A New Approach for Data Retrieval from Hierarchical File System to Handle Priority Satellite Data Products Dissemination

C Pradeep[1], ANSV Murali Krishna[2], G Prasad[3]

[1]National Remote Sensing Centre, ISRO, Department of Space Hyderabad, Telangana, India (Phone: 9573908556; Email: pradeep_c[at]nrsc.gov.in)

[2]National Remote Sensing Centre, ISRO, Department of Space Hyderabad, Telangana, India (Phone: 9986121142; Email: muralikrishna_ansv[at]nrsc.gov.in)

[3]National Remote Sensing Centre, ISRO, Department of Space Hyderabad, Telangana, India (Phone: 9848017965; Email: prasad_g[at]nrsc.gov.in)

Abstract: ***Satellite Data is being acquired by various Indian Remote sensing satellites and being widely used for various applications. Applications include resource management, disasters, crisis-management and relief support. These applications demand immediate and timely satellite data products support which is of paramount importance in order to make predictions, decision making during disasters like floods, Earthquake, landslides, forest fires, etc. and are considered as Priority products. These Priority products are to be supplied within an hour from the time the request is received from the order processing system. The entire process of ordering, processing the requested data imagery and disseminating the product to the requestor is entirely digitalized. The work flows and data processing is handled by a chain of software applications running on top of the respective IT infrastructure chain in the IMGEOS data center. SAN-large-scale data storage which is one of the major components of IT infrastructure is shared by different operational entities functioning in the organization. The resource sharing model implementation has an impact on processing and dissemination of Priority data products. One of the factors for the delay in data dissemination is translated as the delay in data retrieval from the underlying data storage systems viz., Disk and tape storage medium. This paper presents a customized software framework developed for the efficient data retrieval of the priority data sets from the wide range of archived data store. This is envisaged by lodging multiple files under single data request and enables special priority at the retrieval queue to reduce the request queue latency, which has helped in maintaining the overall Turn-around-Time (TAT) for the product generation and dissemination. Additionally, it improves retrieval requests queue length and improves the health of storage hardware.***

Keywords: DM, IMGEOS, HSM, IT, LTO, LUN, LVM, MD, MSM, NRSC, OS, RF, SAN, TAT, TSM, WOF

1. Introduction

IMGEOS (Integrated Multi-mission Ground Segment for Earth Observation Satellites) at NRSC is conceptualized with the aim to achieve multi-satellite data acquisition, data processing and dissemination operations on resource sharing basis. The satellite ground segment chain comprises of Radio Frequency (RF) tracking systems, Data acquisition and data processing systems for the different types of satellites data being received in IMGEOS. The data center consists of the compute, storage and network Infrastructure to acquire process, archive and disseminate the data to national and international user community.

The data acquired from the satellites is being archived for the life of the mission and will be maintained in the archives even after the mission is declared end of life, so as to enable the scientists to process data up to different levels for analyzing, comparing data for applications like disaster, relief measure, change detection etc. As the data rates of the new launched satellites are very high the incoming data volumes are increased multifold and has to be archived i.e., about 2 TB / day. This flow of incoming data for archiving will keep increasing by several hundreds of gigabytes every day, which is the biggest challenge for storage domain. To cater the needs of such increasing demands, storage system was designed with the combination of disk and tape, using disk for short-term storage and then transferring data to tape for long-term retention, as this can suffice the mission requirements and user needs. For permanent data archival, the storage system is designed to take up on robotic tape library that stores digital data in magnetic medium known as LTO (Linear Tape Open).

As the data is being archived in three tiers, data retrieval especially for priority products during emergency is required to be handled to ensure data availability for the data processing center, dissemination center, application scientists and user community as per the requests and within the turnaround time specified. The requests handled by the storage system is in lakhs every single day catering to all the satellite data processing centers as the data center resources are shared by all the processing centers. To meet the urgent data requests especially during disasters, relief works data should be disseminated to the user within an hour and this poses stringent timing requirement on retrieval from the three tier storage. With the increasing demand for remote sensing data during the natural calamities, there is an increase in the amount retrievals for data from the Hierarchical storage. Hence a frame work of data retrieval for priority data request is designed.

Licensed Under Creative Commons Attribution CC BY

2. IMGEOS Storage Area Network

The storage architecture for storing and archiving large volumes of satellite data is designed with extensible features to support availability, accessibility, reliability, expandability, integrity, retention, compliance and security as shown in Figure 1. All these requirements are provided under a balanced storage paradigm known as the Storage Area Network i. e., SAN. SAN is a specialized, high speed network that provides access to data blocks known as LUNs (Logical Unit Number). SAN storage is typically composed of switches, storage elements, disk arrays and tape drives / libraries that are interconnected using multiple technologies, topologies and protocols. The LUNs are presented to the host computing systems as block devices using iSCSI protocol. A File system configured with HSM (Hierarchical storage management) capabilities was implemented to provide complete storage virtualization under multiple layers of the system. The HSM handles automated migration of data objects among the storage devices, usually based on data activity. In our case we have implemented the three TIER SAN architecture with TIER I & Tier II are disk based storage arrays and TIER III a robotic tape library. These policies are framed for the purpose of systematic identification, categorization, maintenance and retention of satellite data received or products generated in the SAN environment.

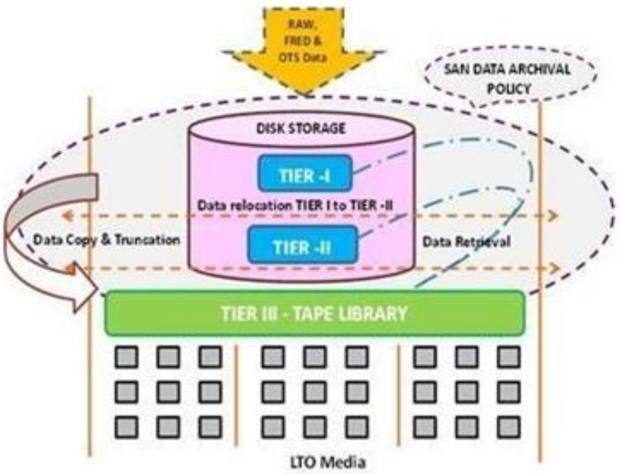

Figure 1: Overview of IMGEOS SAN Architecture

The whole SAN architecture and it operations are configured into 3 layers as 1. Presentation layer 2. Management layer 3. Data layer as shown in Figure 2.

Presentation Layer:

The presentation layer establishes the way in which the data is presented, typically displayed at the hosts / work centers. The Filesystem client agent is installed on these hosts / work centers; therefore it translates information about the data in a way that host Operating System (OS) and applications can understand. The detailed view of data archival structure, type of data, data owners and user permissions, the data blocks / LUNs are presented to the users under Filesystem labels. The mapping context between user, application and the data under the Filesystem label is presented seamlessly irrespective of all data TIER in which it is archived.

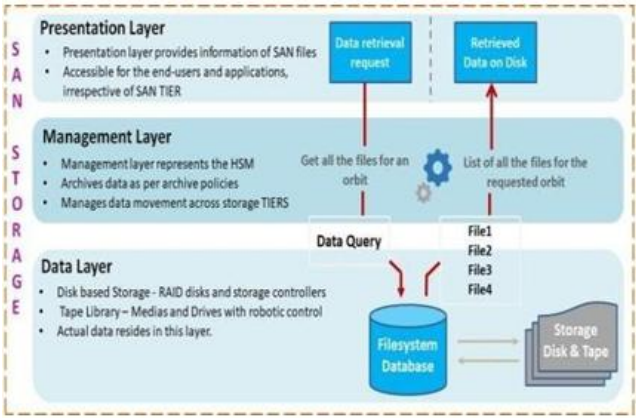

Figure 2: Layers of SAN

Management Layer:

The storage manager the key component that handles the management part of SAN storage, management is tedious at this higher level of aggregation through the use of Device Mappers (DM), Logical Volume manager (LVM) and Multiple Devices (MD). The storage manager alleviates this by providing a friendly unified user interface, which allows applications, administrators and users to run complicated tasks. The storage manager delivers full data lifecycle management from data creation till the end. As the data moves across the three TIER storage system, the manager provides continuous data access.

The data archival polices are created, implemented and maintained with the support of Hierarchical Storage Manager (HSM). The HSM controls the data movement across storage tiers**.** It is responsible for the identification of the tier in which a particular data is stored. It takes care of the data retrieval from LTO media in the Tape library on to disk storage and truncates data from disk. It acts as a software controller for the mechanical drivers in the tape library**.** It works in coordination with the other two lesser level Media storage manager (MSM) and Tertiary storage manager (TSM). It handles the integration of itself with other software components responsible for tape library operations**.** It contains the master database records for the entire SAN storage infrastructure. The master database includes the minute level of information such as LUN IDs of disk storage, starting and ending of data blocks and Media related identifiers. Storage manager will direct and interfaces the storage agents installed in the client systems with the storage server. The storage manager will take care of file migration between storage tiers**.** Keeps track of files/directories creation, modification and deletion.

Data Layer

This layer is the physical file system layer and is concerned with the physical operations of the storage devices. This layer consists of disk controllers, disk arrays, LTO media, and robotics, switching and interfacing components. In this layer clients are physically interfaced via connections to create zones with the storage units. After the zoning process mapping of the data blocks

Licensed Under Creative Commons Attribution CC BY

/ LUNs to the client systems are performed here. There are separate GUI interfaces for the disk controllers, which will be used for mapping LUNs / data blocks to the destined client system. The actual data resides here in its native form.

3. Data Processing and Dissemination

By utilizing the IMGEOS IT infrastructure raw data of Satellite is acquired and processed for correcting various errors and distortions introduced by its sensors, platform, atmosphere, etc. The data has to be packed and presented in specific format with in the specified accuracy (radiometric/geometric) for utilization. Data processing involves applying high level algorithms to the raw data and converting it to meaningful information products for use by agencies/scientists.

During data dissemination the data will be stored in the dissemination server for the user to download. Data exchange systems use both „push" and „pull" methods to get and post data from data processing centers from the internal network to the external network server in the Internet domain. The Data exchange systems use specialized OS kernel to ensure secured data exchange.

4. Existing Method of Data Retrieval and Flow of Priority Products

The current dissemination system for priority products is of the client-server type. The functional flowchart for the data product generation is shown in Figure 3. The user formulates a request based on his area of interest in the User Order Processing System portal. The request is sent to the related data flow system. The work flow manager parses the request and generates a work-order file against the request. The fetched information along with the work-order file is sent to the respective data processing centers. The scheduler unpacks the work-order and checks for the data product in the SAN storage. If the data is available on disk the data is made available to the scheduler to complete the work-order request. If the requested data file / data set is available on tape archives, then it has to retrieve the data from the tape on to the disk. After data is available on the disk, it is made available to the scheduler to complete the work-order request. The generated product is transferred to the external network through a Data Exchange Gateway system. A reply packet that indicates the availability of data product and delivery information is sent to the user with the credentials.

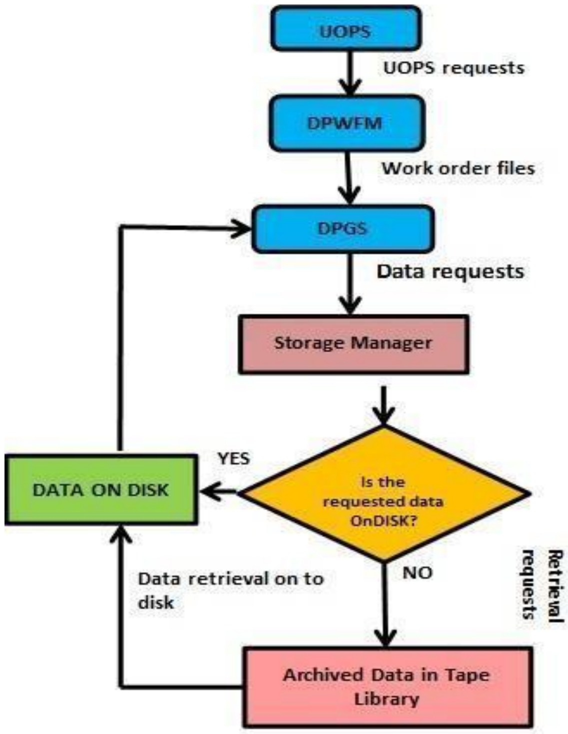

Figure 3: Data request and retrieval mechanism

Existing method of PRIORITY / EMERGENCY Data Retrieval

As the data centers handles millions of operations the storage domain has request priorities for various internal processes. The storage servers systems are ignorant about the data request type that is received from the Data Processing Work-Flow Manager. As per the dissemination policy the PRIORITY / Emergency products are to be delivered within 1 hour to the end-user. It has been observed that the delivery of priority products took exceptionally longer time to reach the intended system. One of the factors for the delay in data dissemination is translated as the delay in data retrieval from the underlying data storage systems viz., Disk and tape storage medium. To address the delay of data retrieval from tier III, we have considered two practical scenarios of data request handled by storage systems.

Scenario 1: If the requested data is available on disk, then the end-user, application can access the data immediately.

Scenario-2: If the requested data is on the tape media, a retrieval request is raised for the data to be retrieved from the tape media. Then, the retrieval request hits the storage manager; it gets into the queue and waits for its turn for retrieval. If there are multiple files to be retrieved for a product, then there will be multiple requests and the waiting / retrieval time will be significantly higher.

As there are numerous requests handled by the storage manager, the PRIORITY product request doesn"t get any kind of privileges among the other requests. Hence, the retrieval requests for the PRIORITY product are placed in queue along with other requests. The prime issue is the time constraint not being able to deliver the PRIORITY Products in an hour. As the priority request is also like any

Licensed Under Creative Commons Attribution CC BY

other request to the storage manager, it gets accumulated as a member of the retrieval queue requests thus, increasing the request queue length, which will in turn delay the retrieval of the priority requests in the queue as shown in Figure 4. .

SAN PATH =/ots/APA/KGS/K3/A/2018/FEB/K3A19FEB20180307330000010035KGSSTUC00GODF							
S.No.	REQUEST TIME	COMPLETION TIME	REQUEST ID	FILES IN SAN PATH	MEDIA ID	DRIVE ID	NO. OF FILES
1	12:16:27	12:26:35	34501606	BAND_META.txt'	NR0116	/dev/sg709	1
2	12:26:35	12:38:57	34518984	K3A19FEB20180307330000010035KGSSTUC00GODF_meta.xml'	NR0116	/dev/sg710	1
3	12:38:57	12:52:44	34531979	K3_20180219081320_30730_05351077_L1G_br.jgw'	NR0116	/dev/sg926	1
4	12:52:44	13:05:15	34542398	K3_20180219081320_30730_05351077_L1G_R_rpc.txt'	NR0116	/dev/sg222	1
5	13:05:16	13:14:00	34551557	K3_20180219081320_30730_05351077_L1G_N_rpc.txt'	NR0116	/dev/sg710	1
6	13:14:00	13:26:27	34557829	K3_20180219081320_30730_05351077_L1G_R.tif'	N10278	/dev/sg928	1
7	13:26:28	13:34:51	34566604	K3A19FEB20180307330000010035KGSSTUC00GODF.jpg'	NR0116	/dev/sg928	1
8	13:34:51	13:47:01	34572437	K3_20180219081320_30730_05351077_L1G.shx'	NR0116	/dev/sg709	1
9	13:47:01	14:01:05	34580910	K3_20180219081320_30730_05351077_L1G_P.tif'	NR0116	/dev/sg438	1
10	14:01:22	14:11:25	34590685	LevelProductSchema.xsd'	NR0116	/dev/sg928	1
11	14:11:25	14:22:30	34597527	K3_20180219081320_30730_05351077_L1G.shp'	NR0116	/dev/sg222	1
12	14:22:30	14:32:18	34604814	K3_20180219081320_30730_05351077_L1G_Aux.xml'	NR0116	/dev/sg928	1
13	14:32:18	14:40:53	34611016	K3_20180219081320_30730_05351077_L1G_G_rpc.txt'	NR0116	/dev/sg439	1
14	14:40:53	14:48:52	34616293	K3_20180219081320_30730_05351077_L1G_G.tif'	N10278	/dev/sg439	1
15	14:48:52	14:58:01	34621247	K3_20180219081320_30730_05351077_L1G_N.tif'	N10278	/dev/sg926	1
16	14:58:01	15:06:57	34627119	K3_20180219081320_30730_05351077_L1G.dbf'	NR0116	/dev/sg709	1
17	15:06:57	15:14:39	34632791	K3_20180219081320_30730_05351077_L1G_P_rpc.txt'	NR0116	/dev/sg439	1
18	15:14:39	15:20:39	34637404	K3_20180219081320_30730_05351077_L1G_B.tif'	N10278	/dev/sg438	1
19	15:20:40	15:30:14	34641190	K3_20180219081320_30730_05351077_L1G_B_rpc.txt'	NR0116	/dev/sg709	1

Figure 4: Results showing TAT with existing method

While retrieving multiple requests for a single product, the media is loaded and unloaded several times, leading to technical concerns because of multiple read/write of the LTO media. LTO media wear out after repeated load/unload cycles, such wear can cause an increase in media errors and leads to media failure. Tape Drive Read/write recording head utilization hours increases the MTBF, which will degrade Tape drive performance and then leads to failure. Mechanical movement of the Robotic arms increases the MSBF, which may lead to failure.

5. Designed Solution and Implementation

Considering the time taken in retrieval of Priority products and technical concerns imposed by the earlier approach on the tape library, the following idea is proposed.

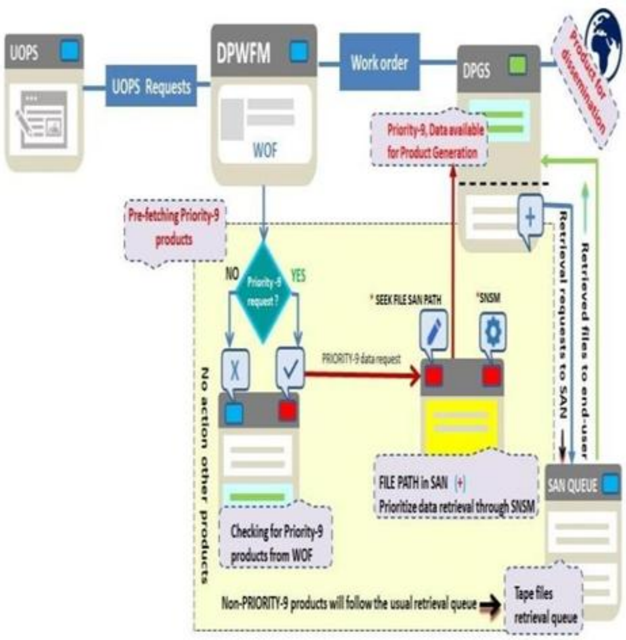

Figure 5: Idea to better handle Emergency data requests

As shown in the Figure 5. A new approach is designed in which, data retrieval of Priority products will be done in two steps without disturbing the conventional method. In the new approach, the script will constantly poll for PRIORITY products from the WOF being received. When Priority products are detected, the SAN data path from the workflow manager is obtained and data retrieval of that requests with the aid of storage manager software will be taken up immediately. The retrieval request given for File/Files will be assigned as one single request by the storage manager software. File/Files get retrieved on to disk at one instance instead of multiple interleaved file retrieves. This application is primarily focused on retrieval of PRIORITY products, and can be extended to other requests which include bulk data retrievals.

Advantages on implementing the solution to the system:

By adopting the new method reduction in high request polling to the storage manager, reduction in request queues, and reducing the waiting time of requests was observed. The new approach keeps the Tape media load / mount count under threshold and extend its course of usage. This increases sustainability of Tape Media, when there is reduced no. of winds and un-winds. Relatively, reduces number of read / write errors and mount /un-mount errors that prevent the tape drive going offline. Reduces the number of Tape drive usage hours and prevents pre-mature tape drive failures. Reduces the Mechanical movement of the Robotic arms and extends its life span.

The following case study explains and compares the advantages achieved by retrieving data of 25 GB from tier III by the old method and new approach. With the existing approach initiated data retrieval from tier III which took approx.3 hours and 5 minutes for retrieving the 25 GB of data.

From, the above table it is seen that 19 data files are required to generate a data product. In this particular case, 14 out of 19 files are from the same cartridge / LTO media; but they are getting retrieved on multiple instances leading to time delay in forming a complete data set.

The same data set of 25GB was retrieved from tier III using the new approach and in the output shown in Figure 6 it clearly indicates the improvement in retrieval time using the newly developed method. It was observed that the time taken for retrieval of the same set of files has reduced from 3 hours and 5 minutes to only 10 minutes, which is 18 times faster.

```
[root@MDS03 TapeReports]# date;fsretrieve -B -a /tmp/30377.txt
Thu May 27 14:48:13 IST 2021
FS0655 27 0043855358 fsretrieve interim: Currently processed 0 out of 19 files.
FS0390 27 0043855358 fsretrieve completed: 19 out of 19 retrieves were successful.
[root@MDS03 TapeReports]# date
Thu May 27 14:58:00 IST 2021
```

Figure 6: Results showing TAT with new approach

Licensed Under Creative Commons Attribution CC BY

6. Conclusion and Future Scope

When considering the possibilities of developing a customized application for this particular aspect of improving the TAT on retrieving and delivering the PRIORITY / EMERGENCY Products. A combinational collection of shell scripting and java programming has aided to develop the software that meets the exceptional requirements that will not affect the existing chain and setup in the organization. These also provided inputs for future data archival methods that can be adopted i. e. by archiving the data sets required for a product in bundled form i. e. TAR/zip/ the data retrieval process can be further optimized i.e. one request for one product. The software further can be extended to other type of user requests where the urgency can be on top of the other requests in the retrieval queue on adhoc requests from users. The activities handled by this software logged in a database and displayed in a dashboard with statistics for better data management and aggregation of storage space.

References

[1] IMGEOS Project Report, NRSC-DPA-IMGEOS-Apr09-TR64.
[2] IMGEOS Infrastructure Document, NRSC-DPA-IMGEOS-JAN10-TR147.
[3] IMGEOS Software Interface Control Document, NRSC-DPA-IMGEOS-JAN10-TR141.
[4] Preliminary Design Review Document, NRSC-SDAPSA-FEB-2014-TR-584.
[5] Data Ingest and Processing Interface control Document C2C/NRSC/Level
[6] -0/ICD/Ver-1.0/May 2015.
[7] Work Flows Interface Document, NRSC/DPPA&WAA/SG/C2C-DPWFM/ICDV1.0/Dec2015.
[8] Critical Design Review Document, NRSC-SDRISA-JAN-2016-TR-788.
[9] SNIA"s "Computational Storage Architectural and Programming Model" v0.8 rev0 August 2021.
[10] SNIA"s "education /tutorial s/ 2008 – Storage tiering for file systems.
[11] Quantum Tiered Architecture and software 6-68507-01 v5.4.1 July 2017

Author Profile

C. Pradeep received B. E. degree in Electronics and Communication Engineering from Anna University, Chennai and M. Tech degree in Software Systems from Birla Institute of Technology and Science, Pilani. Email: pradeepcps4[at]gmail.com

ANSV Murali Krishna, received B. E. degree in Electronics and Communication Engineering from Osmania University, Hyderabad, Email: murali.annadanam[at]gmail.com

G. Prasad received M. Tech degree in Advance Electronics, from Jawaharlal Nehru Technological University, Hyderabad, Master of Business Administration in Operations Research from Indira Gandhi National Open University, New Delhi and Ph. D. in VLSI Design from Jawaharlal Nehru Technological University, Hyderabad. Email: gurram_p[at]yahoo.com

Licensed Under Creative Commons Attribution CC BY