Building a COTS Archive for Satellite Data

Ken Singer, Dave Terril
Federal Data Corporation
4800 Hampden Land
Bethesda, Maryland 20814-2933
+1-301-457-5505
{ksinger, dterrill}@saa.noaa.gov

Jack Kelly
L.A. Systems
9027 Greylock Street
Alexandria, Virginia 22308-2726
+1-703-360-1000
jkelly@nesdis.noaa.gov

Cathy Nichols
NOAA/NESDIS/IPD
Federal Building 4
Suitland, MD 20746
+1-301-457-5243
cnichols@nesdis.noaa.gov

Abstract

The goal of the NOAA/NESDIS Active Archive was to provide a method of access to an online archive of satellite data. The archive had to manage and store the data, let users interrogate the archive, and allow users to retrieve data from the archive. Practical issues of the system design such as implementation time, cost and operational support were examined in addition to the technical issues. There was a fixed window of opportunity to create an operational system, along with budget and staffing constraints. Therefore, the technical solution had to be designed and implemented with regard to the practical issues. The NOAA/NESDIS Active Archive came online in July of 1994, meeting all of its original objectives.

Introduction

The functional requirements of the NOAA/NESDIS Active Archive were quite similar to most other archives. The NOAA/NESDIS Active Archive had to perform the following functions: 1) provide a means to manage and store a great number of large datasets 2) give users access to interrogate the archive 3) give users the ability to retrieve data from the archive. In addition, the following technical features were also desired: scalability so new and future datasets could be included, a modular architecture to allow enhancements, and security since the archive was intended to be accessed across the

Internet. All of these requirements and features could be implemented in a straightforward manner using hardware, software and a support staff focused entirely on creating the archive. The challenge faced in designing and implementing the NOAA/NESDIS Active Archive was to successfully accomplish the same task by using existing hardware to minimize cost, commercial off the shelf (COTS) software to minimize software development, and existing support personnel to reduce new staffing requirements.

The implementation of these functions and features had to be tempered by the fact that there was a window of opportunity to implement the archive, a limited budget, and other everyday work to still be accomplished. The greatest potential enemy was the goal itself, the design and implementation of the archive. If the design was too complex, it might take too long to implement and may never actually happen. If the design did not utilize existing hardware and software, cost might prohibit the project from moving forward. If "leading edge" became the buzzword for too many components, then the project would be throttled by the effort needed bring these components into an operational state. If the skills needed by programmers and the support staff were not available, time for the training would extend the time needed for development. If too many changes to operational procedures and the "usual way of doing business" were needed, management might not agree to the proposed changes. By considering these issues ahead of time and understanding their implications, the solution had to avoid these "potholes" as much as possible.

Characteristics of Archive-stored data

The initial purpose of the NOAA/NESDIS Active Archive was to store (level 1B) datasets from the Advanced Very High Resolution Radiometer (AVHRR) instrument flown on NOAA's current series of polar orbiting satellites. In the future it is very likely that additional datasets from the current satellites and new datasets from future satellites will become candidates for inclusion. The AVHRR datasets vary in size from 50 to70 Megabytes (MB), so 60MB is used as an average for calculations. Each operational satellite transmits approximately 45 datasets per day. Today there are 2 satellites, NOAA 12 and NOAA 14, downloading AVHRR data. So daily data volume is 5.4 Gigabytes (estimated).

Issues considered

The design of the NOAA/NESDIS Active Archive had to achieve a balance of the following issues: implementation time and complexity, overall system cost, commercial availability of hardware and software, reliability, future growth and scalability, and the migration path from existing systems

Solution Approach and Architecture

The solution selected takes advantage of the strengths of 2 different families of computers: the IBM mainframe and UNIX workstations. The mainframe offered high reliability, strong I/O capabilities, established connectivity to mass storage devices and time proven Hierarchical Storage Management (HSM) software. UNIX workstations were chosen for their reasonable price for performance, the availability of tools for developing user interface programs, and strong TCP/IP performance for Internet user access and data delivery. So the function of data management and storage would be done by the mainframe, and the functions of interrogating the archive and retrieving the data would be handled by the UNIX workstations.

The function of interrogating the archive is done totally by the UNIX workstations, with no assistance from the mainframe. This was done for the following reasons. First, the performance of interrogations of the archive would be more consistent by maintaining the database of metadata that describes each dataset on the UNIX workstation. The mainframe is heavily used by many other batch oriented jobs and thus has many periods of peak utilization. This could affect the response the user sees. Second, storage and generation of the browse images is done on UNIX. To assist the user in narrowing down the list of desired datasets, a browse image is provided on request. The underlying goal of interrogating the archive is to help the user narrow down and limit the number of datasets that are of interest. This saves the user time because only the data truly desired is obtained. And, it helps to minimize the load on the data storage and management function because fewer datasets are requested. Finally, reliability of the system should be higher because each function is independent. If the archive interrogation function is not running, the data management and storage function can continue to accept incoming data. Or if the data management and storage function is temporarily unavailable, the user can still interrogate the database and submit requests to retrieve data (although the request may take longer to fulfill).

This approach described above provides scalability. One of the true strengths of the mainframe is the attachment and throughput to storage devices. This strength will probably continue well into the future, allowing for growth. In addition it is possible to have the archive interrogation programs query multiple data servers, which would be another technique to increase capacity. Modularity is also emphasized with this approach. The user interrogation function is separate from the management and storage of data. Similarly, retrieving the data is accomplished without the user seeing or having to understand the storage and management of the data. These abstractions allow changes to be made to any components of the solution without changes to the other components. For example, new tools for interrogating the archive can be added without affecting the storage and management of the data. Or, new data storage devices can be utilized without the user having to change how the data is retrieved. Security is also provided because no user can directly access the data storage and management function. Users only interact with application menus on UNIX, which then cause other events to occur elsewhere in the system.

An additional benefit of this approach is that the NOAA CEMSCS mainframe was already there, already handling the ingest of these datasets, already connected to a mass storage device, and already had a knowledgeable support staff. So, part of the needed solution was in place and functioning. Plus, the processing that already existed on CEMSCS, such as creating other NOAA satellite data products, could utilize the data in the active archive as well.

While the hybrid approach offers many positive features, there are some tradeoffs. There is administrative overhead for coordinating the metadata database with the real archive. There could be missing entries or incorrect entries, each which would cause different problems. Also, operational support issues are more complex maintaining a system that spans across two different computing platforms.

Solution Overview

The description of the solution is based on the functional requirements stated earlier: the ability to interrogate the archive, the ability to retrieve data from the archive, and dataset management and storage. The description of interrogating the archive and retrieving data will be in the section below labeled "Interaction with the Archive", while the dataset management and storage description will be in a section of the same name. Each section will cover how the function is implemented, as well as a description of the hardware and software used.

Interaction with the Archive

User Interrogation of the Archive

A user wants to see what is in the archive based on set of criteria. For example, "Is there any data in the archive from the month of May,1994 over Greenland?" The metadata provides these answers. The NOAA/NESDIS Active Archive provides access to the metadata through an application called the Satellite Active Archive (SAA). SAA is responsible for collecting the metadata after the dataset is available on CEMSCS. In addition, SAA also provides a series of menus to allow the user to query the metadata and view the results. The results are a list of identifiers that point to specific datasets in the archive. As a final aid to the user, SAA provides a browse image for each AVHRR dataset in the archive. The browse image is at a lower resolution than the actual data, but should be of great value to the user. For example, by viewing the amount of cloud coverage in a scene, the user can further reduce the number of datasets that are of interest.

Retrieving Data from the Archive

After the user has selected some datasets of interest, now the user wants to go and get the data. Once again, the user interacts with the SAA application. SAA presents a series of menus to allow the user to order the datasets. The user can order the data for electronic delivery or delivery on tape media. In addition, the user can order the complete dataset or an extracted portion of the dataset. Presently SAA imposes a restriction on the amount of data that can be delivered electronically to insure reasonable network performance. The extract capability makes this possible because the user can limit the size of the delivered data by reducing the geographic area desired or by requesting fewer channels of data.

After the user is done interacting with the SAA application, the dataset management and storage function is finally called into action. A job is submitted to the dataset management and storage function to retrieve the requested data. First, the hierarchical storage management software gets the raw data where ever it may reside. Then the extract function is performed to cut out a piece of the data and add proper headers. Finally, the SAA application sees that the dataset management and storage function has the data ready and waiting. Then SAA picks up the data and delivers it to the customer in the requested manner.

UNIX Workstation functions

The IBM RS/6000 UNIX workstation running AIX was chosen to implement the UNIX based functions. Although most UNIX workstations could have performed the necessary functions, the RS/6000 was chosen for 3 main reasons: First, an existing contractor had strong knowledge and skills to support the RS/6000. Second, the RS/6000 offered ESCON channel connectivity to the CEMSCS mainframe. This could be used as a highly reliable, high performance point to point communication link using standard TCP/IP applications like FTP and NFS. Finally, the original basis for the front end application was the Global Land Information System (GLIS) from the US Geological Survey EROS Data Center (EDC) in Souix Falls, SD. GLIS had been ported to the RS/6000, so the local software developers had a strong head start.

The front end application (SAA) that the user interacts with utilizes both ASCII and Xwindows based screens. The SAA screens and menus were based on GLIS. Strong similarities can be seen between the systems today and may continue due to future cooperation. The cooperation and help provided by the EDC staff was of invaluable assistance in getting the SAA prototype off the ground so quickly. Similarly, the SAA metadata database uses INFORMIX, following the recommendations of EDC.

The electronic data delivery capabilities of SAA were a brand new function that had to be added since GLIS did not support electronic data delivery. The delivery functions are designed to provide reliability and flexibility for growth. Reliability was a necessity, since use of the NOAA/NESDIS Active Archive would probably be minimal if users could not be confident in receiving the data they have ordered. Flexibility for growth was important so the system could be up and running in a short time but easily add new functions such as FTP push and subscriber data delivery. Also, as usage grew, modifications may be needed to insure balanced network performance.

Workstation to Mainframe Communications

The solution architecture necessitates communications between the CEMSCS mainframe and the UNIX workstations on two occasions. The first is to update the metadata database and generate the browse images on the UNIX workstation. The second is to retrieve data from the physical archive for delivery to customers.

The update of the metadata database is accomplished the following way. The UNIX workstation wakes up at a regular interval (presently every hour) and does a directory listing (using NFS) of high level qualifiers where the AVHRR data resides. This list is compared to a list of datasets already in the metadata database. The MVS naming convention uses the Julian date, so only one days worth of data is examined at a time. If any new datasets are found, they are either copied to UNIX or processed across the NFS mounted directory. Some datasets can be processed across the NFS mounted directory because only the header needs to be read. A record of the metadata information is created and then entered into the database. At the same time, the browse image is created and stored.

The process for retrieving data is more complex. First, the SAA application on the workstation submits a request to the mainframe. This request is actually a JCL job that is submitted using FTP. The recall and availability of the dataset needed is handled transparently by the SMS/HSM software on the mainframe. The JCL job runs the extract to subset the dataset if needed. The result is placed in a particular directory and named by the SAA order number. A suffix is added to the name to indicate whether the job is in progress, successfully completed, or failed. Next, the SAA application on the workstation does a directory listing from the mainframe using NFS. If a file named with the SAA order number and the proper suffix is found, that file is copied to the UNIX workstation from the NFS mounted directory.

At this point, the SAA Delivery Server manages the delivery of the data to the user. The Delivery Server is a program that is modeled after a “state diagram”. Each order corresponds to a unique entry in the order database. The Delivery Server wakes up at a regular interval and performs a specific action on each order based on what “state” the order is in. For instance, a job may be submitted to CEMSCS, a check is made if the extract job is complete, the dataset is copied to UNIX, or the data is FTP’ed to the user. The Delivery Server has the ability to retry any state a specified number of times. If a failure is detected, a message is sent to the SAA system administrator.

Data Management and Storage

The Storage Server

The storage server used was the NOAA CEMSCS mainframe. CEMSCS is a multisystem complex of two IBM ES/9000 mainframes and associated peripherals. Workload is scheduled and resources are allocated by the IBM MVS Job Entry Subsystem (JES3). CEMSCS is used create level 1B datasets as well as many level 2 and level 3 products based on the raw data.

Software utilizd on the Storage Server

The storage server software selected on CEMSCS is based on IBM's mainframe System Managed Storage (SMS). The data archiving component of SMS is HSM, which has matured over twenty years of intense DP growth as a COTS solution. CEMSCS had previously elected to utilize HSM as a viable alternative to postpone and minimize expensive DASD acquisition. Since this product had proven successful in managing data for the CEMSCS environment, HSM was reviewed to see if it would satisfy the needs of the NOAA/NESDIS Active Archive. The concern was not the amount of data, which is measured in multiple terabytes, but rather the number of files retained. The number of files is in the area of many hundred thousand. The management of this large number of distinct files approached the limitations of HSM but recent changes addressed this concern. SMS has allowed CEMSCS to minimize personnel requirement, standardize storage retrieval and archiving methodologies, and isolate the installation from the ever changing hardware enhancements.

SMS attempts to optimize placement of data, according to installation directives. This process maximizes automation and minimizes the staffing requirements, but initially requires a higher level of expertise. Complex issues, such as data reference patterns, locality of reference, read to write ratios, etc., are minimized but not eliminated. Once the directives are established, the day to day process requires less expertise. SMS allows minute to minute evaluations about data to be made without requiring manual intervention. while minimizing data access time.

HSM manages the retention and migration of data, again according to installation directives. HSM attempts to ensure that frequently referenced data is maintained on accessible storage while less frequent referenced data is maintained on alternate less expensive storage media. Data may be migrated to less expensive DASD or tape media, depending upon installation criteria, e.g. data, size, importance, or residency time. Data compression can be optionally performed on the migrated data at either the software or hardware level.

The retrieval of data, i.e. moving data from a lower form of the hierarchy to a higher one, is performed with no user intervention. If the data has been migrated and it is referenced, then HSM automatically moves it to an accessible media. If the data has to be brought back from a non DASD device, then a user can be notified that the retrieval may required an "extended" amount of time. With the inclusion of tape robotics, this extended time is less than ninety seconds.

Archive Storage Devices

Access to data must be accomplished from DASD. Once SMS has placed a file on an appropriate DASD device, the file remains on this device until it is migrated by HSM or deleted. Several different DASD media have been utilized at CEMSCS. DASD caching maintains response time, especially for the larger capacity devices. SMS allowed CEMSCS to easily create "pools" of DASD to satisfy the different requirements of the archive. In this context, a pool is simply a grouping of DASD for a specific purpose. As the archive development evolved and moved into operations, so did its requirements. To date, these changes have been easily addressed via SMS mechanisms.

CEMSCS created two pools of DASD for the archive. Initially the satellite data is placed in a pool of four IBM 3390-II and three IBM 3380-III where it resides for one to three days, depending on access. If the data is migrated from this initial pool and subsequently accessed then it is recalled to a separate pool of five IBM 3390-II. This smaller pool has a different migration and residency criteria than the initial pool. The archive concept is that current data is more likely to be accessed: therefore, keep it accessible and do not waste resources migrating it. After time, data is less likely to be accessed; therefore, the data can be migrated. If data is recalled, then the data is placed on the second pool. This method allows recalled data not to impact the management objectives of the current data.

Cost effectiveness has convinced CEMSCS to migrate to robotic tape subsystems. Two different vendors' robotics have been utilized, as well as two different tape media, IBM and STK 3480 and IBM 3490E. The 3490E media is the media of choice for HSM's migration function. The 3490E has the capacity required for the archive and the 3490E has the ability to located records on tape directly. HSM "understands" and takes complete advantage of these hardware enhancements of the 3490E device. Due to the nature of the satellite data, the hardware compaction (IDRC) available with the 3490E does not provide much benefit, less than three percent.

Currently SAA has captured a terabyte of data comprising of 19K files. A subset of this data resides on the two DASD pools of 16 GB. The remainder resides on 1,100 3490E volumes.