==========
==========

M A News
Mainframe Audit News
December, 2003
Issue Number 05

==========
==========

Table of Contents

==========
==========

1) Introducing the Mainframe Audit News

This is the fifth edition of the Mainframe Audit News, a vehicle for sharing information about auditing IBM mainframe computers. For more information on this newsletter, including how to subscribe or un-subscribe, please see section 4).

==========
==========

2) How Dataset Names Work in MVS

A dataset (also called a “file”) is a collection of records treated together. If you have worked with UNIX or Windows computers, you probably have seen filenames containing slashes and dots, like this:

/prod/finance/accounts_payable/checks.oct

In UNIX and Windows, the slashes separate directories: the **prod** directory is at the top of a tree, the **finance** directory is subordinate to it; the **accounts_payable** subdirectory is below the **finance** directory, and the file itself is named **checks.oct**. This might be rendered in English as: the file **checks.oct**, which is in the directory **accounts_payable**, which is in the directory **finance**, which is in the directory **prod**.

You can use control permissions to any of the directories, permitting or preventing a user who wants to access files below that directory in the directory tree.

MVS dataset names do not have slashes. An example might look like this:

PROD.FINANCE.AP.CHECKS.OCT

Note that the letters are all upper case. Note also that the name is separated into pieces by the dots. The pieces are called nodes. They are also called qualifiers. **PROD** is a node. So is **FINANCE**. **PROD** and **FINANCE** are also called qualifiers. (IBM apparently has a policy that once you give something a name, you must give it a second name, in order to confuse people.) The maximum length for a node is 8 characters, and the maximum dataset name length is 44 characters.

The left-most node (**PROD** in our example) is called the HLQ or High Level Qualifer. It is an important factor in:

- Learning what kind of dataset it is
- Finding where the dataset resides, and
- Controlling access to it.

Learning What Kind of Dataset It Is

The High Level Qualifier rapidly identifies the nature of a dataset. If the HLQ is **SYS1**, then it is a system dataset. If the HLQ is a programmer's userid, then the dataset belongs to that programmer. If the HLQ is **PROD** (or in some installations **P**), then it is usually a production dataset. A HLQ of **TEST** (or **T**) usually identifies a test dataset.

When the HLQ is either **PROD** or **TEST**, then the second qualifier often identifies the application that dataset belongs to.

PROD.PAYROLL.OCTOBER.DATA is a production payroll dataset.
TEST.PAYROLL.OCTOBER.DATA is the corresponding test dataset.
SYS1.LINKLIB is a system dataset.
STU.PAYROLL.OCTOBER.DATA belongs to a programmer named STU.

Finding Out Where the Dataset Resides:

To discuss finding where a dataset resides, we need to discuss the concept of the **catalog**. The catalog is a dataset whose only purpose is to contain the locations of other datasets. Each disk pack (like a hard drive, but it holds more data) is assigned a

six digit identifier called the **volume serial number** or **volser**. When a dataset is created on disk, the volser of the disk is where the dataset resides.

If the dataset is catalogued (that is, entered into the catalog dataset), then the catalog dataset will have a record specifying the dataset name and volume serial number of the disk pack. Later, when a program tries to read the dataset, the operating system goes to the catalog to find out which disk pack the dataset lives on.

If the dataset resides on tape instead of on disk, it can still be located through the catalog, since each tape cartridge has a unique volume serial number too.

Originally there was only one catalog on each system Later, as the number of datasets and HLQs increased, this single catalog was replaced with a **master catalog** and a number of **user catalogs**. Each user catalog has records listing dataset names and the volsers of the disk or tape where they lived.

Each user catalog has catalog records only for certain HLQs. For example, all the **SYS1** datasets might be cataloged in one user catalog, while all the **PROD** datasets would be cataloged in another user catalog. System programmers might have their datasets cataloged in a third user catalog, while applications programmers' datasets might be cataloged in a fourth.

While each user catalog contains pointers to datasets, specifying where they are located, the master catalog has pointers to user catalogs, based on HLQs. One record in the master catalog might specify that the pointers for all SYS1 datasets might be found in the first user catalog, while the pointers for all PROD datasets might be found in the second user catalog. Each of these HLQ pointers in the master catalog is called an **alias**.

So when the operating system needs to locate a dataset, it goes to the master catalog, which would point to a user catalog (based on the HLQ). The user catalog has the pointer record describing the tape or disk where the dataset might be found.

When a new online user is added to the system, the security administrator defines the new userid to the security software. If the user is a programmer, then his datasets will usually have his userid as the HLQ. In this case, the security administrator will also need to define a new alias in the master catalog for the userid. If the userid is **STU** for example, the alias would specify that all datasets whose HLQ is **STU** will be catalogued in a certain user catalog.

Controlling Access to the Dataset:

Security software such as RACF and ACF2 usually simplifies dataset security by having one dataset rule cover many datasets, based on the HLQ. For example in RACF, a rule named **PROD.**** would match all dataset names whose HLQ is **PROD**. A rule named **PROD.FINANCE.**** would match all dataset names whose HLQ is **PROD**, and whose second qualifier is **FINANCE**. (The **.**** is translated something like "any characters after that".)

A comparable set of rules in ACF2 might look like this:

$KEY(PROD)
- UID(…..) R(A)
FINANCE.- UID(…) R(A) W(A)

(where in the first line **$KEY** identifies the HLQ, and in the second line the dash (-) matches any value for the second and following qualifiers and in the third line **FINANCE.-** matches any second and following qualifier starting with **FINANCE.**) The **R(A)** and **W(A)** specify read and write permissions for matching UID strings. (Please ignore the the term "UID strings" until we explain ACF2 in greater detail.) The dash in ACF2 is similar to the .** in RACF.

For all three security software products (RACF, ACF2, and TopSecret), any user is automatically granted complete access to any dataset whose HLQ is his userid.

When you initialize a tape cartridge or you create a tape dataset, the operating system writes records on the tape before the dataset. These records are called **labels**, and specify the volser of the tape cartridge and the **dsname** (dataset name) of the dataset. Later when another program reads the tape file, the operating system verifies that the volser and dsname in the label match those specified in the JCL (Job Control Language) for the program. This provides a two-way check: once comparing the dsname requested with that on the label of the tape, and the second passing the dsname requested to the security software.

What This Means to Auditors:

Mainframe auditors will need to identify the naming standards for datasets, in a given data center in order to understand what a given dataset name means, what kind of data it includes, and what the security rules for datasets specify. In particular, on mainframes the security software should be used to provide separation between production datasets and test datasets, preventing programmers and operators from

having any access at all to production data. Use of test datasets in production batch jobs should not be permitted.

One of the advantages of the mainframe computer is that it has the power to provide separation between production data and test data. This is important to provide a controlled environment for the execution of production batch jobs. Auditors will want to address how well the production environment is controlled.

Test of What You Have Learned

In Issue 03, we described **Job Control Language** or JCL, including the **DD card** which describes a dataset to be used by a program. Below is a partial sample of a DD card. How many of the terms in this article can you find in it? (Hint: **DISP** is short for disposition.)

```
//MYINPUT    DD    UNIT=TAPE,VOL=SER=123456,LABEL=(2,BLP),
//                 DISP=(NEW,CATLG),DSN=STU.DATA
```

3) Seminar Information, and the Proverb of the Day

3A) >>>>Seminar Information

This issue we highlight two new seminars for mainframe auditors, and describe two new seminars for new IT auditors:

- **HG64: How to Audit MVS, RACF, ACF2, CICS, and DB2** (November 1-3, 2004 in FL)

- **HG76: How to Audit UNIX Security (Including: LINUX, AIX, and USS)** (November 4, 2004 in FL)

The Henderson Group also offers these "How to Audit..." courses :

- How to Audit **Cross-Platform Applications** (April 1-2, 2004 in MD)
- How to Audit **Mainframe/Internet Connections** (March 1-2, 2004 in FL)
- How to Audit **CICS** (March 3, 2004 in FL)
- How to Audit **RACF** (April 15-16, 2004 in MD)
- How to Audit **MVS** (April 14, 2004 in MD)

To learn more about them, please go to

http://www.stuhenderson.com/XAUDTTXT.HTM

Two one-day seminars are now available for new IT auditors:

? **Introduction to IT Audit**
? **Exercises for IT Audit**

For more information please contact KBSoni at (301) 590-7121 or Kbsoni@aol.com.

3B) >>>>This Issue's Proverb of the Day

_____

(Paraphrased from Peter Drucker) "*If you can't measure it, then you can't manage it*."

Sometimes our audit scope addresses not financial controls, nor security controls, but management controls for effectiveness. For a mainframe data center, these management controls might include: response time management, problem management, chargeback for computer usage, hardware capacity planning, hardware maintenance, and others. If we note for example that response time for online transactions for a given application averages three seconds, how do we determine whether this is appropriate or not? One way would be to compare it to service level agreements between the data center and the online users. Another would be to compare the average response time to what it was in the past (perhaps a chart with time on the horizontal axis). Another way would be to see what variation there is in response time throughout the day. Yet another way would be to compare the average response time to the response time recorded when there is no other work on the computer or on the network to compete with the online transaction. Each of these could provide a method to evaluate the three second response time.

Of course, if the data center manager doesn't measure the response time (or have someone measure it for him), then these evaluation techniques are useless. She can't be managing the response time if she doesn't measure it. What other measures would you look for as indicators of management control in a data center? In applications programming? In a financial audit? In an information security audit? In an applications control review?

==========
==========

4) About the Mainframe Audit News; How to Subscribe/Unsubscribe

The MA News is a free, email, newsletter for auditors who need (or suspect that they will need) to be auditing IBM mainframe systems (primarily MVS, OS/390, z/OS, and the system software associated with them). This software includes: CICS, DB2, JES, VTAM, MQSeries, TSO, USS (UNIX System Services), TCP/IP, and others. It also includes the Websphere server software which connects a mainframe to the Internet. (Note, we will expand each of these acronyms and explain how the software works over the course of the next several issues.)

With this issue, we intend to start issuing shorter newsletters, but more frequently.

The MA News is meant for auditors who are new to IBM mainframes, as well as for experienced MVS auditors who want to keep up to date with the latest developments from IBM. We will not make the list of subscribers available to anyone else for any reason.

4A) >>>>To Subscribe, Unsubscribe, or Request Back Issues for the Mainframe Auditors' Newsletter (MA News)

Send an email to: **stu@stuhenderson.com** with the subject field set to: **MA News** and in the body of the email just the phrase you want: **SUBSCRIBE** or **UNSUBSCRIBE** or **BACK ISSUES: 1, 2**

4B) >>>>To Get Questions Answered from the Mainframe Auditors' Newsletter (MA News)

Send an email to: **stu@stuhenderson.com** with the subject field set to: **MA News** and in the body of the email the word: **Question**: (followed by your question, and tell us whether you want your name included or not if we decide to publish your question and answer) Whether we print your question and answer or not, we will try to email you back an answer to your question within five business days. If you need a faster answer, please phone your question to (301) 229-7187, leaving the question on the machine. Please repeat your phone number slowly and clearly.

==========
==========

4C) Tell Us What You Think

We'd love to hear from you, in particular on these topics:

- What do you like/not like about the MANEWS?
- What websites do you know that you want to share with other auditors?
- What topics and/or columns would you like to see in future issues?
- Is your mainframe connected to the Internet and do you plan to audit the security implications of this connection?

Please email your comments to stu@stuhenderson.com. Thanks.

==========