

● MIRIA CUSTOMER SUCCESS • RECHERCHE AVANCÉE / HPC

DiRAC

DiRAC déploie la solution Atempo Miria Archivage

Les données jouent un rôle primordial dans la mission du DiRAC. Les besoins en puissance de calcul (mesurés en pétaflops) et en stockage de données (en pétaoctets) sont au cœur des projets soutenus par le DiRAC.

« Miria Archivage est extrêmement puissant et riche en fonctionnalités, Miria Archivage devrait répondre à nos besoins futurs. Nous avons été très impressionnés par Miria, même si nous n'avons entrevu qu'une petite partie de ses possibilités. »

— DR ALASTAIR BASDEN • RESPONSABLE TECHNIQUE DU CENTRE MEMORY INTENSIVE

L'ORGANISATION

DiRAC (Distributed Research Utilizing Advanced Computing) est le centre de calcul intensif intégré pour la modélisation théorique et la recherche basée sur le calcul haute performance (HPC) en physique des particules, astronomie et cosmologie. Ce centre est un dispositif clé du Science and Technology Facilities Council (STFC) du Royaume-Uni. Les chercheurs du monde entier peuvent accéder aux puissantes installations informatiques du DiRAC.

L'équipe de Durham a eu pour mission d'assurer la transition des données du stockage Spectrum Scale (GPFS) vers le système Lustre de DDN. L'objectif était, selon Dr. Lydia Heck « d'augmenter efficacement la capacité de stockage en utilisant une solution moins coûteuse ». Le besoin croissant de mémoire de calcul génère en effet d'importants volumes de données dans les environnements HPC.

Les besoins en calcul haute performance des centres de HPC génèrent en effet d'importants volumes de données et des besoins accrus en matière de stockage et de hiérarchisation de données. Ce stockage des données ayant atteint 20 Po en 2022 avec le déploiement de la phase 3 de financement du DiRAC. Le plan de gestion des données du DiRAC comprend l'archivage des bases de données de recherche et des documents de recherche scientifique finalisés, révisés par des pairs.

À PROPOS DU DIRAC ET D'ATEMPO

Le centre « Memory Intensive » du DiRAC situé sur le campus de l'université de Durham a récemment fait appel à Atempo, le spécialiste de la protection et du transfert de données, et à son partenaire britannique OCF pour l'archivage de plusieurs pétaoctets de données de recherche stockées sur ses systèmes Lustre et Spectrum Scale (GPFS).

Le projet DiRAC Memory Intensive, septième cluster HPC à Durham, permet de fournir aux chercheurs une puissance de calcul de 452 nœuds et 12 656 cœurs.

Nous avons rencontré le Dr. Lydia Heck, Directrice Technique et le Dr. Alastair Basden, Responsable Technique du DiRAC qui nous ont donné leur avis sur la solution Miria Archivage d'Atempo, et comment elle répond à leurs exigences d'archivage des projets de recherche sur bande.

LES ARCHITECTURES HPC DU DIRAC

- Quatre universités britanniques, Cambridge, Durham, Edimbourg et Leicester, assurent la gestion des architectures HPC du DiRAC.
- Ces universités fournissent les installations de calcul haute performance principales, ainsi que l'expertise nécessaire à la réalisation de multiples projets de recherche, de soutien, de transfert de connaissances, et de partenariat industriel.

LA SOLUTION EN PRATIQUE : MIRIA ARCHIVAGE

Du stockage principal vers quatre lecteurs de bandes LTO

Le rôle de Miria est d'archiver les données de recherche et libérer de l'espace de stockage disque coûteux en transférant les données de recherche du stockage principal vers quatre lecteurs de bandes LTO. Les besoins actuels de stockage sont de 6,5 Po et passeront à plus de 20 Po avec la phase 3 du DiRAC HPC. Ceci multipliera par 10 la vitesse de traitement et de création de données, et par conséquent des besoins en stockage.

La solution d'archivage en place était lente et pas suffisamment évolutive. Atempo a proposé fin 2018 une démonstration de faisabilité d'abord sur Spectrum Scale (GPFS) puis sur Lustre. Atempo a rapidement démontré que Miria Archivage était idéal pour répondre aux besoins d'archivage du DiRAC. Un serveur d'archivage Miria et son Data Mover dédié accèdent au système de fichiers Lustre et gèrent efficacement les flux d'archivage de données.

L'interface utilisateur de Miria Archivage permet aux utilisateurs d'effectuer leurs propres tâches d'archivage rapidement et de restaurer directement leurs données. Les chercheurs sont libres d'organiser l'arborescence de chaque projet de recherche lors de l'archivage de leurs données. Les équipes du DiRAC utilisent Miria à la fois pour sauvegarder les répertoires personnels des utilisateurs et archiver des fichiers sélectionnés par ces derniers.

« Chaque action de l'administrateur et de l'utilisateur est gérée par le protocole HTTP, ce qui signifie que nous pouvons très facilement mettre en place un tunnel SSH et permettre aux utilisateurs d'archiver leurs fichiers où qu'ils soient dans le monde. »

« Les performances d'archivage sur les flux de données des systèmes de fichiers Lustre sont optimales, ce qui nous satisfait pleinement. »

— DR ALASTAIR BASDEN · DIRAC

Les bonnes pratiques d'archivage imposent la création de deux copies distinctes sur bande, LTO7 et LTO8 dans le cas du DiRAC, ce qui est parfaitement supporté par Miria.

LES PERSPECTIVES

Miria n'est pas seulement un outil d'archivage. Le composant Miria Sauvegarde assure également la protection des données des utilisateurs. Atempo a aussi démontré les capacités de son option de FastScan pour GPFS. FastScan optimise la détection rapide des fichiers ajoutés et modifiés sur un stockage et permet d'éviter de multiples et fastidieux balayages du système incompatibles avec ce type de stockage massifs. DiRAC dispose ainsi de la sauvegarde avancée offerte par Miria avec son option FastScan pour la protection de son système Lustre de DDN.

Le DiRAC possède plusieurs sites à connexion haut débit - entre les universités membres (bande passante pouvant atteindre 20 GO/s) et prévoit de pouvoir transférer, protéger et stocker les données en dehors de ses sites. L'organisme envisage par ailleurs la possibilité de faire une copie secondaire des données d'un site et de les stocker sur un autre site.

CONCLUSION

En conclusion, la solution Miria for Archiving a produit des résultats très positifs en matière de performance, d'efficacité et de sécurité. Les équipes d'Atempo et du DiRAC continueront à travailler dans un esprit de coopération et de mutualisation des ressources, afin de confirmer ce premier succès.

LES BÉNÉFICES MÉTIER

● Réduction des coûts de stockage	● Archivage & sauvegarde rapide & haute capacité	● Évolutivité adaptée aux besoins
---	--	---

Plateforme unique de data management

4 SERVICES DE DONNÉES · 1 PLATEFORME

● Sauvegarde Protéger rapidement les stockages de fichiers et garantir la continuité d'accès aux données	● Migration Migrer efficacement volumes et milliards de fichiers entre stockages hétérogènes	● Archivage Libérer l'espace des stockages primaires et assurer la pérennité des données	● Mobilité Déplacer et synchroniser les jeux de données entre stockages et sites
---	---	---	---

CONTACTER UN EXPERT ATEMPO
links.atempo.com/DM-contact

CONSULTER LA FICHE PRODUIT
links.atempo.com/DM-presentation