A Solution for Maintaining File Integrity within an Online Data Archive

Daniel M. Scholes, Thomas C. Stein, Edward A. Guinness

Washington University
Earth and Planetary Remote Sensing Laboratory
Campus Box 1169, 1 Brookings Drive, St. Louis, MO 63130 USA
EMail: scholes@wunder.wustl.edu, stein@wunder.wustl.edu, guinness@wunder.wustl.edu

ABSTRACT

Online data archives can be threatened by corruption from accidental changes and deletions, software errors, hardware failure, and malicious changes by viruses and hackers. An archive may solely rely on an offline backup copy as the solution for resolving archive corruption. While this strategy resolves basic data restoration issues, it does not detect file changes, locate missing files, or contain a mechanism for verifying the status of backup files. These are more complex activities to incorporate into the maintenance of an online archive.

The Geosciences Node, a discipline node of the Planetary Data System (PDS), is archiving approximately 40 TB of planetary data and adding 1-2 TB of data a month. The Node maintains primary and secondary online copies and a tape backup. The Node has developed an archive management system to actively monitor its multiple archive copies and to track valid file changes. An archive baseline of valid directory statistics and file checksums is cataloged in a database. The database is updated when data are added or replaced in the archives. Validation scans are executed on a monthly schedule to compare primary and secondary online archive copies to the cataloged baseline to ensure archive integrity. The tape backup copy is compared to the baseline checksum to verify any restored data. Through the use of this system, the Geosciences Node has improved its ability to accurately maintain online data archives.

Keywords: PDS, Data Archive Management, File Integrity, Checksum, Validation

INTRODUCTION

The Geosciences Node of NASA's Planetary Data System (PDS) is responsible for developing and maintaining long-term archives of geological data from NASA's orbital and landed missions to Mars, Venus, Mercury, and the Moon. The Node currently maintains archives from over 20 NASA missions, which consists of 40 TB of data stored in 13 million files. The Geosciences Node is the international science community's primary source for most of these data. Thus, the Node's archives are a valuable scientific resource, which makes maintaining their integrity an important part of operating the archives.

There are many threats to file integrity of long-term data archives (Table 1). This is especially true when data archives are actively growing and evolving with new and replacement files. Files can be affected by accidental changes and deletions during the update of a data set. Software errors can inadvertently corrupt or modify files. Hardware failures can also cause corruption and data loss problems. Malicious changes by viruses and hackers are ongoing threats. Natural disasters and other catastrophic events can also destroy an entire archive system.

Entity at risk	Size	What can go wrong
File	Up to 4 GB	Corrupted media, disk failure
Tape	Up to 1600 GB	+ Simultaneous failure of 2 copies
System	~10 TB	+ Systemic errors in vendor software, Malicious user, or Operator error that deletes multiple copies
Archive	~1 PB	Natural disaster, obsolescence of standards

Table 1: Digital entity at risk and threat (adapted from Berman [1])

Many of these threats can be avoided through proper firewall settings and strict security policies. Despite the best efforts of systems managers, problems can still occur. An archive may solely rely on an offline backup copy as the solution for resolving archive corruption. While this strategy resolves basic data restoration issues, it does not detect file changes, locate missing files, or provide a mechanism for verifying the status of backup files. These activities are more complex than backup and recovery operations, and comprehensive consideration is required to incorporate them into the maintenance of an online archive.

As a long-term data archiving and data distribution facility, it is essential that the PDS Geosciences Node present accurate data archives to its customers. A significant investment of time and monetary resources is made to collect, calibrate, and format the data archived in the PDS. Any loss of PDS data files is not an acceptable scenario.

While updating its disaster recovery strategy in 2006, the Geosciences Node began exploring ways of improving the process of ensuring and validating data integrity. Over the last year, the Geosciences Node has designed and implemented a comprehensive strategy to ensure the file integrity of the archives it maintains.

INITIAL DATA INTEGRITY STUDY

The Node's first studies included evaluating various hash checksum algorithms and their processing efficiency. The MD5 checksum algorithm was eventually selected for its speed, size, and wide-spread usage [2].

The initial trial solution was primarily a manual process. It served as both a temporary solution and an opportunity to test the requirements for the process that the Geosciences Node was developing. The process involved manually creating MD5 checksum text files of each data set file archive with the public domain application MD5Deep. The checksum files were then periodically compared to the existing data archives with MD5Deep to confirm data archive file integrity. While this method was technically successful, it had many drawbacks. As a manual process it was inherently time consuming, difficult to keep organized and required much attention from the individual managing the process. First, the process required the maintenance data set archive lists and a history of each data set archive validation. It was time consuming to create the first data archive checksum baselines. The next challenge was to accurately merge and update the checksum files when data archives had legitimate changes, such as file updates. In addition, the checksum comparison reports had to be manually reviewed to identify any corruptions. With many manual steps, comparisons, command line programs executions, and batch scripts to be maintained, the process required a skilled user with a focused effort to manage the validation checks.

SYSTEM REQUIREMENTS

The initial study provided valuable experience that led to the planning of a more comprehensive solution. The Geosciences Node compiled the following functional requirements for the long-term Archive Management System (AMS).

- **Maintain a catalog of archive contents.** The AMS will maintain a catalog of the data archive directory structure, file listing, directory file count, directory and file size, directory and file modify date, directory and file name case, and file checksum value. This detail is needed to identify even very subtle changes to the data archives.
- **Maintain a catalog of primary archives that can be mapped to secondary and additional archive copies.** The archive catalog will be created from the primary archive copy, but the same catalog configuration can be used to validate a secondary archive copy (warm spare) or a restored tape backup. The system will allow for an unlimited number of files and archive copies.
- **Track and add new and updated data to the archive baseline catalog.** Ongoing missions continue to deliver newly released data and recalibrated updates to the Geosciences Node data archives. The system needs to be able to easily add new data files to the archive baseline, as well as make sanctioned updates when files are replaced in the data archives.
- **Compare archives to the stored archive baseline catalog to verify the integrity of the data archive.** The application needs to validate the integrity of existing archived files against the stored configuration baseline (archive catalog). This includes verifying archive directory structure, directory file count, directory and file modification date, directory and file size, directory and file name case, and file checksum value. The application needs to identify files that are found on the disk, but are not in the archive catalog. Files that exist in the archive catalog, but are not found on the disk must be flagged, as well.
- **Validate data holdings every month.** The application needs to have the performance speed to verify the entire Geosciences Node's data holdings on a monthly basis.
- **Verify an external checksum file against an archive location.** Functionality is needed to validate a newly delivered data set. The Geosciences Node now requires that data providers include a checksum manifest with the data files that are delivered to the Geosciences Node. The checksum file can then be used to confirm that no files are corrupted during delivery to the Geosciences Node.
- **Provide improved application usability.** The Graphical User Interface (GUI) must provide an effective means to control the archive management database and applications. It must also provide intuitive status summaries and validation reports. The archive integrity management process should require less user "hands-on" time. The bulk of the processing should be automated with operator confirmation at specified checkpoints.

DATA ARCHIVE ARCHITECTURE

The Geosciences Node incorporated the considerations of security policy, file accessibility, multiple backup file copies, and file integrity validation when designing its overall data archive architecture. A brief description of these components explains how the Archive Management System is a piece of the overall design. The security policy is an initial guard against malicious attacks and accidental modifications. The file accessibility is handled by redundancy in the network and system architecture to avoid file access outages. Multiple file copies are maintained by system redundancy and backup methods to correct data loss situations. The AMS focuses on ensuring file integrity during normal operations and file restoration situations.

A standard firewall system and strict security policies block hacker and virus entries. Servers and user workstations use standard virus protection software to reduce the likelihood of malicious software causing destruction to data archives. User specific access and edit permissions limit archive modifications to Geosciences Node's senior data archivists.

The Geosciences Node's system configuration is created with multiple network switches, fail-over server configurations, RAID (Redundant Array of Inexpensive Disks) network storage devices, and a secondary data archive. The RAID systems provide high availability and data protection because their storage configurations can sustain the lost of multiple physical hard drives without the loss of files or the

logical storage device. The Node has invested in these redundant systems to provide greater archive reliability, and reduce the likelihood of file inaccessibility during hardware failures. To prevent the loss of files, multiple copies of the data archives are maintained by the Geosciences Node. The file copies include primary online archives, secondary online archives, tape backups, and an off-site deep archive.

As the size of the data archive holdings at the Geosciences Node has continued to grow, it has become more time consuming to create full tape backups from the live data archives. Additionally, up to five days could be required to restore all of the data archives from tape during a disaster recovery scenario, assuming a compatible hardware infrastructure is available. The Geosciences Node began using the secondary online data archive approach to provide a faster data restoration solution for cases of file corruption and disaster recovery.

In the mid 1990s, the Geosciences Node relied on a fiber network SAN (Storage Area Network) RAID system for its data archive storage platform. This system was expensive to purchase and maintain, but it provided a stable, fault-tolerant storage environment. A manual process was used to create a secondary copy on a less expensive NAS (Network Attached Storage) device. A network tape backup device was used to create an offline backup of the data archives from the primary SAN.

In 2009, the PDS Geosciences Node migrated its data archive holdings to an iSCSI network SAN RAID system. It is an easier and less expensive system to maintain and includes useful management software. The iSCSI SAN also provides greater flexibility by allowing more servers to directly connect to the storage device. A secondary replication storage site in a different building on the Washington University campus was created with an identical storage system as a file backup and warm failover site. The replication configuration is scheduled through system software to duplicate the primary archive changes to the secondary storage site on a weekly basis. A change history containing previous versions of replaced files is also maintained on the replication site until it is removed by the system administrator. This provides extra recovery protection if an errant change occurs on the primary site and then is replicated to the secondary site. If the error is discovered and the previous version can be verified, the file can be restored to the primary data archive site. The replication site can be configured to serve as the primary data archive site within hours depending on extent of infrastructure damage in a hardware failure or disaster situation. A tape backup system continues to provide an additional file backup and disaster recovery solution. Two data archive incremental tape backups are executed each week. Full tape backups of the entire data archives are created on a quarterly basis. Copies of the backup tapes are stored offsite.

An additional data preservation step is taken by submitting copies of data archives to the National Space Science Data Center (NSSDC) as part of the PDS charter. The NSSDC provides a deep archive backup function. The Geosciences Node submits copies of data archives to NSSDC via data bricks (external hard drive systems). NSSDC copies the files and returns the data bricks. In the event of a catastrophic disaster, the Geosciences Node's PDS data could be retrieved from NSSDC.

The Geosciences Node's Archive Management System (AMS) application provides the data archive integrity verification component to the overall data archive architecture. The AMS catalogs and verifies the data archives under normal operating circumstances. It is also used to confirm the integrity of data archive backups for use in recovery situations.

THE ARCHIVE MANAGEMENT SYSTEM APPLICATION

The Geosciences Node initially searched for a commercial or open-source application to meet the Archive Management System (AMS) requirement specifications. Applications used by internet hosting companies to verify website file integrity were considered, but these programs are typically intended for a much smaller number of files. After an extensive search for alternatives, the decision was made to create a custom application to meet the Node's specific requirements.

The Geosciences Node's AMS was created in a common development environment with the C# programming language. The data archive attributes are stored in a relational database. The database server provides an efficient platform for queries and data storage, with built-in data integrity checks for the database, and is configured with scheduled database backups.

The application is divided into three primary components (Figure 1). A graphical user interface (GUI) application provides for the operator interaction. A command line application executes the CPU intensive validation operations on network servers. The last component is the relational database, which contains the data archive configuration catalogs and validation result records.

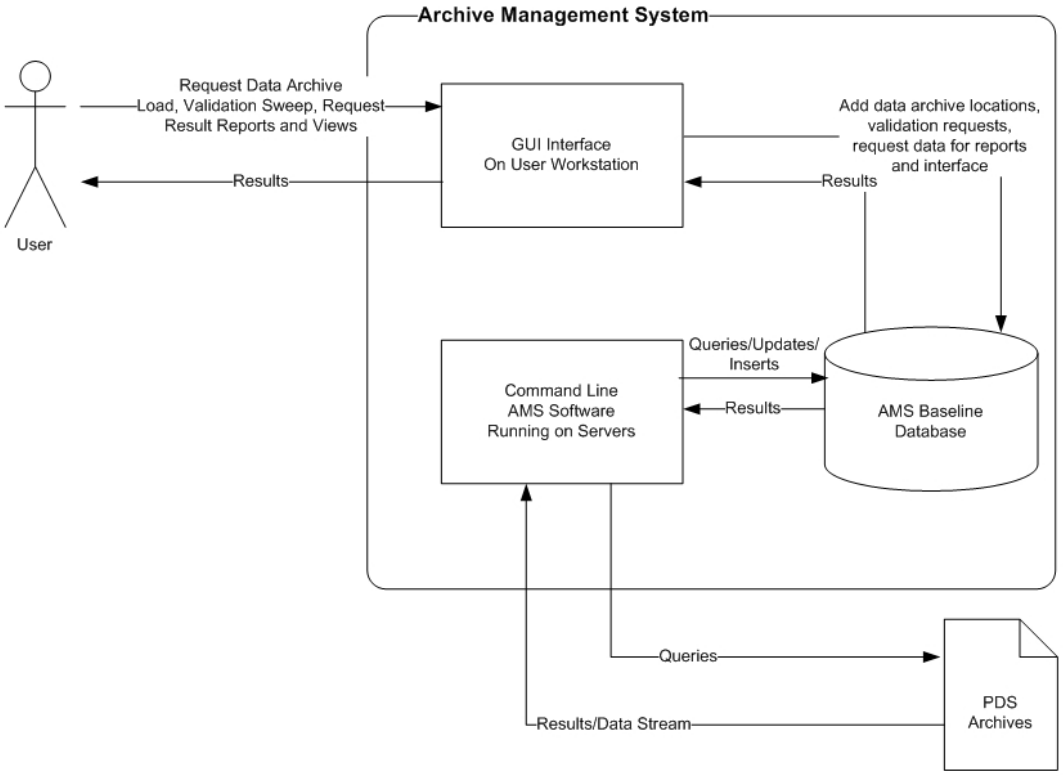

Figure 1: Geosciences Node's Archive Management System Component Model

The GUI application resides on the user's local machine and is designed to provide quick access to the most common functions and requests of a user. It contains functionality for interacting with the cataloged archives, adding new archive locations, adding validation jobs to the job queue, and retrieving report information. Only limited processing is executed in the GUI application on the user's workstation.

The command line application is installed and configured to execute on servers that directly interface with the network RAID disk system. This program executes jobs that have been added to a job queue from the GUI application. The application is scheduled to check for new processing requests multiple times a day as a server background task. Most processes are executed in a multi-threaded fashion to achieve greater performance. For example, a queue of the files to be validated in a data set archive is created. Then the application's sub processes (processing threads) begin pulling files from the queue to validate several files in parallel. The multi-threading greatly reduces processing time by executing several processes simultaneously and better utilizing the multiple CPUs of the servers. The processing threads are typically limited to the number of CPU cores of the server. The number of processing threads and other application settings are controlled by a configuration file included in each application installation.

The database uses a normalized relational database design, which protects record integrity and eliminates redundant data. This also makes the database easier to query and maintain. To optimize performance, the database makes extensive use of user-defined stored procedures. These compiled code procedures are utilized for often executed queries, and record inserts, updates, and deletes. Stored procedures can execute more efficiently because they are precompiled code that does not need to be analyzed by the database server before execution.

The Geosciences Node's PDS data archives are cataloged into the database to create archive baseline catalogs when the data are first delivered to the Node. Each PDS data set is saved as a separate archive catalog. The archive catalogs contain an initial "correct" snapshot of the data archive configuration. The catalogs contain records for each directory and file of the data set archive. Attributes including object name, modify date, size, relative path inside the data set, and file MD5 checksum are stored. Very large data sets, such as the 12 TB MRO CRISM TRDR data set, are divided by PDS volume for archive catalogs to allow for faster and more specific validations. Monthly archive validation scans are executed after the baseline catalogs have been created.

The archive validation process provides two levels of scans (Table 2). The full validation scan verifies all of the attributes of the archive including archive directory structure, file listing, directory file count, directory and file size, directory and file modify date, directory and file name case, and file checksum value. A quick validation scan is also available, which verifies all of the attributes of the full scan except the checksum value. The quick scan process can complete much faster and can identify most problems. Any differences of either validation scan will be flagged as a possible archive corruption. The error records at the file or directory level are recorded in a list containing all the differences that were discovered during a data set archive's validation.

Scan Type	**Attributes Validated**	**Advantages**	**Disadvantages**
Full Scan	Directory structure File listing Directory file count Directory and file size Directory and file modify date Directory and file name case File MD5 checksum	The most thorough process to validate archive attributes	Requires much processing time, network bandwidth, and network storage device activity
Quick Scan without Checksum	Directory structure File listing Directory file count Directory and file size Directory and file modify date Directory and file name case	Very fast processing speed Capable of identifying most accidental changes	Will not detect subtle corruption changes to the bytes of a file Will not detect a file modification that did not update a file's modify date or size

Table 2: Validation Scan Type Summary

When validating an archive, the results are classified into four categories (Table 3). The first includes all of the files and directories that have been verified as unchanged from the archive catalog, which are considered "correct". The second group contains files and directories that do not match the stored baseline archive catalog. This category includes corruptions, and directory and file changes. The changes can be legitimate if there has recently been a revised data release. The third category contains a listing of files and directories that were found during the validation scan, but do not exist in the catalog baseline stored in the database. These results are typically newly released data files for the PDS archives, but the list could include files that were errantly copied into the data archive. The last validation result category contains files that exist in the cataloged baseline, but no longer physically exist in the PDS data archive. This scenario can include files that have been mistakenly removed from the archive and files that were replaced with differently named files during a PDS data release.

Result Category	Result Interpretation
Correct Entries	Correct - Files and directories correctly validated against the archive catalog
Files or directories did not match the archive catalog	Correct - Revised data files and archive directories (new or revised PDS data release) Error Discovery – File corruption or modification Error Discovery – Directory modification or content changes
Files or directories found, but not in the archive catalog	Correct – New data files and archive directories (new PDS data release) Error Discovery – Files that were accidently copied into the archive
Files or directories in the archive catalog that do not exist on the disk	Correct – Revised data files and archive directories were removed and replaced with differently named recalibrated data (revised PDS data release) Error Discovery – Files were mistakenly or maliciously removed from the data archive

Table 3: Four Validation Result Categories

The Geosciences Node's data archivists notify the AMS operator when PDS data sets are updated. The operator can then process newly added data into the AMS and run specific validations based on this information. The validation results are usually easily reviewed and resolved by the AMS operator. Additional Geosciences Node staff members are queried to verify unusual results.

RESULTS

The Geosciences Node's Archive Management System has been in use for nearly a year. The application has successfully validated the entire Geosciences Node holdings ten times with file checksums, since it has gone into operation. All of the Geosciences Node's PDS data archives (40 TB) can be validated in nine days with this validation option. The application can currently validate maximum of 2 GB to 4 GB of data a minute based on various factors, including other application requests against the network storage device, the size and number of files in the data archive being validated, and the number of validation streams being executed. The quick validation scan, which validates all of the attributes except the checksum, can validate the entire archive in 18 to 28 hours. The speed of this validation is affected by other data requests from the SAN and the number of validation streams executed simultaneously. While various factors affect the performance of the AMS, the use of AMS has caused no noticeable performance impacts on other Geosciences Node's operations or system configuration. To date the system has found two accidental archive changes, which were quickly resolved. No file loses or corruptions have been identified.

FUTURE

The Geosciences Node's AMS has successfully met its outlined system requirements. Through the use of this system, the Geosciences Node staff can more confidently guarantee the integrity of its PDS archives. In the future, additional updates will be made to the application when new needs are identified.

At this time, the primary AMS operator is the application's developer. Eventually the operator role may shift to a data archivist.

The validation process performance will be analyzed in the near future to prevent processing from becoming too time consuming as the Geosciences Node's data archive holdings continue to grow.

Network switch configurations, server configurations, disk performance, simultaneous processing stream options, and possible application code modifications will be reviewed to identify changes that can be made to improve the validation performance.

REFERENCES

[1] – Berman, Fran: Developing Cyberinfrastructure for Data-Oriented Science and Engineering. Next Steps in Using Combustion Cyberinfrastructure Workshop (2007)

[2] - Stein, T.C., Guinness, E.A., Slavney, S.: Establishing a Mechanism for Maintaining File Integrity within the Data Archive. PV 2005 Conference (2005)

ACKNOWLEDGEMENT

This work is supported under NASA Grant NNG05GB73G, “Planetary Data System Geosciences Node.”