

UN RETOUR D'EXPERIENCE SUR LA STRATEGIE D'EVOLUTION DU SYSTEME D'ARCHIVAGE STAF (Service d'Archivage et de Transfert de Fichiers).

Anne JEAN-ANTOINE PICCOLO

Centre National d'Etudes Spatiales (CNES)
18, Avenue Edouard Belin
DEE/IR/ISM/IS
31401 TOULOUSE CEDEX 4

Anne.Jean-Antoine@cnes.fr

Résumé

Le Centre Informatique du CNES propose à la communauté des scientifiques et responsables de missions spatiales du CNES (et de ses filiales ou partenaires) des moyens logiciels et matériels pour assurer la conservation à long terme des données scientifiques et techniques. Ces données sont de différentes origines : des données satellitaires (télémesures, données brutes et produits à valeur ajoutée), des données de simulation issues de calculs (orbitaux, altimétriques, géodésiques, ...) et des données d'expérience de toute nature.

Depuis 1995, le Centre Informatique propose le Service d'Archivage et de Transfert de Fichiers (STAF) en tant que service garantissant sur le long terme la pérennité physique des données sous forme de fichiers. Ce service offre des fonctions élémentaires d'ingestion et de diffusion reposant sur un transfert binaire des fichiers, quel que soit le format adopté par le projet spatial. La pérennité du format reste de la responsabilité du projet.

Ce service a évolué sur ces 8 dernières années en fonction des besoins nouveaux des projets spatiaux. En effet, les besoins en volumétrie sont passés de quelques giga-octets globalement stockés en 1997 à une cinquantaine de téra-octets constatés actuellement. Le service STAF, s'appuyant sur une architecture technique conçue il y a une dizaine d'années, a dû faire face à cette progression exponentielle des volumes. De plus, l'évolution de la technologie de stockage associée au système STAF, reposant sur de la technologie des gammes IBM et StorageTek a largement contribué à adapter le système à la progression de la volumétrie stockée.

Face à cette montée en charge de l'utilisation des moyens du STAF, le Centre Informatique a mis en place une stratégie de stockage prenant en compte des besoins très hétérogènes des projets spatiaux (gestion de fichiers de faibles versus de fortes tailles). La communication s'attachera donc à présenter un retour d'expérience critique sur la solution actuelle en ce qui concerne sa capacité à s'adapter à l'évolution technologique permanente et aux besoins fonctionnels des projets spatiaux. Elle abordera dans un deuxième temps les perspectives technologiques et l'état de réflexion du Cnes sur la rénovation du système STAF à l'horizon 2004.

Abstract

The CNES Computer Center provides hardware and software to ensure long-term preservation of scientific and technical data. Users are data managers who belong to CNES or its partners (subsidiaries, search labs). Data are the results of varied experiences as satellite telemetries, simulation data (in orbital, geodesic, altimetric fields) and other added value data.

Since 1995, File Transfer and Archiving Facility (named STAF in French) functions have consisted in providing physical keeping of data files. It offers elementary archiving and retrieving facilities based on a binary file transfer, whatever the format may be. Data managers are responsible for data format preservation.

For 8 years, due to spatial project needs, the archival system has been in constant evolution. Stored data represented a few gigabytes in 1997 and today, the system stores more than 50 terabytes.

The STAF system is based on a technical architecture, which was designed ten years ago : it has to cope with a very important rise of stored volume. In addition, the STAF storage architecture, which is mapped on IBM and StorageTek technologies, has gained from technological developments to adapt the archival system to a permanent volume rise.

Faced with that rising use of STAF facilities, the Computer Center has worked out a storage strategy in order to consider heterogeneous project needs (about little and big data file storage).

This presentation will deal with a critical experience feedback of the STAF technical solution, about its adaptability to permanent technological developments and new functional needs. It will set out, in a second time, lines of technological prospects and the state of thinking in the Cnes, about STAF system renovation to 2004.

1. Présentation du STAF

Le Service de Transfert et d'Archivage de Fichiers est proposé par le Centre informatique du Cnes depuis 1995 aux projets scientifiques et techniques en vue de préserver leurs données sur le très long terme.

1.1. Sa mission

Le STAF est un service d'archivage et de restitution de fichiers qui prend en charge, de façon transparente pour l'utilisateur final, la gestion des fichiers et des supports associés. Sa mission est d'offrir :

- le stockage structuré des données sous forme d'arborescences de fichiers,
- la pérennité des données et la gestion transparente des supports physiques,
- la confidentialité et l'intégrité des données,
- la restitution et la mise à disposition des données dans des environnements techniques hétérogènes.

1.2. La nature des données archivées

Ce sont des données à caractère scientifique et technique, d'origine diverse : données d'observation de la Terre, télémessures brutes, résultats d'expérience, données de calculs de simulation, Généralement, ces données sont uniques, non reproductibles et nécessitent d'être préservées sur le long terme car elles constituent les données de référence du domaine scientifique correspondant.

Le STAF archive un train de bits et conserve le format initial de la donnée. Ainsi, la mise au format neutre des données selon les standards en vigueur reste de la responsabilité des utilisateurs.

1.3. Les fonctions offertes

Les fonctions offertes à l'utilisateur final sont sous la forme d'une ligne de commande identique quel que soit le système d'exploitation du moyen de traitement de l'utilisateur. D'abord disponible sous les systèmes Nos/Ve (Control Data), Unicos (Cray) et Solaris (Sun), cette ligne de commande a été portée sur HP-Ux (HP) et très récemment sur Aix (IBM).

Les fichiers archivés sont structurés dans une arborescence privée à un projet spatial. En plus des fonctions d'archivage et de restitution de fichiers, STAF met à disposition un ensemble très complet de fonction de gestion des données tels que l'attribution de droits d'accès, la caractérisation des fichiers, la navigation dans l'arborescence des fichiers et la gestion des utilisateurs propres à chaque projet.

1.4. Le principe de stockage

L'utilisateur précise au moment de l'archivage, pour chacun des fichiers indépendamment de sa localisation dans l'arborescence, un délai de mise à disposition du fichier et un nombre de copies souhaitées. Ces deux caractéristiques définissent une classe de service (CS).

En face de chaque classe de service, le Centre Informatique fait correspondre des supports de stockage dont la gestion (localisation, identification, changement, ...) reste transparente pour l'utilisateur final.

A l'heure actuelle, six classes de service sont disponibles :

TAB. 1 : Classes de service du STAF

Classe de service (CSx)	Duplication (Oui/Non)	Délai maximal de mise à disposition sur le réseau	Taille maximale conseillée	Média correspondant en juin 2002
CS1	O	Accès rapide	< 50 Mo	Disques magnétiques
CS2	N	< 1 Minutes		
CS3	O	Accès moins rapide	> 50 Mo	Cartouches near-line de 400 à 800 Mo – Lecteurs 4490 et 9490.
CS4	N	< 8 Minutes	< 350 Mo	
CS5	O	Accès moins rapide	> 50 Mo	Cartouches near-line de 20 Go – Lecteurs 9840.
CS6	N	< 10 Minutes	< 1Go	

1.5. L'architecture informatique

Le service STAF repose sur une architecture client serveur basée sur le protocole de communication RPC (Remote Procedure Call) pour gérer les dialogues et assurer le transfert des fichiers de données. Le serveur STAF fonctionne sur un ordinateur IBM (système OS/390) et présente deux composantes :

- **Le gestionnaire de l'application**, qui assure la gestion des entités fonctionnelles (projet, utilisateur, fichiers, ..) et des données référentielles (correspondance nom de fichiers utilisateur / nom de fichier stocké). Il s'appuie sur une base de données DB2 d'IBM.
- **Le gestionnaire de stockage**, qui assure la gestion du fichier stocké et la mise en oeuvre des classes de service. Il repose sur une hiérarchie de média sous DFSMS/DFHSM d'IBM : le premier niveau assurant un accès direct (*on-line*) se base sur des disques magnétiques, le niveau suivant moins performant se base sur des cartouches accessibles automatiquement (*near-line*) depuis des bibliothèques robotisées.

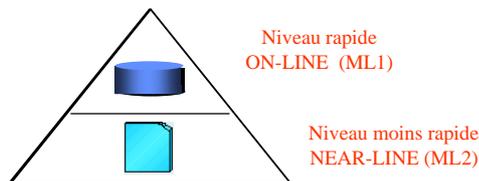


FIG. 1 : Hiérarchie du stockage dans le STAF

De par ses fonctions, le « hiérarchiseur de média » (DFHSM d'IBM) consiste à migrer automatiquement en fonction de sa classe de service, le fichier du niveau ML1 vers le niveau ML2 et à l'inverse, permet le rappel en toute transparence lorsque l'appliquatif STAF accède à un fichier migré.

2. Evolution de la configuration du service STAF depuis 1995

2.1. Historique

De 1995 à 1998, le Centre Informatique du CNES propose le service STAF sous la forme d'un seul serveur multi-projets pour assurer le stockage des données scientifiques. Le service est alors sollicité pour sa fonction d'archivage car les premiers utilisateurs ré-habilitent l'archive existante au CNES pour la conserver dans le STAF. La volumétrie des premières données archivées au STAF est faible en comparaison de celle stockée à l'heure actuelle. La taille des fichiers varie de quelques kilo-octets à quelques méga-octets. L'espace de stockage associé au serveur est constitué de média de faible capacité – cartouches au format 3480/3490E (de 250 Mo à 800 Mo) rangées dans deux bibliothèques automatisées StorageTek 4400 (11000 emplacements).

En 1998, le CNES décide de conserver sur les moyens du STAF, l'imagerie SPOT (Satellite Pour l'Observation de la Terre). La taille moyenne des fichiers est de 80 Mo et peut atteindre 300 Mo. Le service STAF s'étend alors sur un nouveau serveur réservé à l'archivage des télémétries image de Spot. L'espace de stockage, qui doit croître à raison de 1.7 Téra-octets par mois, est dans un premier temps constitué de média de 10 Go (cartouche Redwood) rangés dans une nouvelle bibliothèque automatisée de 5500 emplacements. Les nouveaux lecteurs 9840 (20Go) de StorageTek présentant une meilleure fiabilité, sont mis sur le marché en 1999 et remplacent les lecteurs Redwood : la migration technologique est effectuée sans perturber l'archivage en cours de l'archive numérique Spot.

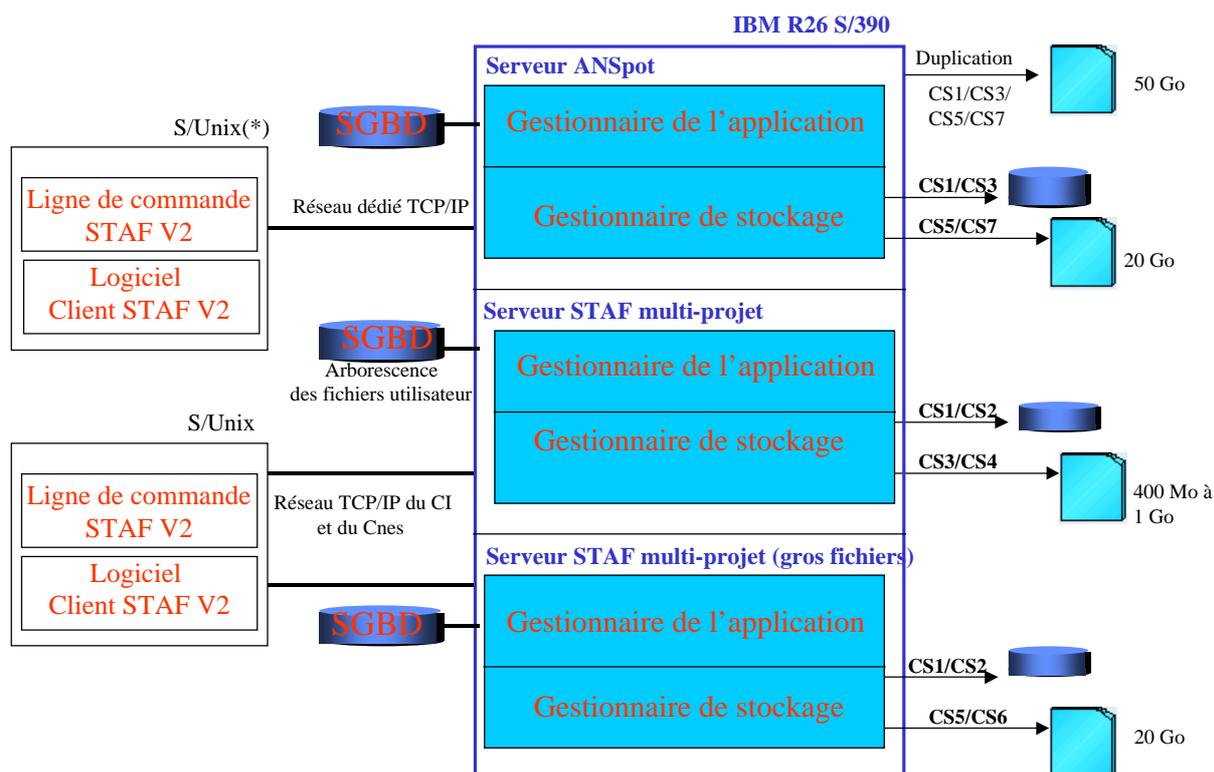
A partir de 2000, le service STAF atteint une activité en archivage de 100 Giga-octet par jour et accueille des projets d'archivage de plus en plus volumineux. La puissance du calculateur IBM est alors multiplié par 2,5 pour répondre au nombre croissant de projets. Enfin, courant 2000, le service offre un troisième serveur multi-projets réservé à l'archivage de données à forte volumétrie (fichiers de 50 Mo à 1 Go). Le stockage se base sur des médias capacitifs 9840 de StorageTek. Les bibliothèques sont alors partagées entre le serveur STAF de première génération et ce nouveau serveur, et présentent des lecteurs de gamme différente : 4490, 9490 et 9840 de StorageTek.

TAB. 2 Evolution du parc de médias

	1998	1999	2000	2001	2002
Cartouches STAF première génération	4490 9490	→	Recyclage des cartouches anciennes	→	→
Cartouches STAF dernière génération	SD3 (Redwood) 10Go	Migration SD3 10Go vers 9840 20Go	9840 20Go	→	→
Cartouches de sauvegarde	4490	Migration 4490 vers SD3 50Go	→	→	Migration en cours SD3 50Go vers 9940 60Go

En 2002, le service STAF se dote d'une nouvelle baie de disque IBM (ESS) de 500 Go pour temporiser les données récemment archivées et présentent de nouvelles interfaces réseaux (Ethernet Gigabit). Les volumes journaliers en archivage sont de l'ordre de 200 giga-octets par jour globalement sur les serveurs STAF de nouvelle génération. Le serveur STAF de première génération est maintenant sollicité en restitution de fichiers à raison de 2000 montages de cartouches par jour.

Ainsi, selon le profil de ses données, un projet qui souhaite archiver ses données sur les moyens du STAF, est orienté par l'équipe d'exploitation vers le serveur STAF le plus adéquat. Les serveurs STAF multi-projets sont accessibles depuis les moyens de traitement des utilisateurs et offrent l'image d'un service unique.



(*) Solaris 7 et 8 de SUN, Unicos de CRAY, Hp-Ux 11 de HP

FIG-2 : Configuration du service STAF – Juin 2002.

2.2. Evolution de la volumétrie stockée

Le service STAF accueille aujourd'hui une quarantaine d'applications opérationnelles, qui disposent d'une ou plusieurs arborescences de leur archive. Globalement, environ 2.4 millions de fichiers sont référencés pour un volume de 66 Téra-octets (hors copies) en juin 2002.

TAB. 3 : Evolution du stockage disque en volumétrie (Go) et en nombre de fichiers.

Stockage disque	Fin 1998	Fin 1999	Fin 2000	Fin 2001	Mi-2002
Volumétrie en Giga-octets	13	27	93	110	156
Nombre de fichiers	4950	79250	245965	417540	541980

TAB. 4 : Evolution du stockage sur cartouche de faible densité (de 400 Mo à 2Go) en volumétrie (Go) et en nombre de fichiers.

Stockage cartouche faible densité	Fin 1998	Fin 1999	Fin 2000	Fin 2001	Mi-2002
Volumétrie en giga-octets	2 125	1 950	2 320	2 340	2 700
Nombre de fichiers	967200	619760	1151396	1298900	1367280

TAB.5 : Evolution du stockage sur cartouche de forte densité (20Go) en volumétrie (Go) et en nombre de fichiers.

Stockage cartouche forte densité	Fin 1999	Fin 2000	Fin 2001	Mi-2002
Volumétrie en giga-octets	8 551	29 470	50 430	64 010
Nombre de fichiers	78395	255230	429540	482890

2.3. Critique du système actuel

2.3.1. Les avantages d'un système automatisé et centralisé

Le système technique du STAF présente un fort degré d'automatisation. Les fonctions offertes aux utilisateurs sont automatisées et ne nécessitent pas l'intervention d'un exploitant :

- le Centre Informatique a fait le choix de ne pas gérer de médias « *off-line* » pour faciliter les migrations technologiques des supports,
- les copies de sauvegarde sont directement accessibles par le système STAF sans intervention de l'exploitant dans le cas où la donnée originelle ne serait pas accessible,
- les opérations liées à la migration technologique des supports sont planifiées par l'exploitant mais reposent sur des mécanismes de relecture et de contrôle automatiques,
- le service est accessible 24H/24 et 7j/7 – en dehors des opérations de maintenance préventive (maintenance de la robotique, évolution des serveurs, des logiciels...) qui limitent la disponibilité du système à un taux de 98.5%.

Le service STAF bénéficie des avantages d'un système centralisé : en effet, les moyens de stockage sont des ressources partagées entre différents projets spatiaux et leur évolution profite à l'ensemble. La gestion des changements technologiques, qui touchent particulièrement de tels systèmes est prise en compte à un seul endroit par des équipes compétentes dont la mission est d'adapter le système existant à de nouvelles technologies de stockage. Le Centre Informatique recense régulièrement les besoins en archivage des projets utilisateurs pour dimensionner les moyens techniques du STAF et s'appuie sur les plans de développement de ses fournisseurs pour étudier les axes possibles d'évolution, que ce soit en terme de puissance des serveurs ou en terme de supports de stockage.

2.3.2. Evolutivité de l'architecture technique et ses limites

Malgré la montée en charge du service, l'architecture technique du STAF conçue il y a maintenant près de 10 ans, fournit des performances en archivage et restitution conformes aux besoins des utilisateurs.

Elle a su tirer profit des évolutions technologiques par ailleurs :

- l'évolution des débits du réseau de communication du 10 Mbit/s à 100 Mbit/s voire Gigabit/s, ce qui a permis d'augmenter significativement la bande passante en archivage et en restitution de données,
- l'évolution des temps d'écriture / lecture des lecteurs de médias, ce qui a permis de prendre en compte l'augmentation des volumes de données à stocker
- la diversité des choix de lecteurs de média adaptables à la robotique existante
- l'augmentation de puissance du calculateur IBM, ce qui a permis d'une part d'augmenter le nombre de sessions utilisateurs simultanées et d'autre part de traiter un flux de données en migration (HSM) et en sauvegarde de plus en plus volumineux.

Le système STAF, comme tout autre système d'archivage n'est pas performant pour des fichiers de très faible taille (< 2 Mo). En effet, pour ces fichiers, le temps de transfert reste négligeable comparé au traitement du fichier dans le système (contrôle d'intégrité et de droits d'accès, catalogage, montage de cartouche en restitution).

L'application STAF, sur laquelle repose le service, est très peu évolutive du fait de l'intégration de progiciel d'ancienne génération seulement disponible sur un environnement OS/390. Cependant, elle est restée compatible avec les montées de niveau successives du système OS/390 et de DB2. L'applicatif client, quant à lui, a été porté sur des systèmes d'exploitation différents depuis sa mise en service.

La gestion de la montée en charge du service a été facilitée par la multiplication des serveurs de traitement, qui adressent des besoins d'archivage différents :

- serveur d'archivage multi-projets pour des données de faible taille (<300 Mo)
- serveur d'archivage multi-projets pour des données de forte taille (<1 Go)
- serveur dédié à un projet.

La mise à disposition d'un nouveau serveur STAF nécessite aujourd'hui de dupliquer la configuration logicielle du serveur STAF sur une partition du ordinateur IBM, qui dispose de ses propres interfaces vers le réseau de communication et vers les moyens de stockage partagés (baie de disques et bibliothèques robotisées).

Enfin, le choix des médias et de leur densité a été fait en fonction du profil des données (longue durée de conservation et taille) pour assurer d'une part un nombre de relecture depuis le média le plus élevé possible et d'autre part, conserver une performance d'accès conforme aux classes de service.

Le « hiérarchiseur de média » DFHSM d'IBM ne fournit cependant pas la possibilité d'adresser différents médias au deuxième niveau de sa hiérarchie – ce qui limite la souplesse de l'architecture actuelle.

2.3.3. Fiabilité du service et respect de l'intégrité des données

STAF est un système d'archivage fiable, qui depuis sa mise en exploitation garantit l'intégrité des données stockées : la fonction d'archivage assure dans un premier temps la complétude et l'intégrité de la donnée à conserver, puis, une fois stockée, la donnée est régulièrement relue et déplacée sur des supports différents. Hormis dans le cas des restitutions par les utilisateurs, la relecture des données a lieu dans les cas suivants :

- la phase de dé-fragmentation des médias : le hiérarchiseur de média ne permet une relecture des données au moment de la restitution que par déplacement de la donnée du deuxième niveau (cartouche near-line) vers le premier niveau (disque) – ce qui entraîne la fragmentation des cartouches.
- le recyclage des cartouches tous les 5 ans
- la migration technologique des supports lorsque ces supports deviennent obsolètes
- les audits sur l'espace de stockage.

Pour se parer des éventuels accidents matériels et logiciels, les disques magnétiques sont de technologie RAID 1 (disque miroir) et RAID 5 (contrôle d'erreur réparti sur un agrégat de disques).

Pour l'ensemble des médias, l'utilisateur signale au moment de l'archivage la duplication ou pas de la donnée stockée (via les classes de service). Près de 100% des données stockées sur des médias *near-line* sont dupliquées. Les versions de copies sont accessibles automatiquement depuis une bibliothèque robotisée et dédiée. Cette bibliothèque est localisée dans une salle informatique différente de celle où se trouvent les bibliothèques contenant les versions originelles.

3. Projet de rénovation du STAF

3.1. Objectifs

Le projet de rénovation du STAF a comme objectifs principaux de :

- rénover certains modules logiciels qui sont devenus obsolètes
- favoriser l'ouverture de l'architecture technique vers des gestionnaires de stockage de nouvelle génération
- dimensionner le système à des besoins en performance d'archivage trois fois plus élevés.

tout en conservant la compatibilité ascendante de l'interface utilisateur et en intégrant le gestionnaire de stockage existant sous IBM OS/390.

Ce projet donne lieu à un appel d'offres auprès des intégrateurs en septembre 2002 – le développement et la qualification sont envisagés sur l'année 2003 et premier semestre 2004. Le CNES vise une mise en service sur le deuxième semestre 2004.

3.2. Les grands principes d'évolution de l'architecture technique

Globalement, la future architecture technique doit présenter une meilleure modularité et flexibilité pour prendre en compte rapidement les montées en charge du service sur les prochaines années.

Pour cela, on vise une dissociation complète entre le gestionnaire de l'application et le gestionnaire de stockage. Le gestionnaire de l'application reposera sur une ou plusieurs plates-formes Unix, qui communiqueront avec un ou plusieurs gestionnaires de stockage.

L'optimisation des fonctions d'archivage et de restitution passe par la dissociation entre les flux de données, les flux de contrôle et les flux de requêtes client serveur. Chacun des flux reposera sur des protocoles de communication adaptés et standards :

- Le flux de données (transfert de fichiers), conformément au modèle de référence d'un système de stockage de masse ^[1], sera direct entre l'émetteur (source des données) et le récepteur (destinataire des données).
- Le contrôle du transfert sera assuré par le gestionnaire de l'application en tant que tierce partie (conformément au modèle de référence OSSI ^[2]).
- Le dialogue client serveur sera assuré par un logiciel (middleware) standardisé et évolutif.

L'ouverture possible du système à de nouveaux gestionnaires de stockage est souhaitable pour d'une part bénéficier de l'apport des nouvelles technologies dans le domaine et d'autre part, se parer contre l'obsolescence éventuelle des matériels et logiciels en place.

L'objectif est de prévoir dans le nouveau système les mécanismes :

- qui permettront d'intégrer à moindre coût un nouveau gestionnaire de stockage (logiciels et matériels standards du marché),
- qui faciliteront la migration des données d'un ancien gestionnaire vers un nouveau.

Dans un premier temps, le CNES vise à intégrer les gestionnaires de stockage actuels (partitions OS/390) pour éviter de déplacer l'ensemble des fichiers stockés depuis 1995.

3.3. Quelques informations de dimension du futur système

D'après les études de besoins, le futur système devrait délivrer un débit en archivage de l'ordre de 70 Go/heure et de 25 Go/heure en restitution. Le projet vise 8 millions de fichiers référencés à l'horizon 2006 pour une progression du volume stocké de l'ordre de 600 Giga-octets par jour.

4. Conclusion

Le service STAF fait maintenant partie intégrante des applications spatiales pour la préservation physique des données et assure son rôle de conservatoire. Le système technique est en perpétuelle évolution pour s'adapter aux nouveaux besoins de stockage et parer à l'obsolescence des logiciels ou des matériels qui le composent. Au travers du projet de rénovation, le CNES confirme sa volonté de moderniser le service d'archivage et garantir sur le long terme la continuité opérationnelle de STAF.

Bibliographie :

[1] Mass Storage System Reference Model Version 4 (May, 1990) – IEEE – Edité par Sam Coleman, Steve Miller.

[2] Open Storage Systems Interconnexion (OSSI) Reference Model – Version 5 – IEEE P1244
http://www.ssswg.org/public_documents/MSSRM/V5.html