

CAS CLIENT

L'OBSERVATOIRE DE PARIS

STOCKAGE ET ARCHIVAGE DE PHOTOS DU CIEL EN ACCÈS LIBRE POUR LA COMMUNAUTÉ SCIENTIFIQUE



Observatoire de Paris - Site de Meudon © Observatoire de Paris

L'Observatoire de Paris a choisi Active Circle pour mettre en place une infrastructure de stockage qui augmente le niveau de sécurité des données et qui permet de gérer le stockage comme un service.



Exemple d'image stockée : galaxies de l'hémisphère sud (dite du sombrero) – crédit CAI-Gepi-VOParis

L'OBSERVATOIRE DE PARIS

Fondé en 1667, l'Observatoire de Paris est une des plus grandes institutions mondiales de recherche en astronomie et astrophysique et le plus important centre de recherche européen dans ce domaine. Il est installé sur trois sites : Paris, Meudon (Hauts-de-Seine) et Nançay (en Sologne). Les activités de recherche sont regroupées au sein de cinq départements, d'une unité scientifique (Nançay) et d'un Institut, des entités formellement associées au Centre national de la recherche scientifique. L'Observatoire est également tutelle du laboratoire APC, hébergé à l'Université Paris-Denis Diderot.

UNE REFONTE DU STOCKAGE

En octobre 2007 un appel d'offres est publié par la DIO (Division Informatique de l'Observatoire) pour effectuer une refonte du stockage scientifique. La DIO est un service commun qui a pour rôle de fournir des services informatiques à l'ensemble des usagers de l'établissement.

« Nous voulions doter la DIO d'une solution de stockage ayant pour vocation de remplacer de multiples baies saturées et sujettes à des pannes. De plus nous cherchions un système de stockage permettant de faire face à de fortes évolutions de capacité, d'un rapport prix/capacité compatible avec l'usage des données stockées. Enfin il devait être sécurisé, pour pallier tous les risques possibles de perte de données. » précise Joël Marchand, directeur de la DIO.

Ce projet comporte un caractère pilote pour une solution de stockage novatrice qui pourrait être étendue à l'hébergement de données des laboratoires, si ceux-ci sont convaincus par les qualités de la solution retenue.

DES DONNÉES SCIENTIFIQUES ET IMAGES

Les données concernées par ce projet sont exclusivement des données scientifiques, qui sont le produit d'observations (images, spectres et données 3D), ou de simulations numériques. Le périmètre du projet comprend le stockage de « l'Observatoire Virtuel », qui a pour objectif de stocker les images du ciel et de les mettre à disposition de la communauté scientifique.

Les caractéristiques des données sont les suivantes : grands volumes (plusieurs dizaines de Teraoctets), compression native très forte, très grands systèmes de fichiers comprenant des millions de fichiers, une conservation des données sur une durée très longue, et enfin une valeur importante des données car elles sont issues d'un long travail de recherche ou d'observations uniques. L'accès devra s'effectuer via le protocole NFS.

LE CHOIX

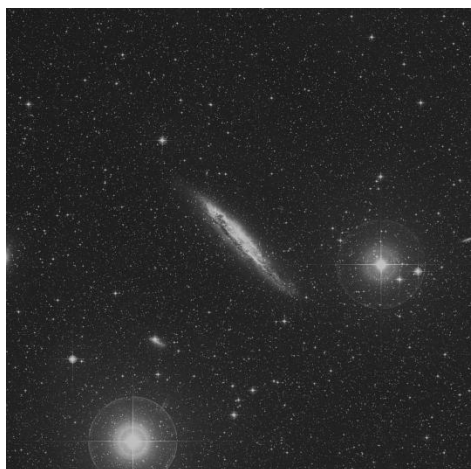
Active Circle est retenu au terme d'un appel d'offre public, et donc d'une compétition très large. Parmi les critères déterminants, la solution devait pouvoir évoluer sur le plan matériel sans remise en cause de l'architecture. Le stockage à deux niveaux avec HSM sur disque et bande a été décisif, ainsi que la possibilité de répliquer les données entre les sites tout en gardant une gestion et un système de fichier global. Enfin l'Observatoire de Paris a apprécié le format TAR pour l'archivage et le modèle tarifaire compétitif adapté à un fonctionnement en mode service pour fournir du stockage sécurisé aux laboratoires.

« Nous cherchions un système de stockage permettant de faire face à de fortes évolutions de capacité »
 « Il devait être sécurisé, pour pallier à tous les risques possibles de perte de données. »

Joël Marchand – Directeur de la division informatique

« Active Circle nous permet de gérer le stockage en mode service. Nous mutualisons les ressources et pouvons proposer un stockage sécurisé aux différents laboratoires de l'Observatoire. »

Joël Marchand – Directeur de la division informatique



Exemple d'image stockée : galaxies de l'hémisphère sud (dite du sombréro) - crédit CAI-Gepi-VOParis

UNE INFRASTRUCTURE LOGICIELLE SUR MATÉRIEL STANDARD

La solution choisie est constituée de la plate-forme logicielle Active Circle qui fournit une virtualisation des systèmes de fichiers qui inclut les fonctionnalités de gestion des données : réplication, protection continue, HSM, et archivage. Les serveurs sont au nombre de deux : un SunFire X4500 sous Solaris d'une capacité disque de 36 To bruts, choisi pour son rapport capacité/cout et le support du système de fichiers ZFS ; un serveur HP Proliant sous Linux qui dispose de 5 To de stockage disque et qui pilote deux bibliothèques HP et Overland d'une capacité totale de 100 To avec 3 lecteurs LTO4.

UN STOCKAGE SÉCURISÉ ET HIÉRARCHISÉ

Ces espaces de stockage disques et bandes sont organisés en entrepôts. Les deux serveurs sont situés physiquement en des endroits différents sur le campus de l'Observatoire, et les données sont répliquées pour une protection contre les sinistres et les pannes matérielles. Le versioning protège contre les erreurs. De plus, pour certains systèmes de fichiers, les données les plus anciennes sont automatiquement migrées sur bande LTO, et restent en ligne en cas d'accès. « Nous apprécions particulièrement la possibilité de mixer le disque et la bande » déclare Joël Marchand.

UNE GESTION AUTOMATISÉE DU CYCLE DE VIE DES DONNÉES

Les données sont organisées en systèmes de fichiers en fonction de leur nature. Toute la gestion du cycle de vie des données - localisation, niveau de redondance, déplacement vers la bande - est automatisée selon une règle appelée « classe de service » qui est définie pour chaque système de fichiers. « Active Circle est une bonne solution pour faire de l'archivage longue durée sur des gros volumes : le travail s'effectue automatiquement en tâche de fond, avec reprise automatique en cas d'erreur ou de coupure » souligne Joël Marchand.

ARCHIVAGE AU FORMAT TAR

Régulièrement un export des systèmes de fichiers est effectué sur un pool de cartouches LTO pour mise au coffre. Cette opération constitue un moyen supplémentaire de protéger ce patrimoine scientifique. Active Circle propose de manière native cet export au format tar, pour garantir la relecture des données sur n'importe quel système, même hors de l'application Active Circle.

BILAN ET EXTENSIONS

« Après une phase de mise au point en lien étroit avec l'éditeur, qui a surtout travaillé la scalabilité en dizaines de millions de fichiers, le service est considéré comme mature » déclare Joël Marchand, directeur de la DIO. « Active Circle nous permet de gérer le stockage en mode service. Nous mutualisons les ressources et pouvons proposer un stockage sécurisé aux différents laboratoires de l'Observatoire. Ce service se présente aux utilisateurs sous la forme de services NAS (CIFS/SMB, NFS, FTP) classiques accessibles à tout type de clients (Windows, MacOS, Linux/Unix), qui n'ont besoin d'aucun logiciel pour utiliser un partage réseau proposé par la solution. »

Créée en 2002, la société Active Circle est un éditeur de logiciel spécialisé dans le stockage et l'archivage numérique. Active Circle s'adresse aux organisations qui gèrent de grands volumes de contenu vidéo, images, données scientifiques et techniques, données utilisateurs, et leur permet de maîtriser le cycle de vie des données tout en optimisant les coûts de possession et d'administration du stockage.



Active Circle est une marque déposée par Active Circle SA. Les autres noms ou représentations sont mentionnées à des fins d'identification et appartiennent à leurs propriétaires respectifs.

© ACTIVE CIRCLE 2009 - Ce document ne peut être copié ou reproduit sans accord écrit.
www.active-circle.com