted by the operator, the request is passed to the hardcore.

Installed in system 28-4.

40. Hardcore Check on Volume Connection.

The hardcore is changed by this task to require the connection call from ring 1 before allowing a process to initiate a segment on a demountable volume. (The root logical volume is never demountable and other "public" volumes can be declared not demountable.) This insures that ring 1 is not bypassed, and makes sure that all programs using segments on removable volumes execute independently of whether some other process has caused a pack to be mounted. The list of demountable physical volumes which the process is connected to will be stored in the KST.

Installed in system 28-4.

41. Phase VI: Follow-up Installation at MIT.

Operational experience will lead us to make many improvements to the interface and behavior of the storage system. Performance measurements under actual load may also show use where to concentrate our programming effort in order to speed

the system up; if these improvements are possible we will install them soon.

Target date: March 1976.
Actual installation was spread over several systems.

42. Command System Changes.

These changes are the ones specified in paragraph 27. In addition to the changes to handle new error and state conditions, the create_dir command must accept and check the new parameter which specifes the logical volume in which storage will reside, and the status command must be modified to show this attribute.

Finished.

43. Phase VII: Install MR 4.0 at MIT.

Target date: April 1976.

This installation was broken down into smaller parts, and was installed as follows:

28-1	3/19	fixes
28-2	4/01	fixes, speedups
28-4	4/26	mounting, registration
28-5	5/12	demounting
28-7	6/10	reliability
28-8	6/16	reliability

CURRENT TASKS

44. Phase VIII: Release MR 4.0

Target Date: June 1976. Currently on schedule.

52. Continuing Performance Improvement

This task includes the investigation and metering of the system's performance as installed at MIT, and the identification and removal of performance bottlenecks. For example, the locking and I/O strategies used by the supervisor will be studied. Some of the results of this investigation were included in systems 28-2, 28-4, and 28-5.

24. Implementation of Hardcore Primitives for Backup.

The hardcore primitives to support the new backup system must be able to maintain the list of modified segments on each physical volume for the use of incremental dumping; and to activate and dump or reload a segment by volume ID and VTOC

index without referencing the branch.

25. Implementation of Backup Dumping Programs.

The new complete and incremental dumping programs can be much simpler than the current dump programs, since all hierarchy walking and access forcing code is eliminated. The hardcore primitives do most of the work. These programs are easy given the format of the output records to be produced.

26. Implementation of New Reload and Retrieve.

The reloading and retrieval programs will use the output of the dumping programs to reconstruct volumes and to recover the contents of single segments.

32b. Directory Control Checking.

This task adds to directory control new code for appropriate in-line checks of the new directory fields. MTB-220 describes the details of this change.

53. Improve Disk DIM

This task covers improvements in several areas: handling of "device in standby" statuses, reassigning page addresses on error, and recording of bad track addresses.

54. Connection Check while System is Running

The current method of performing a connection check from the VTOC back to the hierarchy requires complicated manipulations of disk packs and is error prone. This facility will be replaced by a privileged program which runs under Multics and which can collect free VTOC entries while the system is using the volume.

55. Volume Maintenance Programs

This step provides utilities for moving segments from one physical volume to another upon (privileged) request. This facility can be used to compress a logical volume. A facility to grow or shrink partitions and VTOCs will also be provided.

56. Salvager Improvements

This step rewrites the directory salvager to operate in a cleaner and more reliable fashion. We hope that we will be able to permit users to invoke the on-line salvager to cause a directory to be compacted via hcs_.

57. Enforce Directory Quota

Directory quota is not currently enforced, and a determined user could cause system problems by creating many directory records. When users can control their usage (i.e. when they can compact directories on demand) and when the system is known to function correctly even when a directory cannot be expanded, the call side of directory control will be modified to enforce the directory quota.

FURTHER ENHANCEMENTS

36. Backup Integration.

This task integrates the new backup mechanisms into the system and ties backup in with salvaging.

Although most of the parts of the new backup system will be available by the time release 4.0 is frozen, the new backup has not had sufficient testing and operational experience to allow us to depend upon it.

51. Change Volume Mounting to use RCP.

Modify the mount and demount commands to use RCP for mounting and demounting volumes.

When a logical volume is to be mounted, the LVRF must be consulted to find the list of physical volumes to be mounted. Calls must then be made to RCP to mount each of the physical volumes, the volume labels must be checked, and the hardcore must be called to tell it that the volumes are accepted.

46. Keep Duplicate Copies of Selected Volumes.

Once this task is completed, crucial volumes in the system can be maintained in duplicate; all modified pages will be written out to both devices. In a configuration which places the secondary copy on a different disk subsystem from the primary copy, the cost of maintaining two copies will be very low.

47. Automatic Use of Secondary Volume on Error.

Once the duplicate copy facility is available, the system can be modified so that when a disk record is unreadable, the system automatically switches to the use of the secondary copy.

49. Calls to Initializer Process During Connection.

This step causes RCP to pass all connection requests through the system control process, so that charging can be done, mount messages can be routed, and so that operator commands affecting the request can be issued.

50. Billing.

Modifications must be made to the administrative and billing package to enhance the administrator's ability to manage the system resources. Some of these improvements cannot be specified until we have obtained some operational experience.