

RAPPORT D'ACTIVITÉ

■ DCI ■ DS2I ■ DAD ■ DARH ■ FORMATION



L'éditorial



Francis DAUMAS

Directeur

Un premier coup d'œil dans le rétroviseur nous rappelle l'installation en 2008 de la première tranche du supercalculateur JADE, modèle ICE8200 de la société SGI, acquis par GENCI. C'est grâce à la qualité des infrastructures disponibles et de son expérience dans le calcul intensif que le CINES a pu offrir cette opportunité. JADE a été classée 12^{ième} mondiale en terme de performance et a été la première étape d'une progression spectaculaire des moyens de calcul nationaux contribuant au retour de la France dans le concert européen du calcul de haute performance.

A ce stade le CINES a mis à la disposition de GENCI et des autres centres nationaux l'application DARI en ligne qui permet encore aujourd'hui de gérer les appels à projets, les évaluations des experts et le suivi statistique des consommations.

L'importance primordiale de disposer d'infrastructures performantes a depuis conduit le CINES à mettre en service en 2010 un nouveau « pôle énergie » et à lancer en 2012 l'étude de la construction d'une nouvelle salle machine de 600 m². Celle-ci permettra dès 2014 d'accueillir les futures générations de supercalculateurs. Ces travaux s'accompagneront de l'installation d'équipements électriques et de climatisation performants et limitant l'empreinte environnementale par la prise en compte des technologies de refroidissement des processeurs par de l'eau tiède et du « free cooling ».

La construction du nouveau « pôle énergie » a permis en 2010 la jouvence de JADE, à mi-parcours de son existence. Elle affiche maintenant une puissance de 267 Tflops, confirmant ainsi la volonté de faire évoluer une configuration qui donne entière satisfaction aux chercheurs. L'environnement de calcul a également été complété par des investissements du CINES sur les espaces de stockage (disques et robotique) et les machines de pré et post-traitement : machine CRISTAL constituée de nœuds de type BullX équipés de GPU avec des fonctionnalités de visualisation à distance. Parallèlement des collaborations ont été mises en œuvre sur le « green computing », notamment avec IBM. Enfin depuis 2011 un lien Renater à 10 Gigabits/s dédié relie le CINES aux deux autres sites nationaux et au-delà au réseau des supercalculateurs des autres pays de l'Union Européenne.

L'ouverture à l'Europe à travers les contributions aux projets PRACE, HPC Europe2 et Eudat, est aussi un des faits essentiels de ces dernières années. Le centre assure le développement de l'application « Prace Peer Review », contribue au tremplin de codes stratégiques vers plus de parallélisme, travaille sur différents prototypes technologiques européens, accueille des chercheurs de l'Union Européenne et prépare l'arrivée prochaine des grandes masses de données. Dans ce domaine l'expertise du CINES sur l'archivage pérenne des données numériques est un atout important.

Grâce aux très bonnes performances de JADE et à l'implication des personnels du CINES, de nombreux scientifiques ont pu réaliser des avancées spectaculaires dans leurs disciplines. Dans cette aventure les comités thématiques, et particulièrement leurs présidents, et le comité des utilisateurs du CINES (C4) ont joué un rôle prépondérant.

Concernant la conservation à long terme des données numériques, à côté des réalisations opérationnelles d'archives numériques (Thèses, revues PERSEE, ...), de nouveaux projets se sont manifestés (l'INSERM, l'IRHT, l'Université de Lorraine, ...) dont certains adressent de très gros volumes de données. Quelques points méritent d'être mis en exergue.

Le CINES a obtenu pour sa plate-forme d'archivage (PAC) l'agrément du Service Interministériel des Archives de France (SIAF) pour les archives intermédiaires et participe à la mise en place de la norme ISO 16363. Ainsi les communautés scientifiques peuvent trouver au CINES une vraie solution, respectant les normes et standards, à leur problème de conservation des données stratégiques (documentaires ou issues d'observations ou de calculs). Le Centre est en particulier partenaire de l'Equipex GEOSUD (« GEOInformation for Sustainable Development », piloté par le Cemagref, pour l'archivage d'images satellites à haute définition. Il est utile de rappeler ici l'obligation faite avant toute publication de prouver que les données utilisées ont bien été archivées.

Le CINES co-pilote le groupe de travail sur l'archivage pérenne dans le cadre du projet national d'infrastructure BSN (Bibliothèque Scientifique Numérique) mis en place par le Ministère.

Le CINES est le nœud français de Eudat (« European Datafore-science »), projet européen de 3 ans démarré le 1 octobre 2011, regroupant 23 partenaires dont l'objectif est de mettre en place en Europe une infrastructure de données collaborative (CDI) inter-disciplinaire pour le partage, la duplication et la conservation à long terme des grandes masses de données. Des services (« data services ») seront développés et mis en place sur les nœuds du réseau Eudat à destination des communautés d'utilisateurs et en prenant en compte leurs besoins. Les communautés déjà structurées, CLARIN (linguistique), EPOS (science de la terre), LIFEWATCH (environnement), VPH (santé) et ENES (climat) sont actuellement associés à la démarche.

En tant que nœud français de Eudat, le CINES dispose de capacités de stockage, de réplication à distance et de conservation à long terme de grandes masses de données pour les communautés qui le souhaitent. Les communautés scientifiques nationales sont invitées à communiquer leurs besoins spécifiques auprès du CINES afin que ceux-ci soient pris en compte dans la réflexion qui s'est engagée.

Au-delà des moyens matériels et logiciels mis en place, l'implication des personnels du Centre permet d'assurer la meilleure exploitation possible et un support de qualité aux utilisateurs. Les nombreuses réalisations et projets en cours témoignent de l'expertise unanimement reconnue des collaborateurs du CINES.

DCI

LE DÉPARTEMENT CALCUL INTENSIF (DCI) 6

- a Les missions 8
- b L'équipe 8

LES CLUSTERS DE CALCUL 9

- a Les matériels 9
- b Les logiciels et outils systèmes 10
- c Les logiciels applicatifs 11

LES ACTIONS 12

- a La procédure DARI de demandes d'heures 12
- b Le support aux utilisateurs 13
- c La visualisation à distance 14
- d Le pré/post traitement 18
- e Le C4 18

LE CINES ACTEUR DU HPC 20

- a Au plan national 20
- b En Europe 21
- c Au niveau régional 25

3 ANS DE CALCUL AU CINES RÉSUMÉS EN 7 GRAPHIQUES 26

- a L'activité de la machine Jade sur 36 mois 26
- b L'activité de la machine Yoda sur 36 mois 26
- c Le nombre de projets retenus par CT et par an 27
- d La consommation de chaque CT par an 27
- e L'origine géographique des projets retenus 28

DS2I

LE DÉPARTEMENT 32

- a Les missions 32
- b Structure 32
- c Moyens humains 32
- d Les formations 32

ADMINISTRATION DES SYSTÈMES ET DES DONNÉES 33

- a Le périmètre 33
- b Les serveurs d'archivage 33
- c Les données scientifiques 33
- d Convergence : partage mobilité, sécurisation et archivage des données scientifiques 34

LES SERVICES RÉSEAUX ET LA SÉCURITÉ 35

- a Le réseau évolue 35
- b Les services 35
- c La sécurité 35

L'EXPLOITATION 36

- a La suspension et le suivi des services 36
- b Le centre des services : SVP 36
- c L'astreinte 37
- d ITIL 38
- f L'hébergement pour des tiers 39
- g L'activité en chiffres 39

BILAN ET PERSPECTIVES 39

DAD

LES SERVICES D'ARCHIVAGE PERENNE DU CINES 42

LES NOUVEAUX UTILISATEURS DU SERVICE PAC 44

- a Archive ouverte HAL 44
- b TGE Adonis (projet pilote avec le CRDO Paris et le CRDO Aix) 44
- c Bibliothèque de l'université Pierre et Marie Curie (BUPMC) 44
- d Bibliothèque CUJAS 44
- e École française d'Extrême Orient (EFEO) 44
- f Bibliothèque Interuniversitaire de Santé 45
- g Sainte-Geneviève (BSG) 45

LES CHIFFRES CLÉS 45

- a Évolution de la volumétrie 45
- b Répartition des objets archivés en 2010 et 2011 47
- c Facile_ outil de validation du format d'archivage du CINES par analyse et expertise 48

LA DÉMARCHE QUALITÉ 50

- a En route vers la certification 50
- b Formalisation de l'activité et amélioration de la qualité de la documentation 51
- c En route vers la certification 52

PARTAGE DE L'EXPERIENCE ET DE L'EXPERTISE DU CINES AVEC LES COMMUNAUTÉS (ENSEIGNEMENT SUPÉRIEUR, RECHERCHE, ARCHIVISTIQUE, SERVICES PUBLICS, ETC.) 53

- a Mise à disposition de la documentation 53
- b Participation à divers groupes de réflexion et de travail 54
- c Publications 54
- d Formations 54

PERSPECTIVES D'ÉVOLUTIONS 55

- a L'archivage intermédiaire des données scientifiques 55
- b L'implication dans les projets européens 55

DARH

LE DÉPARTEMENT 58

- a Ses missions 58
- b Ressources humaines 58
- c La gestion financière, le contrôle de gestion 58
- d Les marchés 58

ORGANISATION DU CINES 59

- a Le conseil d'administration 59
- b Le comité hygiène et sécurité 59
- c Le conseil du centre 59

LA FORMATION

- a Les formations dites CIREN 60
- b Les formations de type «expert en archivage» 61
- c Les formations calcul 61
- d Département Services Informatiques et Infrastructures 62

LE DÉPARTEMENT CALCUL INTENSIF (DCI) 6

Les missions 8
L'équipe 8

LES CLUSTERS DE CALCUL 9

Les matériels 9
Les logiciels et outils systèmes 10
Les logiciels applicatifs 11

LES ACTIONS 12

La procédure DARI de demandes... 12
Le support aux utilisateurs 13
La visualisation à distance 14
Le pré/post traitement 18
Le C4 18

LE CINES ACTEUR DU HPC 20

Au plan national 20
En Europe 21
Au niveau régional 25

3 ANS DE CALCUL AU CINES RÉSUMÉS EN 7 GRAPHIQUES 26

L'activité de la machine Jade sur ... 26
L'activité de la machine Yoda sur...26
Le nombre de projets retenus... 27
La consommation de chaque CT par an 27
L'origine géographique des projets... 28

Système de données # Exploitation # Réseau, hébergement et infrastructure

Le département calcul intensif (DCI)

a Les missions

La création du Département Calcul Intensif en 2008 avait pour premier objectif de regrouper au sein d'une même entité l'ensemble des forces concourant au développement de la mission historique du CINES : le calcul. Ce regroupement permettait également de donner, à travers l'organigramme du Centre, une visibilité immédiate et concrète à la mission calcul intensif du CINES. Le DCI est en charge des environnements de calcul hébergés au CINES, de l'aide aux utilisateurs pour la parallélisation et l'optimisation des codes, et de la promotion des techniques de calcul parallèle. Il fournit l'expertise informatique aux

nombreuses disciplines scientifiques qui utilisent les supercalculateurs du Centre pour la résolution de problèmes qui exigent des puissances de calcul extrêmes associées à de grandes capacités mémoire.

b L'équipe

A la fin de l'année 2012, le DCI regroupe quinze ingénieurs – onze titulaires et quatre contractuels - dont l'activité se décline principalement selon trois axes : architectures et systèmes, support aux utilisateurs, projets et collaborations internationales.



L'équipe du DCI devant la machine JADE

Les clusters de calcul

a Les matériels

Le CINES met à la disposition de ses utilisateurs trois supercalculateurs à architecture parallèle :

- ✚ Jade (constructeur SGI)
- ✚ Yoda (constructeur IBM)
- ✚ Cristal (constructeur BULL)

La machine Jade



Date d'installation : Avril 2008 (tranche 1) et mars 2010 (tranche 2)
Constructeur : SGI
Type : ALTIX ICE
Puissance crête : 267 T.flop/s
Processeurs : INTEL Quad-Core E5472 (tranche 1) et E5560 (tranche 2)
Nombre de cœurs / nœuds : 23 040 cœurs / 2 880 nœuds
Mémoire par nœud : 30 ou 34 Go
Coprocesseurs graphiques : aucun
Système d'exploitation : Suse SLES 11 SP2
Réseau d'interconnexion : QDR 40 Gbits/s
Système de fichiers : Lustre 2.1.4
Espace-disque fichiers temporaires : 700 To (12 armoires)

Le cluster Jade est un supercalculateur scalaire parallèle constitué de 46 armoires dont 45 sont dédiées au calcul. Une armoire frontale de connexion assure l'accès à l'ensemble des 2 880 nœuds de calcul. Le refroidissement est assuré par un système à porte froide, haut rendement air/eau.

La machine Yoda



Date d'installation : Mars 2012
Constructeur : IBM
Type : IBM P1600 / Power 7
Puissance crête : 6.6 T.flop/s
Processeurs : Power7 P755 (3.3 GHz)
Nombre de cœurs / nœuds : 256 cœurs / 8 nœuds
Mémoire par nœud : 128 Go
Coprocesseurs graphiques : aucun
Système d'exploitation : Linux RedHat 6
Réseau d'interconnexion : Infiniband DDR
Système de fichiers : GPFS
Espace-disque fichiers temporaires : 10 To

La machine Cristal



Date d'installation : Octobre 2011
Constructeur : BULL
Type : BULL X
Puissance crête : 0.6 T.flop/s
Processeurs : INTEL ES-2660 et ES-4260
Nombre de cœurs / nœuds : 16 cœurs / 9 nœuds
Mémoire par nœud : 128 Go et 1 To
Coprocesseurs graphiques : 4 GPUs Quadro FX5800 de 4 Go (NVIDIA)
Système d'exploitation : Centos 6.3
Réseau d'interconnexion : Infiniband FDR (56 Gb/s)
Système de fichiers : NFS plus Lustre Jade
Espace-disque fichiers temporaires : 20 To

La machine Cristal est une machine spécialisée dans les travaux de pré ou post-traitement et de visualisation à distance.

b Les logiciels et outils systèmes

Les trois principaux clusters du CINES sont administrés avec des outils logiciel différents. Tout d'abord, quelques indications techniques sur ces trois clusters. Les nœuds sont imagés par **Xcat** (IBM) pour Yoda, **Tempo** (SGI) via **SystemImager** pour Jade et **Cobbler** pour Cristal. Les commandes Linux d'administration sont exécutées sur les nœuds de calculs via leur BMC avec **Ipmitool** ou **Cpower** pour Jade ou, plus classiquement, par **pdsh**. D'un point de vue réseau haute performance, l'**OFED** et l'**OpenSM** gèrent les fabrics infiniband de Yoda, Jade et Cristal.

Les informations relatives aux utilisateurs et à leurs projets sont gérées par un serveur **LDAP** sécurisé. Il est interrogé par Yoda, Jade et Cristal. En particulier, Jade procède par extraction de la base LDAP et propagation d'un fichier passwd dans les nœuds de calcul.

Les environnement SHELL utilisateurs sont simplifiés et positionnés de façon cohérente grâce à l'outil **module**.

Les espaces disques attachés sont de type **GPFS** pour Yoda, **Lustre** pour Jade avec des quotas non pas coercitifs sur l'écriture de fichiers mais sur la soumission de jobs. **RobinHood** enregistre régulièrement les métadonnées Lustre dans une base **MySQL** à des fins statistiques et de définition puis d'application d'une politique de gestion de données.

RRDtool couplé à **Drraw** fournit une interface web interactive produisant des graphiques statistiques et/ou temporels sur l'activité Lustre.

Les trois clusters sont contrôlés par un outil de monitoring **Nagios** central. Il est hiérarchique sur Jade. De plus, un système de remontée d'alertes envoie des courriels à svp@cines.fr. Ainsi, les alertes sont enregistrées automatiquement par l'outil de gestion d'incidents **OTRS**.

Dans le cadre de PRACE, une infrastructure web **INCA** assure la surveillance fonctionnelle et des mesures de performances. Les accès (10Gbe) et les transferts vers les nœuds PRACE des autres centres de calculs sont réalisés avec **GSI-SSH**, **Grid FTP** et monitorés avec l'outil **Iperf**.

Ganglia enregistre des métriques mémoire, charge CPU et fournit une interface web interactive produisant des graphiques temporels.

Htop affiche en temps réel et de façon synthétique l'origine des activités I/O, réseau et CPU d'un nœud.

Les travaux utilisateurs sont ordonnancés par **LoadLeveler** et isolés à l'intérieur d'un nœud avec la fonctionnalité **cgroup** sur Yoda. Pour Jade, **PBS** est utilisé avec l'application d'une politique d'ordonnancement favorisant les travaux des projets sous-linéaires par rapport à leur consommation annuelle. Le contrôle des ressources est assuré par **Arrayd** (SGI) et un épilogue tue les processus inutiles et libère

la mémoire des nœuds à l'aide notamment des commandes **bcfree** et **ipcrm**. Sur Cristal, l'ordonnanceur **SLURM** est couplé à un portail web **Sysfera** capable de réservations. Une fois le job de visualisation démarré, ce portail simplifie la redirection des flux X11 et OpenGL vers un navigateur web d'un poste client distant.

LLVIEW permet d'avoir une interface graphique donnant une vue globale du placement des jobs sur la machine Jade.

Flexlm gère les licences de nombreux logiciels commerciaux et est couplé avec un décompte de ressources spécifiques à chaque logiciel dans PBS.

C Les logiciels applicatifs

LOGICIELS	DOMAINE	MACHINES	STATUT
ABAQUS	Calcul de structures	Jade	Libre
ABINIT	Density Functional Theory	Jade, Yoda	Libre
ADF	Density Functional Theory	Jade, Yoda	Commercial
AMBER	Modélisation moléculaire	Jade	Commercial
CHARMM	Modélisation moléculaire	Jade, Yoda	Commercial
CP2K	Dynamique moléculaire	Jade, Yoda	Libre (licence)
CPMD	Dynamique moléculaire	Jade	Libre
DALTON	Dynamique moléculaire	Jade	Libre (licence)
ELK	Chimie computationnelle	Jade	Libre
FLUENT	Dynamique des fluides	Jade	Commercial
ENSIGHT	Visualisation	Cristal	Commercial
GAMESS	Chimie quantique	Jade	Libre
GAUSSIAN	Modélisation moléculaire	Jade, Yoda	Commercial
GROMACS	Dynamique moléculaire	Jade	Libre
GULP	Simulations sur matériaux	Jade	Libre
HYPRE	Bibliothèque scientifique	Jade	Libre
JAGUAR	Chimie quantique	Jade	Commercial
LAMMPS	Dynamique moléculaire	Jade	Libre
MATERIALS STUDIO	Multi-fonctions	Jade	Commercial
NAMD	Dynamique moléculaire	Jade, Yoda	Libre
NUMPY	Multi-fonctions	Jade	Libre
NWCHEM	Chimie computationnelle	Jade	Libre
OPENFOAM	Dynamique des fluides	Jade	Libre
PARAVIEW	Visualisation	Cristal	Libre
QUANTUM ESPRESSO	Calcul de structures	Jade	Libre
RPY	Post-traitement	Jade	Libre
SIESTA	Simulation électronique	Jade	Accès restreint
SCYPY	Bibliothèque scientifique	Jade	Libre
SUITE SCHRODINGER	Simulation	Jade	Commercial (licence)
TURBOMOLE	Calcul de structures	Jade	Commercial
VASP	Simulation	Jade, Yoda	Commercial (licence)
VISIT	Visualisation	Cristal	Libre

Les actions

a La procédure DARI de demandes d'heures

L'application nationale DARI de gestion et de suivi des demandes et des attributions d'heures de calcul sur les centres nationaux Tier-1 s'est significativement enrichie au cours de ces dernières années. Parmi les nombreux développements réalisés, on peut citer notamment :

Rubrique utilisateur

✚ **Consommation d'un projet sur les trois centres nationaux** : avant 2010, les utilisateurs n'avaient accès qu'à leur consommation sur le CINES. A partir de 2010, une procédure a été mise en place pour enregistrer quotidiennement et de manière automatique dans la base DARI la consommation des projets tournant à l'IDRIS et au TGCC.

✚ **Pénalité des projets sur le premier semestre** : sur décision de GENCI, les projets ayant peu consommé durant les cinq premiers mois de l'année peuvent se voir appliquer une pénalité au 1er juin.

✚ **Téléchargement du visa** : le visa (première page du dossier signée) était jusqu'à présent transmis au centre de calcul par courrier ou par fax. L'utilisateur a maintenant la possibilité de télécharger lui-même sur le site ce document après l'avoir scanné.

Rubrique Centre de calcul et GENCI

✚ **Expertise technique des centres** : sur demande de GENCI, les centres de calcul sont maintenant sollicités pour expertiser au plan technique les dossiers annuels et les demandes complémentaires du second semestre. Cette expertise peut être consultée par les experts des comités thématiques durant leur propre expertise scientifique des projets.

✚ **Reporting mensuel et trimestriel pour GENCI** : en fin de mois et en fin de trimestre un rapport est envoyé à GENCI synthétisant l'activité de la machine Jade.

Présidents de comité thématique :

✚ **Expertise scientifique et technique** : les présidents de CT ont maintenant la possibilité par simple clic de copier automatiquement dans leurs zones d'expertises les avis et commentaires de leurs experts.

Mésocentres :

✚ à la demande de GENCI, une nouvelle rubrique intitulée « Mésocentres » a été créée. Cette rubrique permet à toute personne d'un mésocentre disposant d'un code d'accès de consulter partiellement la base DARI. En fonction du code d'accès (qui identifie le mésocentre) le serveur DARI liste les projets pour lesquels un laboratoire est situé dans la même région et également les projets pour lesquels un accès à ce mésocentre a été signalé.

Développements étudiés en 2012 qui seront réalisés en 2013 :

✚ **Demandes de type accès préparatoire** : inspiré de la procédure « preparatory access » de PRACE, les utilisateurs du DARI peuvent désormais déposer un dossier de type accès préparatoire (dossier DARI simplifié). Si le projet est accepté, une allocation forfaitaire d'heures est attribuée au projet.

✚ **Régulation mensuelle des projets** : les « gros projets » (allocation supérieure ou égale à 1 M d'heures) sont maintenant étroitement surveillés par une procédure automatique qui, tous les mois, compare leur consommation réelle à une consommation théorique idéalement répartie sur l'année. Si le projet accuse un retard d'un mois de

consommation, le porteur du projet est alerté. Dès que le projet atteint deux mois de retard, le porteur de projet est alerté et le projet se voit appliquer une pénalité d'heures correspondant à un mois d'allocation.

✚ **IPEA** : cette nouvelle rubrique permet aux utilisateurs de déposer un dossier dans le cadre de l'IPEA (Initiative pour le Passage à l'Echelle des Applications). Ce dossier est ensuite évalué par un comité expert et, si le projet est accepté, une aide technique sera apportée. L'objectif est de permettre à certains codes d'utiliser les plus gros moyens de calculs académiques français (en l'occurrence la machine CURIE du TGCC).

b Le support aux utilisateurs

On commettrait une grave erreur si on réduisait le support aux utilisateurs au seul traitement des tickets d'incidents ouverts dans l'outil OTRS. Il va de soi que le groupe constitué de cinq ingénieurs dont l'activité est tournée principalement vers le support pour les applicatifs intègre cette tâche, mais ne réduit pas son activité à cela.

La mission de cette équipe couvre l'ensemble des tâches suivantes :

- ✚ accompagner les utilisateurs débutants dans la maîtrise des outils de calcul ;
- ✚ leur expliquer individuellement ce qu'ils peuvent attendre d'un supercalculateur, quels sont ses atouts, ses différences par rapport aux machines locales avec lesquelles ils sont familiarisés ;
- ✚ les informer individuellement sur tous les aspects qui méritent d'être examinés afin d'améliorer les performances des calculs ;
- ✚ effectuer des « profilings » de performance de certaines applications ;
- ✚ leur suggérer des améliorations dans leur manière d'utiliser une application scientifique sur le supercalculateur : cela va de l'adaptation des parties du programme dépendantes de l'architecture, au placement des tâches ou des threads, en passant, bien évidemment, par le choix des options de compilation ;

- ✚ rédiger des « Guides des bonnes pratiques », des notices techniques ;
- ✚ installer et maintenir les bibliothèques scientifiques, les langages ;
- ✚ mettre à leur disposition les modules associés, vérifier leurs dépendances et leur cohérence dans l'environnement utilisateur ;
- ✚ veiller à l'exploitation correcte des ressources ;
- ✚ identifier les raisons de l'échec d'un travail sur la machine et y remédier.

Cette équipe d'ingénieurs est également placée à l'interface entre les utilisateurs et les administrateurs systèmes des machines. Le travail de diagnostic sur un problème est souvent délicat : qui, du supercalculateur ou de l'utilisation inappropriée d'un code, est responsable de l'échec d'un calcul ? C'est un travail d'investigation très rigoureux et méthodique qu'il faut savoir mener et où l'expérience joue un rôle prépondérant.

Le diagnostic étant posé, il peut conduire à une identification d'un problème hardware ou software (Lustre, PBS, ...). A cette étape, l'incident ou la demande de l'utilisateur passe au niveau de problème (conformément au modèle ITIL adopté par le CINES) et doit être orienté vers les administrateurs systèmes. Toutefois, ceux-ci ont besoin d'être guidés dans la connaissance de l'application.

Cette activité quotidienne, faite d'imprévus, ne pourrait se réaliser dans un contexte aussi dynamique que le HPC aujourd'hui, sans une veille technologique et une veille scientifique actives. Cela se fait grâce à la participation régulière à des meetings nationaux (ORAP, Teratec ...) et internationaux (ISC et SC) où l'on peut suivre l'activité internationale et recueillir des informations que l'on pourra rediffuser auprès des chercheurs. Il s'agit donc, ici, d'une tâche, peu visible et pourtant fondamentale, qui consiste à se former continuellement pour ensuite mieux former et informer les chercheurs calculant au Centre.

GENCI fait également appel aux compétences de cette équipe support pour réaliser des expertises techniques sur les dossiers DARI demandant plus d'un million d'heures.

Les membres de ce groupe s'enrichissent d'expériences sur les autres supercalculateurs via des participations à des projets européens (PRACE notamment) qui les amènent à travailler sur d'autres environnements.

Ainsi, une partie du guide «Best Practices» sur la machine CURIE a été conçu par un membre de l'équipe CINES et deux «Preparatory Access» à CURIE ont été accompagnés par un autre de ses membres. Cela donne lieu à des publications accessibles sur le site de PRACE.

L'équipe support collabore également à des projets ANR (projets CAMPAS, STRASS,...).

Enfin, GENCI fait également appel à ce groupe support pour constituer et suivre le déroulement de bancs d'essais applicatifs dans le cadre des appels d'offres pour le renouvellement des moyens de calcul nationaux.

En résumé, cette mission intitulée trop simplement «Support aux utilisateurs» est plus multiple qu'il n'y paraît d'un premier abord. Elle fait appel à des compétences de plus en plus élargies : dix champs thématiques calculent au Centre, des technologies du HPC en perpétuelle évolution.

C La visualisation à distance

Le contexte

Depuis que le CINES met de grandes infrastructures de calcul à la disposition des chercheurs - soit plus de trois décennies - ces derniers se heurtent toujours aux mêmes obstacles. Les technologies évoluent, mais ces problèmes demeurent. Parmi ces problèmes, figure notamment une question récurrente :

Une fois les calculs terminés, que faire des données ?

Faut-il les effacer ? Les transférer dans son laboratoire ? En totalité ou en partie ? Comment transférer ?

Autrement dit : comment accéder à une chaîne de traitement complète ? Dans une chaîne de traitement, on trouve les phases de pré-traitement, de calcul, et pour finir de post-traitement incluant éventuellement une étape de visualisation. S'il est parfois possible de réaliser le pré-traitement

directement sur la machine de calcul, le CINES a cependant fait le choix d'offrir un environnement dédié. De même, pour le post-traitement et la visualisation, Le CINES propose une plate-forme et des services dédiés. L'ensemble de la chaîne – pré-traitement, calcul, post-traitement et visualisation – constitue ainsi une solution intégrée offrant une gamme complète de services au niveau de ce que les utilisateurs sont en droit d'attendre d'un grand centre national.

Deux obstacles à surmonter

Les calculs génèrent des volumes de données toujours plus importants, et malgré l'augmentation des débits réseau au plan national, il est toujours aussi difficile de déplacer ces données. La course entre l'augmentation des débits réseau et la puissance des machines est toujours à l'avantage des machines. Il apparaît dès lors judicieux de placer à proximité de la machine de calcul, un équipement destiné à pré/post-traiter les données, ou en faire des images. Partant de cette idée, nous nous heurtons rapidement à un second écueil. Si le débit réseau est rédhibitoire pour transférer les données brutes, il l'est aussi tout autant pour déporter les images.

Le passé : des solutions insatisfaisantes

Naguère, le déport de visualisation se faisait de façon assez frustrante via des protocoles réseaux standards, qui présentaient une limitation majeure : ils dégradaient la qualité de l'image.

Le présent : des potentialités nouvelles et intéressantes

Depuis, de nouveaux outils ont vu le jour, qui permettent de tirer pleinement profit des cartes graphiques peu onéreuses qui équipent aujourd'hui les postes de travail des chercheurs. Ces outils permettent enfin de « pousser » - push - des images de grande qualité jusqu'à l'écran de l'utilisateur final. Mais ces logiciels gardent un défaut majeur : ils nécessitent l'installation d'un client logiciel spécifique sur le poste distant.

Le projet du CINES

Convaincu du besoin réel éprouvé par la communauté et ayant bien identifié les derniers obstacles techniques à surmonter pour pouvoir offrir un service fluide et simple à mettre en œuvre le CINES a décidé en 2010 de proposer à ses utilisateurs une plate-forme de visualisation à distance. Une équipe de trois ingénieurs a été constituée, alliant des compétences dans les domaines de la visualisation, du système et du support aux utilisateurs.

La phase prototype

Pour offrir un service à la hauteur des capacités de calcul du CINES, il fallait un environnement de visualisation pluri-machines. Deux machines de la société Bull ont été retenues. Ces machines, de puissances remarquables avaient la capacité d'héberger tout à la fois de nombreux cœurs de calcul, mais aussi plusieurs GPU.

Ces deux clusters (Cristal 1 et 2) ont servi de base pour le projet. Une des machines était attachée à un espace disque local de plus de 10 To destiné à la visualisation. Les utilisateurs pouvaient ainsi réaliser leurs premiers tests, et plusieurs projets ont su tirer profit de ces machines. L'avantage obtenu par la proximité de la machine de visualisation et de la machine de calcul était très apprécié.

Durant cette première phase, des images, et des films ont été réalisés, les utilisateurs devaient

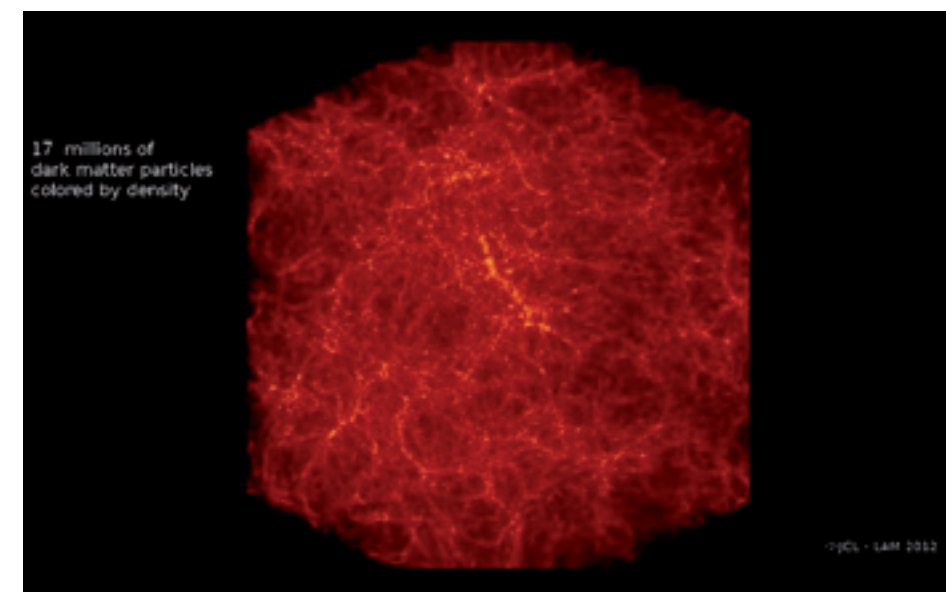
dans un premier temps se connecter via un outil assurant la qualité du déport (virtualGL), et ensuite, s'affichaient sur leur écran les images qui étaient calculées sur le serveur au CINES.

Cette phase, courant 2011, a été mise à profit pour mesurer la difficulté importante générée par l'installation obligée de logiciels sur le poste client. Il a alors été décidé d'ouvrir une nouvelle phase de ce projet. Cette phase 2, phase opérationnelle, devait permettre tout à la fois d'augmenter la capacité de visualisation mais aussi de faciliter l'accès aux ressources. C'est ainsi, qu'est née l'idée d'un portail de visualisation déportée.

Le portail de visualisation déportée

Fin 2011 a été installé un nouveau cluster. Ce cluster offre, dans sa configuration de base, une puissance de calcul très importante, à laquelle des cartes graphiques haut de gamme ont été ajoutées. Ces cartes disposent de GPU de dernière génération couplées à une grande capacité mémoire.

Ce cluster est constitué de huit machines identiques équipées des processeurs Sandy Bridge, de 128 Go de mémoire, et de deux cartes GPU Quadro 6000 de la société Nvidia. Un neuvième nœud a été rajouté à la configuration ; ce nœud, un peu plus gros que les autres, dispose de quatre processeurs Sandy Bridge, de deux cartes Quadro 7000, mais surtout de 1 To de mémoire. Avec cette configuration, le service est désormais en capacité d'adresser de très

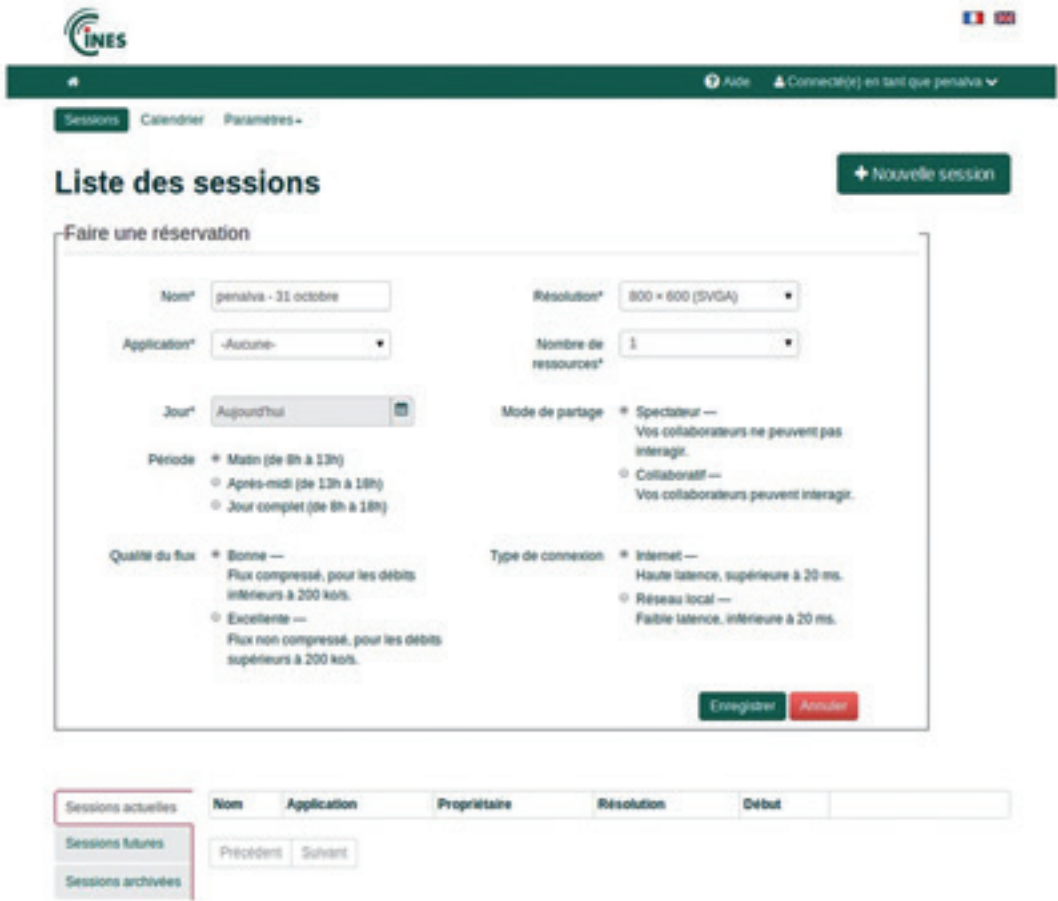


gros volumes impossibles à traiter jusque là. Pour interconnecter ces machines, un réseau Infiniband de dernière génération a été installé (Infiniband 4X FDR 56Gbits/s).

Côté Logiciels

Au plan logiciel, après avoir étudié les solutions en matière de déport, nous avons décidé d'offrir l'accès déporté au travers d'un portail web, cumulant . Ainsi, nous cumulons le double intérêt de ne pas avoir de logiciel client à installer, et de pouvoir offrir à tout moment et à tous la dernière version du logiciel (les mises à jour devenant

complètement transparentes pour les utilisateurs). Les trois principaux fournisseurs de ce type de solution actuellement présents sur le marché (Visuportail de la société Oxalya, Extreme Factory de Bull et le Webboard de Sysfera) ont été contactés dans le cadre de cette étude. C'est sur cette dernière offre que le choix s'est porté. Il s'agit d'un portail léger, facile à configurer et à modifier, et qui convenait tout à fait aux spécifications exprimées. Les tests ont pu commencer dès le début de l'année 2012 et les premiers résultats ont d'emblée été très encourageants.



Offre de service actuelle

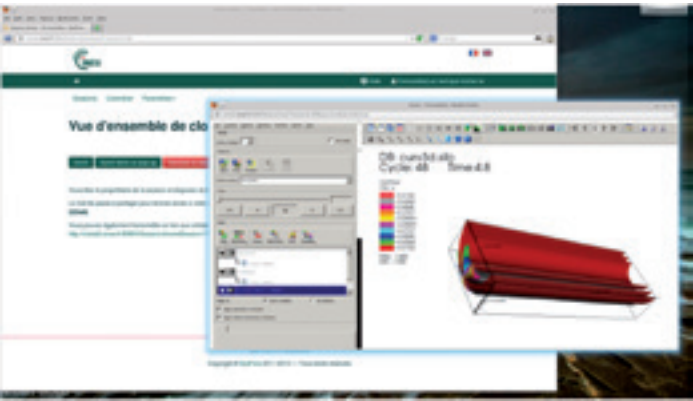
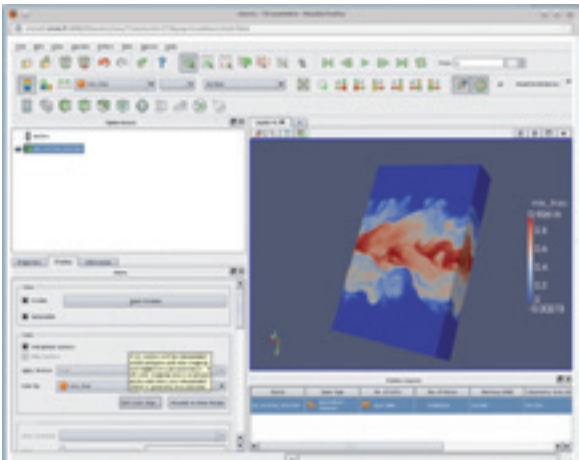
Le portail est encore en phase de recette, des premiers tests ont été réalisés avec succès auprès des utilisateurs. Certains gros volumes de données ont pu être visualisés, permettant enfin de s'affranchir de la barrière de la distance. Ci-contre une copie d'écran du portail au moment de la réservation d'une session dédiée pour la visualisation.

La solution mise en place respecte les deux règles que nous avons imposées dès le départ : simplicité et souplesse. On peut ainsi voir sur cette capture d'écran que seuls quelques champs nécessitent d'être complétés pour pouvoir commencer une session.

Les logiciels applicatifs

Ce portail donne un accès privilégié à quelques logiciels que nous avons sélectionnés. Parmi les logiciels que la communauté utilise depuis de nombreuses années, quelques-uns sont sortis du lot, et font désormais référence. Nous avons retenu :

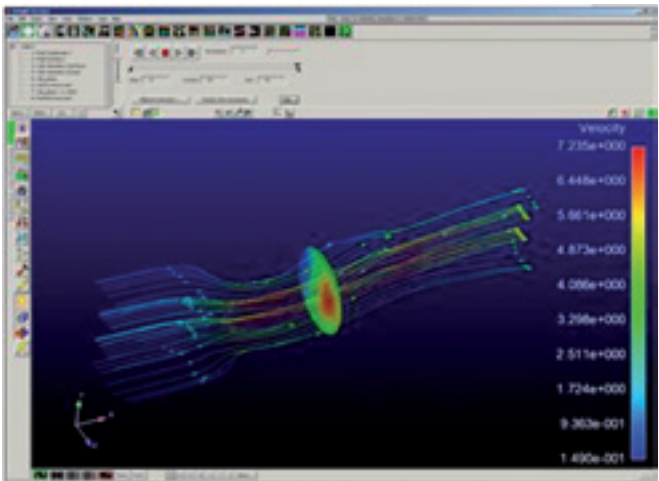
PARAVIEW.....
Le logiciel Paraview est un des standards en matière de visualisation. Largement utilisé par la communauté, il permet l'accès facile, grâce à son interface, à des images extraites de très gros volumes de données. Le choix de ce logiciel s'est imposé très rapidement, car de très nombreux utilisateurs le mentionnaient dans leurs dossiers de calcul.



VISIT.....
Visit est à l'identique de Paraview un des « poids lourds » de la visualisation. Il n'aurait pas été concevable d'offrir un service de qualité sans l'apport de ce second logiciel. Son interface est certes un peu plus basique, mais il est très puissant quand il est question de traiter de très gros volumes de données. Il a déjà été fait des démonstrations avec d'énormes volumes sur de très grosses machines. Ce logiciel est capable de tirer profit du parallélisme du cluster, ainsi que des multiples CPU et GPU de chaque nœud. La capture d'écran montre que l'application est directement accessible au travers du portail de visualisation.

ENSIGHT.....
C'est le seul logiciel commercial proposé. Il s'agit là aussi d'un « must-have » de la communauté. Avec cette offre, le CINES balaye un très large ensemble de fonctionnalités. L'ensemble des formats de fichiers reconnus par ces trois logiciels permet d'adresser la quasi totalité des fichiers générés par le calcul. Il existe aussi, et sur demande auprès de l'équipe support du DCI, la possibilité d'intégrer de nouveaux formats de fichiers en entrée d'un de ces logiciels.

Ce choix de logiciels n'est pas fermé, il pourra évoluer grâce à la souplesse de la solution retenue. Si un utilisateur propose un logiciel qu'il a développé lui-même, et souhaite son exploitation au travers du portail, cela peut être fait sur simple demande.



En conclusion

Le portail de visualisation est aujourd'hui une réussite et nous envisageons d'étendre la notion de portail au-delà du seul service de visualisation avec l'ouverture d'un futur service de type portail orienté calcul. D'autres articulations entre le calcul et la visualisation sont aussi en cours d'étude. Certains logiciels déjà utilisés permettent de lier la visualisation et le calcul pendant l'exécution d'un job. Cela fera sûrement l'objet d'un prochain chapitre lié à la visualisation au CINES.

d Le pré/post-traitement

Les simulations numériques produisent une grande quantité de données qu'il n'est pas toujours aisé de traiter et d'analyser. C'est pourquoi le CINES propose depuis 2010 un moyen de visualisation afin de faciliter l'exploitation des résultats de calculs obtenus sur les machines Jade et Cristal. Ce cluster peut aussi être utilisé en tant que moyen de pré et post-traitement. Le pré-traitement concerne la mise en données préalable à une simulation. Il s'agit de créer un jeu de données d'entrée nécessaire à l'exécution d'un calcul. À la différence de la visualisation, on entend ici par post-traitement l'opération de traitement des données issues d'un calcul sans éprouver le besoin d'interagir directement avec les données.

L'accès à Cristal est aujourd'hui réservé aux projets ayant reçu des heures de calcul via la procédure DARI. Les heures CPU utilisées sur Cristal ne sont pas décomptées des heures attribuées.

L'environnement logiciel proposé sur Cristal est basé sur un système d'exploitation Linux CENTOS sur lequel s'appuient les compilateurs GNU utilisés pour la compilation des programmes de pré/post-traitement. Certains outils et bibliothèques usuels de l'environnement pré/post-traitement sont d'emblée mis à disposition. On retrouvera, par exemple, les bibliothèques IO Netcdf et HDF5 pour la manipulation des données issues des simulations mais aussi les outils tels que gnuplot et ffmpeg pour réaliser graphiques, animations ou films à partir d'images.

Un des objectifs d'une telle machine est de proposer une grande quantité de mémoire par nœud. La capacité à traiter des données d'une

application non-parallèle ou faiblement « scalable » qui est donc, de fait, limitée par la quantité de mémoire par nœud s'en retrouve étendue. En effet, Cristal se compose d'un nœud de login et de neuf nœuds destinés aux travaux de visualisation et de pré/post-traitement. Parmi ces neuf nœuds, huit proposent 128 Go de mémoire chacun et le neuvième 1 To de mémoire (à titre de comparaison, Jade dispose, au maximum, de 34 Go de mémoire utilisable par nœud).

Un autre intérêt de Cristal est de proposer une connexion avec Jade via un lien réseau 10 Gb. Cette connexion, pour un utilisateur ayant effectué ses simulations sur Jade, lui permet d'éviter un transfert coûteux en temps vers une machine de son laboratoire.

Un projet est actuellement à l'étude afin d'établir une passerelle entre l'espace scratch de Jade et Cristal. Ainsi, à terme, il sera possible de travailler directement sur les données issues de simulations effectuées sur Jade.

Dans des domaines tels que la dynamique des fluides ou la climatologie, la réalisation de maillages définissant une structure ou un domaine géographique implique une utilisation importante de la mémoire. Cette opération de pré-traitement est parfois découplée du calcul et doit être réalisée au préalable. Dans ce cadre, ce type d'application serait tout à fait à même de tirer parti des caractéristiques de Cristal.

Parmi les projets DARI ayant eu accès à Cristal, l'un d'entre eux avait pour but de réaliser un post-traitement de données issues de simulations climatiques. Son cheminement est décrit ci-après : Dans un premier temps, les données sont transférées de Jade vers Cristal par l'intermédiaire du lien réseau 10 Gb.

Une partie des données, à l'origine sous format binaire propre au code utilisé, est convertie au format Netcdf. De plus, ces données résultats sont réorganisées pour obtenir un fichier par variable contenant l'ensemble des pas de temps au lieu d'un fichier par pas de temps contenant toutes les variables.

Le reste des résultats est constitué de fichiers au format Netcdf produits par chacun des processus MPI impliqués dans la simulation. Chacun de ces fichiers contient des informations spécifiques à une zone géographique du globe. Afin de représenter les résultats sur l'ensemble du globe, ces fichiers

sont rassemblés en un seul.

Ces différentes opérations impliquent un besoin en mémoire conséquent. Cristal a rendu possible, par la mise à disposition d'une grande quantité de mémoire, la simplification de la réorganisation des résultats d'une simulation afin d'en faciliter l'analyse a posteriori.

e Le C4

Le C4, Comité des Chercheurs Calculant au CINES est un relais crucial permettant de diffuser l'information auprès des utilisateurs du centre. C'est aussi une source principale d'informations pour les équipes du CINES, notamment pour un retour sur les besoins et problèmes. Le C4 joue également un rôle au niveau de la formation, l'animation scientifique (réunion Jeunes Chercheurs Calculant au CINES – Journée CCC), l'exploitation des supercalculateurs du Centre, l'achat des logiciels et l'évolution des équipements.

Durant les deux mandats précédent, j'étais représentant élu des chercheurs pour le Comité 7 «Systèmes moléculaires organisés et biologie» et président du C4. Durant les premières années, (avant l'ère GENCI-Jade) nous avons au sein de ce comité défendu les intérêts des chercheurs en gérant tant bien que mal l'adéquation des demandes et la pénurie des moyens informatiques. Comme président du C4, j'ai été amené à siéger à plusieurs reprises aux réunions du Conseil Scientifique du CINES pour l'octroi des heures de calculs.

Avec les autres membres du C4 et grâce à la mobilisation de l'ensemble des chercheurs, nous avons contribué à convaincre les instances de tutelle de doter le CINES de moyens informatiques dignes de notre communauté. Nous avons défendu la position du Centre et des Chercheurs Calculant au CINES devant le Comité Stratégique du Calcul Intensif (CSCI), et depuis sa création, travaillé étroitement avec GENCI. Grâce à la mobilisation et aux efforts du Ministre de l'Enseignement Supérieur et de la Recherche, de GENCI, du CINES et de plusieurs acteurs scientifiques, 2008 a vu le Centre s'élever à un rang très honorable dans le classement des centres de calcul et la machine Jade nous a donné entière satisfaction. L'accessibilité à cette machine a ouvert de nouvelles voies

à de nombreux chercheurs et a remplacé notre communauté dans la compétition internationale. Comme nous l'avions anticipé, les demandes d'heures de calcul sur ce type d'architecture ont, depuis, augmenté d'une manière significative. L'apport du calcul intensif pour les systèmes biologiques, comme dans les autres domaines est devenu en effet de plus en plus évident pour plusieurs groupes de recherches. L'extension de JADE moins de deux ans après son installation, et sa forte sollicitation par la communauté est venue confirmer cette donne.

Le C4 a durant toute la période entretenu un excellent rapport avec les équipes du CINES pour que les calculateurs et les services mis à la disposition de notre communauté soient à la hauteur des défis scientifiques que nous tentons de relever. Au-delà des contacts permanents entre ses membres et le CINES, des réunions (annuelles ou bi-annuelles) ont rythmé le travail du C4 en partenariat avec le centre. Consacrées principalement à un échange constructif, ces réunions ont été l'occasion pour le centre de présenter en amont ses plans d'action pour l'évolution des services qu'il propose, et, pour le C4, de communiquer les besoins des chercheurs. Ces échanges avec le CINES ont toujours été précédés par des réunions du C4 en conclave, permettant à ses membres d'échanger les informations remontant des chercheurs, et de prendre la mesure des spécificités de chaque comité thématique. De même, une séance plénière dans la programmation des journées, impliquant la participation des représentants de GENCI fut l'occasion souvent de faire remonter informations, requêtes ou doléances auprès des tutelles. Comme à l'accoutumée, la direction et le personnel du CINES ont été à l'écoute, ont fourni des efforts conséquents et constants pour que les ressources et logiciels soient disponibles le plus rapidement possible, permettant ainsi aux chercheurs de mener rapidement à bien leurs travaux de recherches.

Le C4 a également lors de cette période organisé plusieurs journées scientifiques permettant aux jeunes chercheurs de présenter via des interventions orales et posters, les résultats de leurs travaux sur les machines du CINES. Organisés par « thématique », ces journées ont été aussi l'occasion d'inviter un chercheur renommé dans le domaine pour une conférence plénière. L'un des

derniers invités fut le Professeur Martin Karplus, du Laboratoire de chimie biophysique Université Louis Pasteur, Strasbourg et Université de Harvard, lauréat du prix Nobel de Chimie 2013. Ainsi, après quelques années au C₄ et basé sur de nombreux témoignages des utilisateurs du centre, j'aimerais pouvoir exprimer une immense gratitude aux autres membres du Comité et au personnel du CINES dont le professionnalisme et le dévouement reconnus ont contribué à faire en sorte que le service rendu à la communauté soit à la hauteur des attentes. Aujourd'hui, à l'instar des autres centres de calcul (TGCC et IDRIS), GENCI dote le CINES

via un nouvel appel d'offres, d'un supercalculateur scalaire parallèle de dernière génération, dont l'installation est prévue au second semestre 2014. C'est dans ce contexte que je souhaite à la nouvelle équipe du C₄ succès et courage pour continuer à défendre les intérêts des chercheurs auprès du CINES et de GENCI, et de continuer à animer une composante scientifique en soutenant et en organisant des Ecoles thématiques.

Mounir TAREK, Directeur de Recherches CNRS, Président sortant du C₄

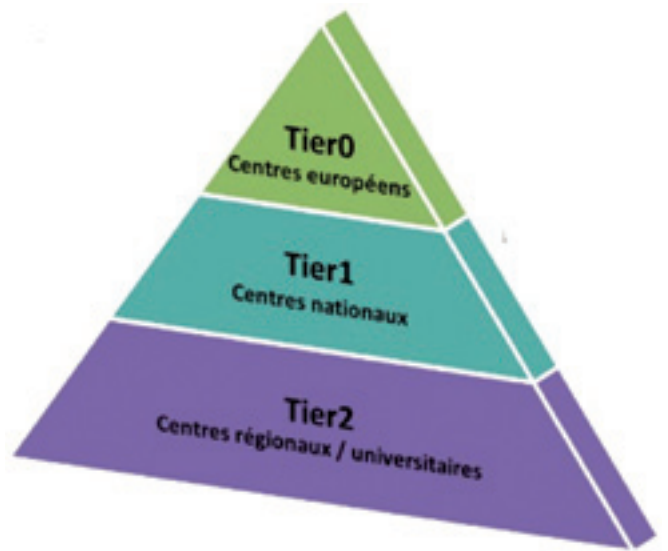
d'un accueil sur site. Une instance spécifique, le CCCC ou C₄ (Comité des Chercheurs Calculant au CINES), permet de maintenir un dialogue permanent entre le CINES et ses utilisateurs. Un recueil de l'expression des besoins et un retour sur la qualité des services proposés par le CINES sont

établis lors de réunions régulières (une ou deux réunions par an). Le CINES propose également des formations à la programmation parallèle et à la visualisation scientifique. Des journées d'animation thématique complètent le dispositif d'accompagnement des utilisateurs.

Le Cines acteur du HPC

a Au plan national

La création en 2007 de GENCI (Grand Équipement National de Calcul Intensif) et l'initiative européenne PRACE à partir de 2008 ont très fortement contribué à harmoniser et à structurer ce que l'on a désormais coutume de nommer l'écosystème du calcul intensif. Le CINES est l'un des trois centres nationaux français de calcul - centres Tier-1 dans la terminologie PRACE - aux côtés de l'IDRIS (CNRS) et du CCRT/TGCC (CEA).



PRACE : la pyramide des 3 niveaux de centres de calcul

Le CINES fournit un service de calcul intensif à la recherche publique nationale. Il donne aux chercheurs des universités et des organismes publics de recherche français accès à un ensemble de moyens informatiques de très grande puissance permettant non seulement d'effectuer des simulations numériques de grande dimension mais aussi d'en stocker les résultats et d'en permettre l'analyse à distance. Les architectures des supercalculateurs mis à disposition autorisent la réalisation de simulations numériques sur plusieurs milliers de processeurs. Ce service est disponible 24h/24h, tous les jours de l'année et il est accessible à distance dans d'excellentes conditions via le réseau RENATER. Les modalités d'accès aux équipements de calcul nationaux sont identiques pour les trois centres et pour tous les chercheurs. Un dispositif national unifié - procédure DARI (Demande d'Attribution de Ressources Informatiques) - permet une attribution annuelle de ressources après évaluation des dossiers de demandes par des équipes d'experts scientifiques et techniques. Selon l'état de la disponibilité des ressources de calcul et l'état d'avancement des projets de recherche, une attribution complémentaire d'heures de calcul est possible en milieu d'année. Le CINES propose une assistance de haut niveau technique à l'utilisation de ses moyens de calcul. Une aide à l'utilisation des équipements et à l'optimisation des codes est proposée, soit à distance (téléphone, courriel), soit dans le cadre

b En Europe

Le CINES poursuit son engagement dans le programme Européen PRACE.



Le partenariat PRACE (Partnership for Advanced Computing in Europe) prépare la création d'une infrastructure paneuropéenne permanente de calcul, se composant de plusieurs centres de classe Tier-O, complétant vers le haut la gamme de supercalculateurs. Les chercheurs européens auront ainsi à leur disposition un « écosystème » de moyens de calcul haute performance du plus haut niveau. Le projet de PRACE fait partie du 7ème programme-cadre de l'UE. Il est structuré en plusieurs « Work Packages » :

- + Gestion du projet ;
- + Préparation de l'infrastructure de recherche ;
- + Communication, formation, contacts ;
- + Gestion de l'écosystème de calculateurs ;
- + Applications clés ;
- + Prototypes pour systèmes multi pétaflopiques.



Sur la période 2010 – 2012, trois programmes spécifiques PRACE se sont succédés et chevauchés, la famille des IP (Implementation Phase) : PRACE-1IP, PRACE-2IP, PRACE-3IP. Alors que 1IP s'est terminé en juin 2012 et 2IP en septembre 2013, 3IP se poursuivra jusqu'en 2015.

L'engagement du CINES dans ces programmes est significatif : Une activité prévue contractuellement de 84 Mois x Homme sur la période répartis en fait sur la moitié de l'équipe calcul intensif, alors que 35 M x H sont encore à venir pour PRACE-3IP.

Préparation de l'infrastructure de recherche

Cette activité a pour objectif la constitution des éléments de fédération des moyens de calcul mis à disposition des utilisateurs européens. Elle comprend la constitution du réseau des machines Tier-0 et celui de machines Tier-1, ainsi que tous les services nécessaires au fonctionnement et aux accès (authentification, outils de transfert, surveillance).

Formations

Le CINES est impliqué dans le programme PATC (PRACE Advance Training Centers) aux côtés de la maison de la simulation (CEA) et de l'IDRIS (CNRS).

Applications du calcul intensif

Deux axes d'activités principaux se distinguent :

✚ Le passage à l'échelle, ou programmes PA « preparatory access », qui offre une assistance au passage à l'échelle d'applications. Le CINES a soutenu les programmes suivants :

✚ PA0398, Navier-Stokes Multi-Block pour l'IMFT, Institut de Mécanique des Fluides de Toulouse.

✚ PA0444, NEMO code de climatologie

✚ Le « Benchmarking » ou passage de bancs d'essais, l'activité a pour objectif la constitution d'une suite d'applications européennes du calcul intensif. Cela comprend le portage, le réglage et les mesures de performances de ces applications pour les architectures émergentes. La suite est rendue disponible sur le site, hébergé au CINES : <http://www.prace-ri.eu/ueabs?lang=en>

Prototypage

L'aspect collaboratif est ici prépondérant. A l'instar de PRACE-PP (2008-2009) avec le prototype ClearSpeed, l'activité prototype a été poursuivie :

✚ AMFT – Advanced Multi Fault Tolerance, une collaboration avec l'INRIA, Argonne National Laboratory, le BSC (Barcelone), GENCI et le CINES a permis l'essor du développement d'une librairie (Fault Tolerance Interface) qui a pour but d'implémenter la résilience au niveau applicatif dans un contexte massivement parallèle et préparer ainsi les challenges soulevés par les futures machines de calcul intensif dont le nombre d'unités de traitement sera de l'ordre du million. Le principe de base repose sur la mise en place de fichiers de reprise en mode RAID sur les nœuds de

calcul ; Le projet a utilisé la machine CURIE (TGCC – Bruyères le Châtel) et la machine Ambre au CINES.

✚ EXASCALE I/O – Une collaboration avec le CEA, qui a pour objectif d'évaluer de nouvelles technologies de système de fichiers parallèles, dans la perspective de l'évolution des plates-formes de calcul et des volumes de données considérables qui seront produits.

Le prototype a été installé et réparti sur les sites du CINES et du TGCC, avec des activités comprenant la mise en œuvre de traitements sur une infrastructure répartie.

Intégration du CINES dans la grille PRACE

L'intégration d'un nouveau site dans la grille européenne PRACE requiert la mise en place de plusieurs éléments : matériels et logiciels.

D'un point de vue matériel, il s'agit de connecter le site à l'infrastructure réseau dédiée PRACE. Le CINES a acquis pour cela en 2011 un lien 10 Gbps Ethernet spécifique et dédié auprès de RENATER. Cette liaison a permis au CINES d'une part de se raccorder à la plate-forme réseau existante sur laquelle les deux autres centres français IDRIS et TGCC étaient déjà connectés, et d'autre part d'intégrer ainsi le réseau européen dédié. Grâce à cette liaison, et avec l'appui de ces deux autres centres nationaux, une reconfiguration du réseau national a été effectuée fin 2012 pour mettre en place une interconnexion directe de ces centres offrant ainsi une infrastructure très haut débit pour les échanges nationaux de données. Cette infrastructure nationale pourra ensuite évoluer pour intégrer d'autres sites.

Pour ce qui est des logiciels, PRACE requiert le déploiement de services sur chacun de ses sites. Ces services sont classés en 6 catégories : calcul, données, suivi de fonctionnement, utilisateur, réseau, autorisation/comptabilité/authentification. Chacun de ces services nécessite le déploiement d'un intergiciel spécifique hérité du projet DEISA et basé sur l'environnement « Globus Toolkit ». L'ensemble de ces outils ont été déployés autour de la machine Jade que le CINES

a ainsi affichée comme composant de la grille européenne.

Liste des services déployés au CINES dans le cadre de l'intégration à la grille PRACE

SERVICES « CALCUL » :

✚ UNICORE (Uniform Interface to Computing Resources) : outil distribué de soumission et de gestion de travaux (et des données associées) sur la grille. L'outil a été installé au dessus de PBS, le gestionnaire de travaux natif de la machine Jade.

SERVICE « DONNÉES » :

✚ GridFTP : outil de transfert de fichiers haut débit. Deux serveurs GridFTP ont été déployés: le premier pour les échanges de données haut débit entre les nœuds de la grille PRACE et le second pour les laboratoires utilisateurs.

✚ Gtransfer : interface conçue pour faciliter l'utilisation des commandes de transfert standard de GridFTP pour lequel nous avons configuré et diffusé la description de l'interface de Jade dans la grille.

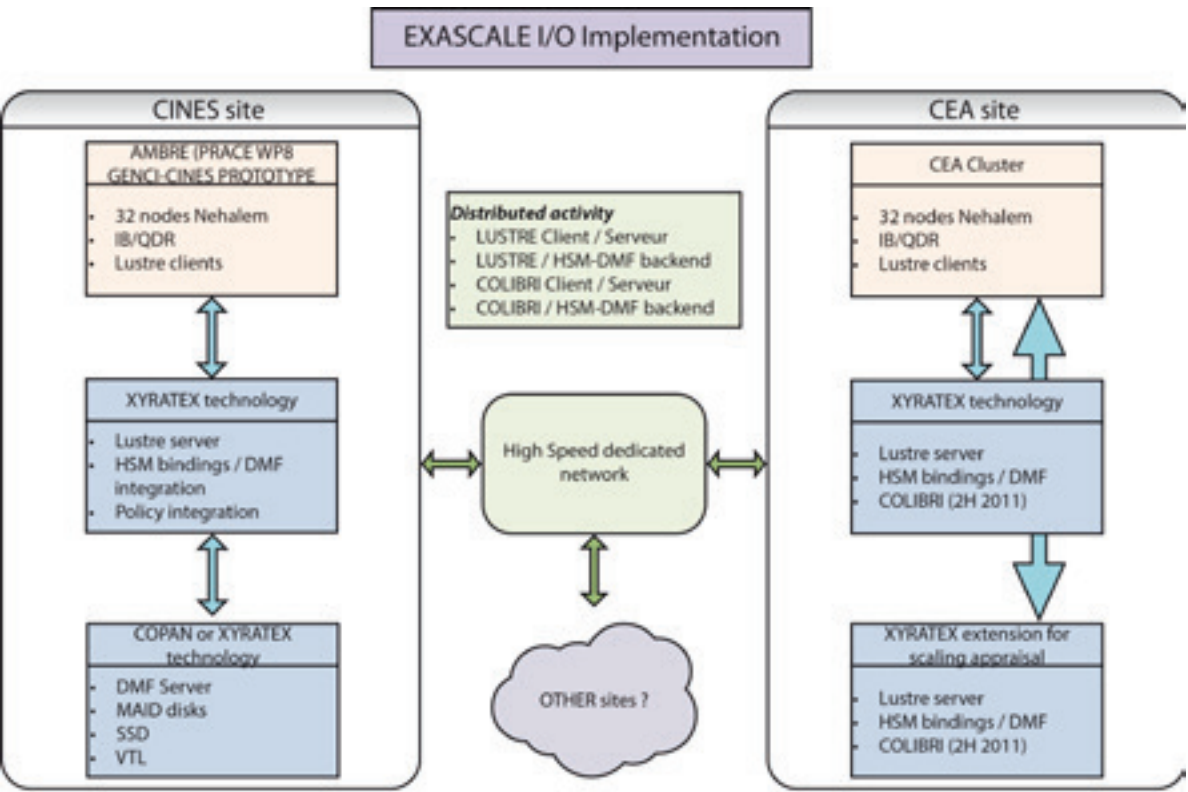
SERVICES « SUIVI DE FONCTIONNEMENT » :

✚ Inca : environnement de suivi de fonctionnement des services PRACE. L'installation s'est faite dans un mode où le CINES maîtrise totalement les contrôles qui sont effectués sur son environnement. Le CINES a installé un environnement local propre (Inca Reporter Manager) qui transmet, à la demande, l'état des services et configurations de Jade spécifiques à PRACE, à un serveur Inca central, unique dans la grille européenne.

SERVICES « UTILISATEURS » :

✚ Documentation : Un ensemble de documentations ont été produites spécialement pour les utilisateurs PRACE. Ces documentations décrivent les protocoles d'accès, les outils de transfert de données et d'utilisation des moyens de calcul du CINES.

✚ PCPE (Prace Common Production Environment) : environnement uniformisé



d'accès aux logiciels sur les différentes machines de la grille.

✚ **PraceService** : outil permettant de connaître les paramètres d'accès aux services déployés sur les sites de la grille.

SERVICES « RÉSEAU » :

✚ **Deisamonctl** : suivi des performances de chacune des liaisons du réseau dédié.

AUTORISATION, COMPTABILITÉ, AUTHENTIFICATION :

✚ **Comptabilité** : mise en place sur le CINES de l'environnement de comptabilité des travaux PRACE.

✚ **LDAP** : mise en place sur le CINES d'un environnement de gestion des utilisateurs PRACE couplé avec le système LDAP du CINES et dans la base centrale PRACE.

✚ **PKI** : outils de gestion des certificats pour l'authentification des utilisateurs.

✚ **GSI-SSH** : outil de connexion aux machines utilisant les certificats.

PPR – PRACE PEER REVIEW

Depuis le printemps 2010, suite à un appel à candidatures européen de PRACE, le CINES s'est vu confier la conception et le développement de l'application de gestion du processus d'évaluation par les pairs des ressources de calcul haute performance sur les machines européennes Tier-O. Cette application nommée « Prace Peer Review Tool », PPR en abrégé, permet de gérer de manière homogène, cohérente et sécurisée les demandes et les attributions de ressources de calcul lors des appels à projets PRACE.

Pour le compte de PRACE, le CINES assure ainsi annuellement la gestion de :

✚ deux appels à projets pour les accès réguliers - regular access – représentant l'évaluation administrative, technique et scientifique d'environ 150 projets par appel souhaitant avoir accès aux ressources des supercalculateurs Tier-O.

✚ quatre appels à projets pour les accès préparatoires - preparatory access - pour un

volume d'environ 30 à 50 projets de portage et de test des codes dont l'évaluation simplifiée se limite aux plans administratif et technique.

A la fin 2012, sept campagnes d'accès réguliers et onze campagnes d'accès préparatoires, soit environ deux mille candidatures ont ainsi été évaluées par l'outil PPR dans le cadre d'un contrat entre le CINES et PRACE-AISBL. Il est prévu que le CINES prenne en charge en 2013 la gestion des processus similaires pour les machines nationales Tier-1 avec une déclinaison de l'outil PPR spécifique au processus DECI. Afin de parvenir à cet objectif, une équipe de trois ingénieurs travaille à plein temps depuis 2010 à la conception, au développement, à l'hébergement et à la maintenance de ces outils. L'objectif à terme du CINES est de mettre en place un outil suffisamment générique et configurable pour permettre la gestion de tous les cas d'utilisation des processus de Peer Review pour la communauté européenne.

HPC – EUROPA

HPC-Europaz est un programme européen d'échange permettant à des chercheurs de tout pays éligible de rendre visite à un institut de recherche étranger adhérent au programme, pour entreprendre une collaboration scientifique pouvant durer jusqu'à trois mois, et avoir ainsi accès aux plus puissantes installations de calcul intensif européennes :

✚ BSC	(Barcelone)
✚ CINECA	(Bologne)
✚ CSC	(Helsinki)
✚ EPCC	(Edimbourg)
✚ GENCI/CINES	(Montpellier)
✚ HLRS	(Stuttgart)
✚ SARA	(Amsterdam)

HPC-Europaz favorise l'accès transnational scientifique interdisciplinaire par la prise en charge de frais de transport et d'hébergement. GENCI et le CINES en assurent la participation au niveau français. L'ouverture transnationale s'entend dans un esprit d'ouverture dépassant le cadre strictement européen comme en atteste la liste des pays éligibles :

L'Albanie, l'Autriche, la Bosnie-Herzégovine, la Belgique, la Bulgarie, la Croatie, Chypre, la République Tchèque, le Danemark, l'Estonie, la Finlande, la France, l'Allemagne, la Grèce, la Hongrie, l'Islande, l'Irlande, Israël, l'Italie, la Lettonie, le Liechtenstein, la Lituanie, le Luxembourg, l'ARYM, Malte, le Monténégro, les Pays-Bas, la Norvège, la Pologne, le Portugal, la Roumanie, la Serbie, la Slovaquie, la Slovénie, l'Espagne, la Suède, la Suisse, la Turquie et le Royaume-Uni.

Le rôle du CINES dans ce dispositif a été d'assurer l'accueil sur site des participants sur une période de quelques jours et de leur apporter l'assistance nécessaire pour mener à bien les aspects calcul intensif de leur projet. Nous avons également géré les évaluations des dossiers de candidatures concernant nos moyens de calcul. Associés au CINES, des laboratoires d'accueil nationaux complètent le dispositif français de cette collaboration internationale, pour la partie recherche scientifique.

Pendant la période 2010 – 2012, le CINES a produit près de 7 millions d'heures sur Jade pour les 131 visiteurs qui ont été accueillis au CINES.

Enfin le CINES a organisé et accueilli en 2009 le TAM (Transnational Access Meeting) qui a réuni sur trois jours à Montpellier plus de 120 chercheurs qui ont présenté les résultats scientifiques obtenus dans le cadre du programme HPC-Europaz.

C Au plan régional

La proximité entre la Région Languedoc-Roussillon et le calcul date du milieu des années 1960 avec l'implantation à Montpellier de l'usine IBM (plus de 2 000 emplois créés). En 1980, la création par le ministère en charge de l'enseignement supérieur d'un centre de calcul national à Montpellier - le CNUC - est venue confirmer la vocation « hautes technologies » de la ville et au-delà de la région toute entière. Les années 2010-2012 ont encore renforcé cette tendance au travers de la création du centre régional Languedoc-Roussillon de compétence en calcul intensif - HPC@LR - initiative à laquelle le CINES a contribué avec d'autres partenaires régionaux.

Le centre de compétence HPC@LR vise à créer en Languedoc-Roussillon un centre d'expertise dévolu au calcul intensif à destination des entrepreneurs

régionaux, des chercheurs et ingénieurs des secteurs public ou privé ainsi qu'aux enseignants. Soutenu financièrement par la Région, l'Europe (fonds FEDER) et l'Université Montpellier 2, le centre HPC@LR propose un modèle original s'appuyant à la fois sur des moyens matériels et humains.

La plate-forme matérielle hybride vise à servir les recherches menées dans les différentes spécialités d'excellence en permettant des comparaisons étendues entre les différentes architectures disponibles.

L'accompagnement permet aux utilisateurs qui jusqu'à présent n'avaient pas accès et n'avaient pas envisagé de telles ressources d'utiliser ces outils de calcul pour faire face à leurs nouveaux besoins : besoins de calculs plus fins, volumes de données de plus en plus importants et de moins en moins réductibles aux méthodes plus classiques. La plate-forme servira également à des fins pédagogiques : formation continue, formations académiques, séminaires, tutoriaux, etc.

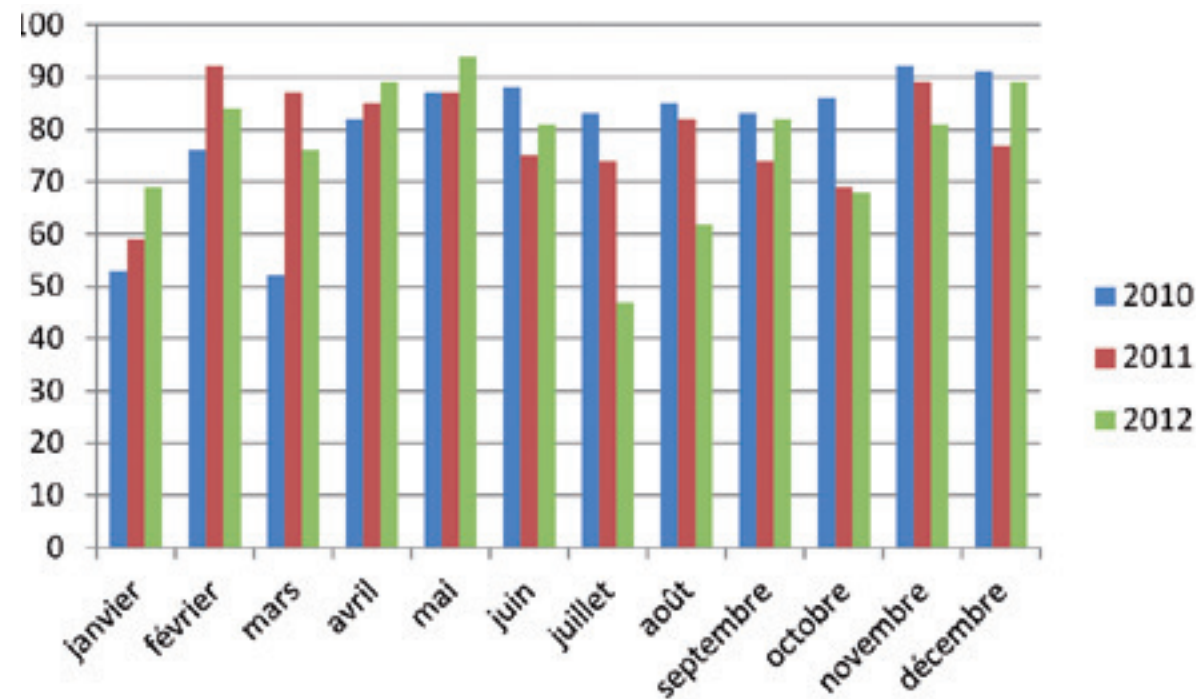
Les services offerts par le centre de compétences HPC@LR viennent compléter l'offre des centres de niveau national, parmi lesquels le CINES, en proposant un processus léger et réactif de sélection des candidats et d'accès aux ressources tout en étant ouvert aux entreprises et au monde universitaire.

La contribution du CINES au projet HPC@LR se situe sur trois plans : dans le montage initial du projet, le CINES a apporté son expérience et son expertise technique ; dans la phase d'exploitation, le CINES est le centre hébergeur des calculateurs et, enfin, il offre une prestation de support et d'accompagnement aux équipes porteuses de projets académiques.

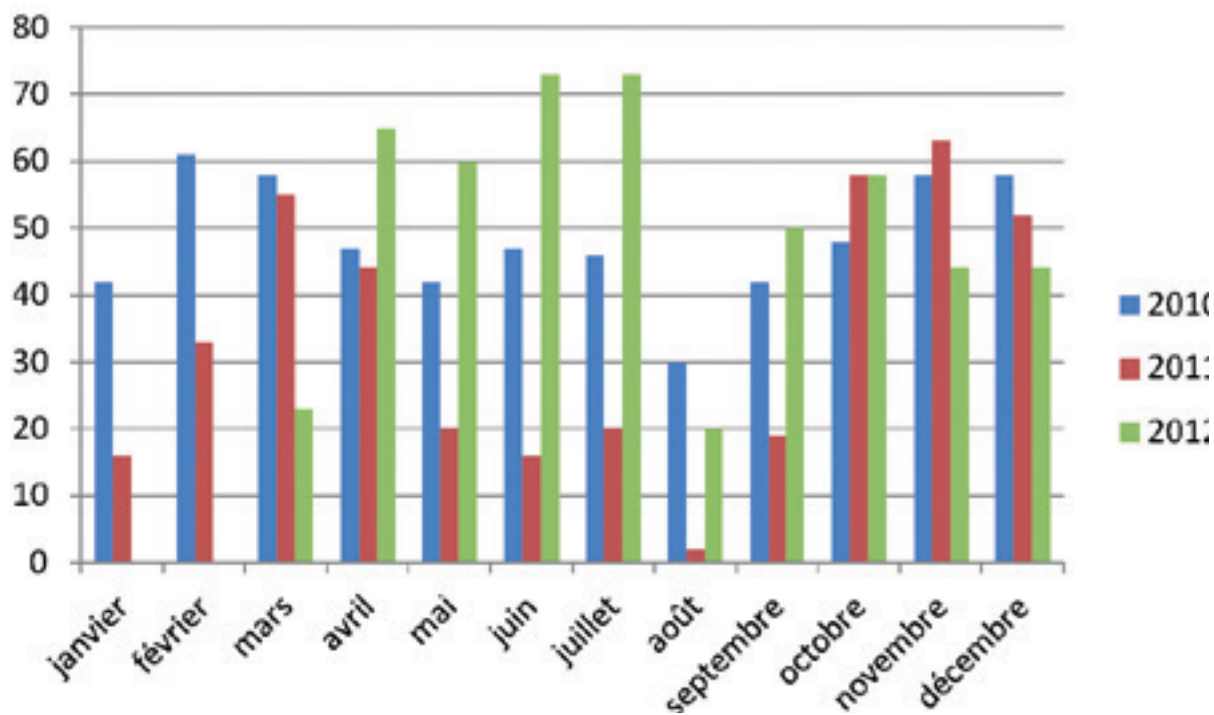
Outre les partenaires financeurs et le CINES, les autres partenaires régionaux de HPC@LR sont les entreprises ASA, SERVIWARE et SILKAN.

3 ans de calcul au Cines résumés en 7 graphiques

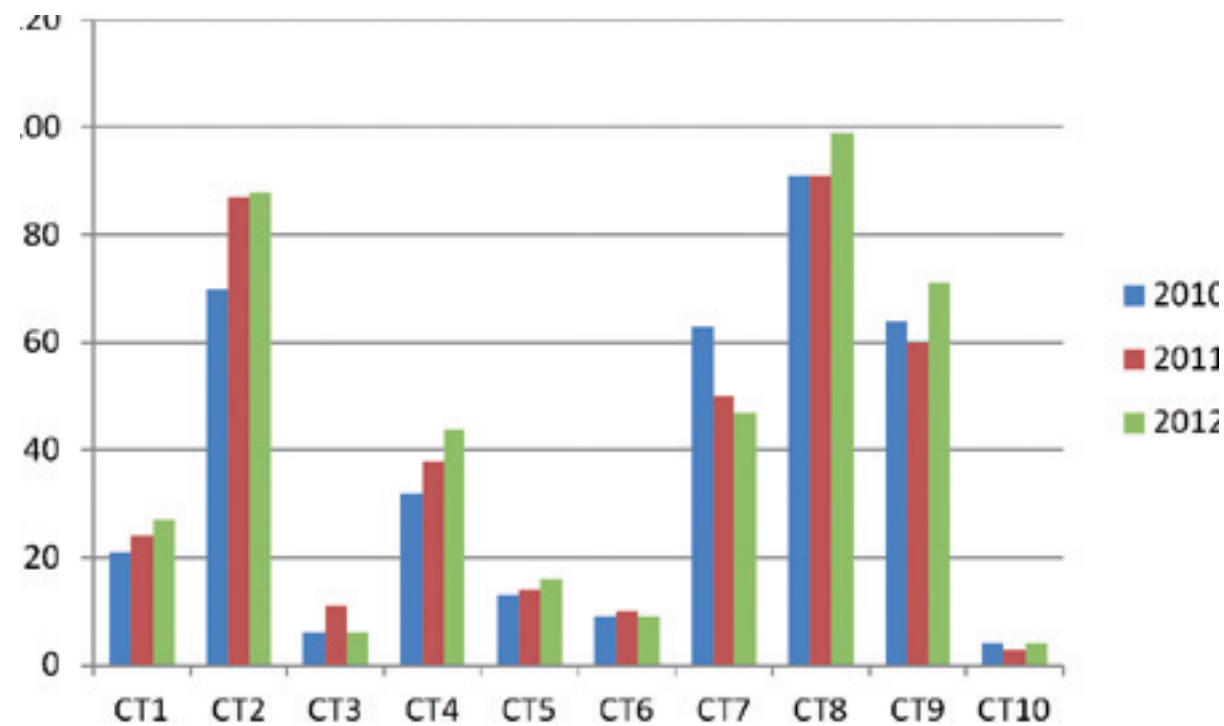
a L'activité de la machine Jade sur 36 mois



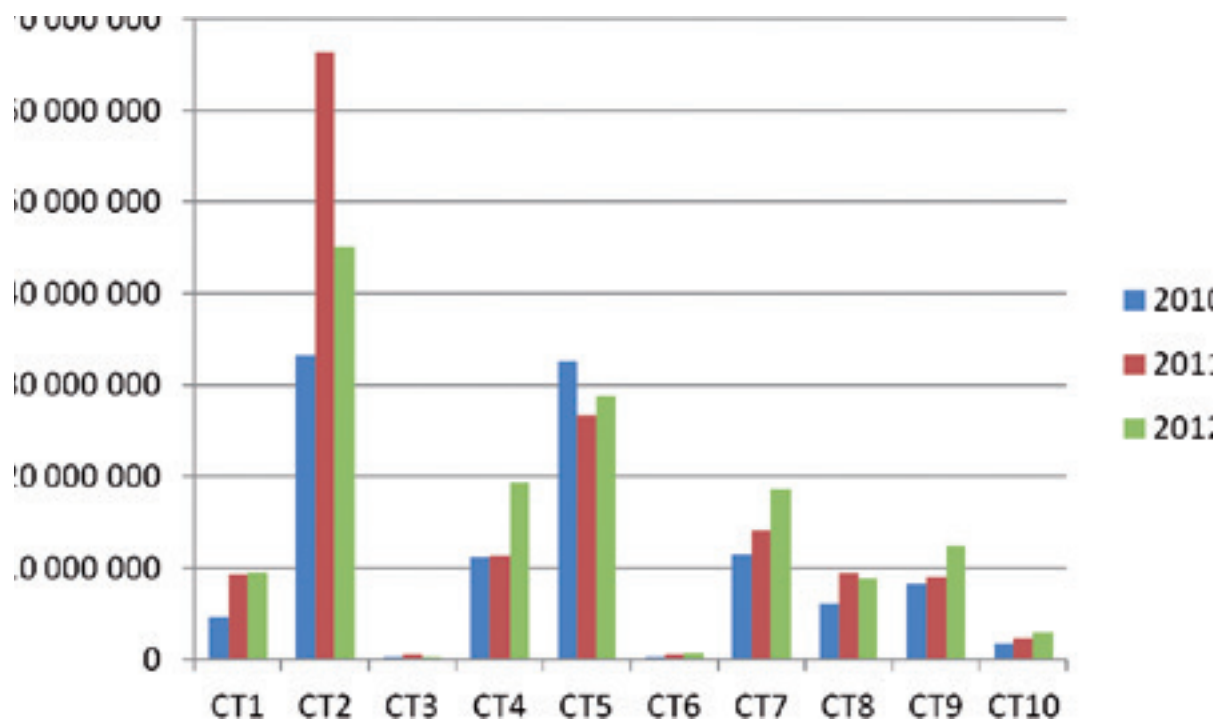
b L'activité de la machine Yoda sur 36 mois



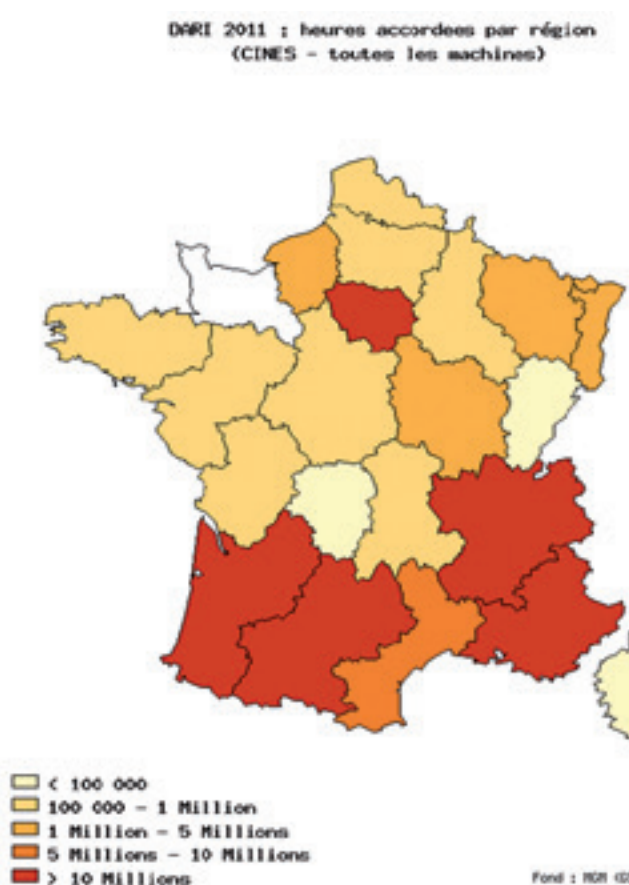
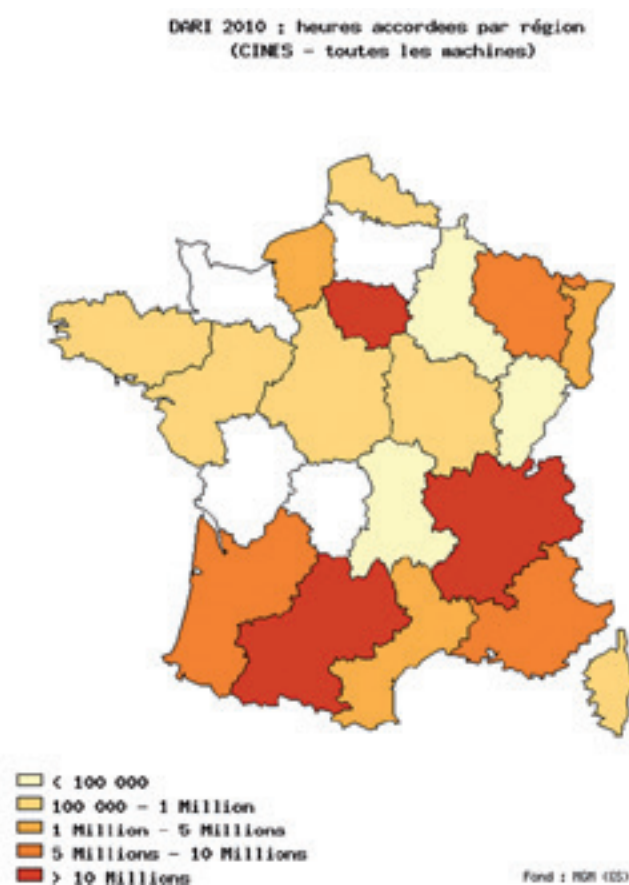
c Le nombre de projets retenus par CT et par an



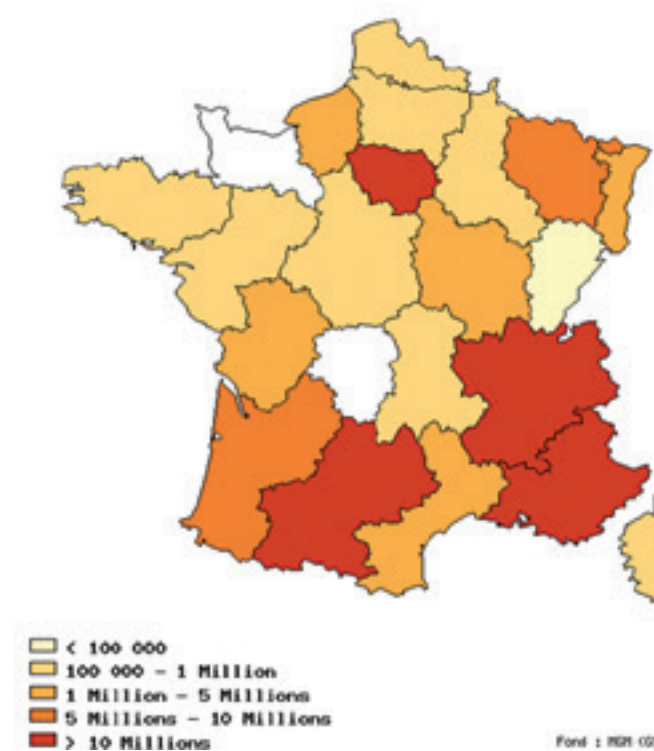
d La consommation de chaque CT par an



e L'origine géographique des projets retenus



DARI 2012 : heures accordées par région
(CINES - toutes les machines)





LE DÉPARTEMENT	32
Les missions	32
Structure	32
Moyens humains	32
Les formations	32
ADMINISTRATION DES SYSTÈMES ET DES DONNÉES	33
Le périmètre	33
Les serveurs d'archivage	33
Les données scientifiques	33
Convergence : partage mobilité, sécurisation et archivage des données scientifiques	34
LES SERVICES RÉSEAUX ET LA SÉCURITÉ	35
Le réseau évolue	35
Les services	35
La sécurité	35
L'EXPLOITATION	36
La suspension et le suivi des services	36
Le centre des services : SVP	36
L'astreinte	37
ITIL	38
L'hébergement pour des tiers	39
L'activité en chiffres	39
BILAN ET PERSPECTIVES	39

Système de données # Exploitation # Réseau, hébergement et Infrastructure



Le département

a Les missions

L'organisation actuelle du CINES a été mise en place au 01/09/2008. A cette date, les missions du Département Services Informatiques et Infrastructures (DS2I) ont été définies comme suit. Le DS2I est en charge des infrastructures informatiques communes sur lesquelles s'appuient les équipements des autres départements. Il veille à la disponibilité des ressources 24 heures sur 24 et à la sécurité des données. Il assure également le centre d'appel de premier niveau.

Le département assure l'administration des systèmes informatiques, excepté les serveurs de calcul. Il fournit les services fondamentaux sur lesquels s'appuient les autres départements. Il assure également l'exploitation 24h/24 des ces services. Il intervient donc sur les infrastructures réseau, sur les serveurs, sur les services. Il assure la surveillance, la métrologie, et pallie les dysfonctionnements.

Le département s'est vu confier également une activité d'hébergement d'environnements informatiques, d'envergure nationale, tirant partie de la mutualisation d'infrastructures performantes.

b Structure

Le département est organisé autour de 3 pôles d'activités :

- ✚ les systèmes et les données,
- ✚ le réseau, l'hébergement et infrastructure,
- ✚ l'exploitation.

Chaque personne est attachée à un ou plusieurs pôles.

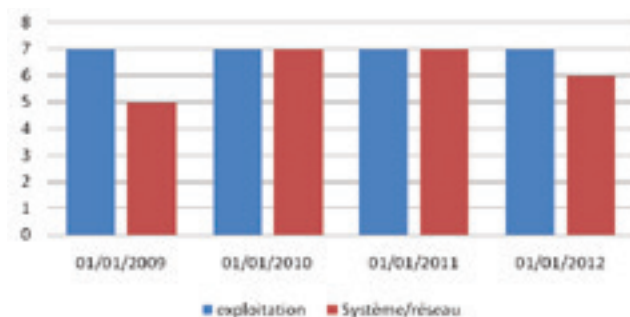
c Moyens humains

Les effectifs du département se situent en moyenne autour de 13 ou 14 personnes, pour moitié des techniciens et pour l'autre moitié, des ingénieurs

d'étude et de recherche. Les effectifs globaux sont stables, il faut mentionner cependant entre le 1er septembre 2008, et début 2012, cinq mouvements de personnes au sein de l'équipe exploitation, et autant dans l'autre partie de l'équipe.

Le département est composé de 13 titulaires. En 2010 et 2011, il y a eu également 2 personnes en CDD, pour un contrat moyen de 12 mois.

Effectifs DS2I



d Les formations

Formations pour les personnels :

La mise à jour des connaissances est indispensable pour maintenir un service de qualité et proposer de nouvelles solutions à nos utilisateurs. En soutien de la veille technologique personnelle, des formations sont proposées, en fonction des besoins individuels et des projets.

Ces formations sont réalisées souvent par le CINES, parfois par des prestataires extérieurs.

Le département a bénéficié et produit 50 jours/homme de formation dans les 12 derniers mois.

Formations externes :

Le département apporte sa collaboration à des formations hébergées au CINES, telles CIREN (CINES/RENATER), en prenant en charge le plateau technique de formations. On compte en moyenne 8 formations par an.

Administration des systèmes et des données

a Le périmètre

Le département assure l'administration des systèmes informatiques, excepté les serveurs de calcul. Il fournit les services fondamentaux sur lesquels s'appuient les autres départements.

On dénombre approximativement 70 serveurs de taille très variable, dont la majorité hébergent un système d'exploitation Linux. La virtualisation des serveurs a été amorcée depuis 2006 ; elle est menée de manière plus systématique depuis 2011. A la fin de l'année 2012, plus de la moitié des serveurs seront soutenus par une architecture virtualisée.

Les techniques de virtualisation, et de Haute Disponibilité sont utilisées afin de réduire les indisponibilités des services.

b Les serveurs d'archivage

Une infrastructure dédiée a été mise en place pour appuyer le service d'archivage pérenne, dont le Département Archivage et Diffusion a la responsabilité.

Les données archivées sont copiées en deux exemplaires sur disques, et sécurisées sur bandes magnétiques. En 2013, il est prévu de compléter ce dispositif d'une copie délocalisée, afin de répondre aux exigences de sécurité et de pérennité, en

conformité avec les normes d'archivage et les recommandations du SIAF.

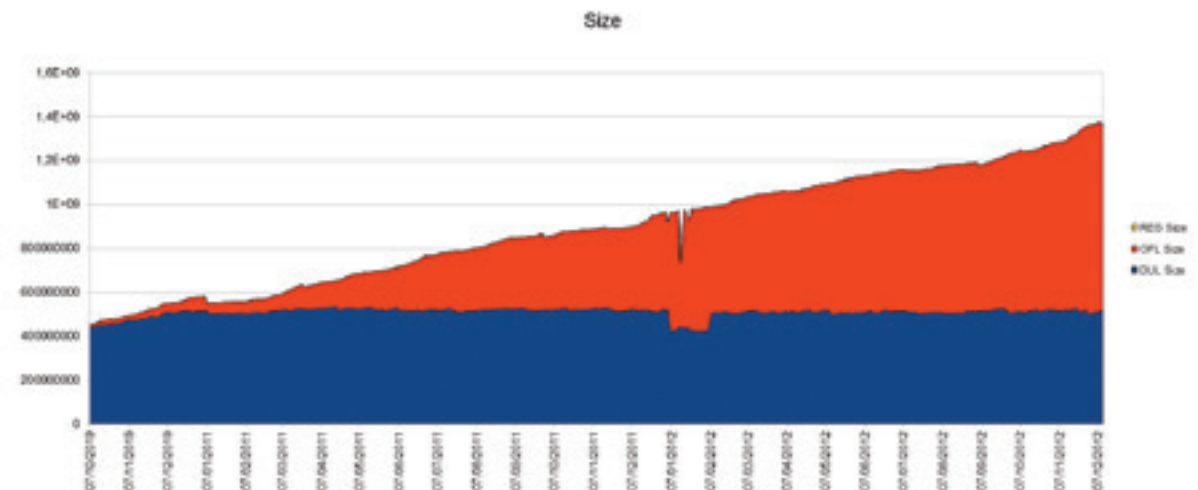
c Les données scientifiques

Les données scientifiques immédiates, nécessaires aux calculs ou produites par les calculateurs sont stockées sur les espaces disques locaux des calculateurs. Ces espaces très haute performance, ne proposent pas de service de sauvegarde.

Le stockage sécurisé et durable des données scientifiques est organisé autour d'un service commun de fichiers. Les données sont stockées sur 500 To de disques et sécurisées sur cartouches en deux exemplaires.

Fin 2010, le département a pris en charge le renouvellement des robotiques qui hébergent ces données. Les lecteurs qui ont été choisis pour leur performance et leur fiabilité, assurent la restitution des données dans les meilleures conditions. A cette période, la migration des 450 To (900 To pour les 2 copies) de données a été réalisée en 1,5 mois.

Le graphique suivant présente l'évolution des données scientifiques sur le serveur de fichiers entre fin 2010 et fin 2012, en volume et en nombre de fichiers.



La partie bleue du graphique représente la partie des données présentes sur disques.
La partie rouge représente les données stockées sur cartouches.



La prochaine évolution du serveur de fichiers est prévue pour fin 2012. Elle porte sur l'amélioration des débits entre le service de fichiers et les plateformes de calcul et de visualisation, afin de fluidifier les transferts de données de et vers les calculateurs. Le réseau actuel, 4x10Gb/s Ethernet, combiné au protocole NFS, ne suffit plus aux besoins exprimés par les utilisateurs. Cela devient un frein à la sécurisation des données.

Cette évolution majeure devrait se traduire par la mise en place d'un service « data-centrique » sur un réseau InfiniBand.

d Convergence : partage, mobilité, sécurisation et archivage des données scientifiques

Le CINES, comme les autres centres de calcul qui hébergent des volumes importants de données, est fortement sollicité pour étoffer ses services liés au traitement de ces informations.

La puissance de calcul destinée à la recherche en France a été multipliée par 80 depuis 5 ans, elle est passée de 20 Tflops en 2007 à 1,6 Pflops en 2012 (la machine Curie du CEA est 20% française et 80% européenne). On constate la même progression en Europe. Les données scientifiques qui sont produites, évoluent quant à elles de manière encore plus importante.

De nombreuses équipes scientifiques utilisent plusieurs calculateurs français Tier-1, et également un ou plusieurs calculateurs de niveau européen Tier-2. Cette activité répartie engendre des flux de

données importants et nécessite des capacités de stockage énormes.

Les projets de recherche tendent à augmenter la collaboration des équipes à travers l'Europe. Ceci impose de savoir répondre aux problématiques nouvelles : gérer des très grosses quantités de données, les rendre disponibles et mobiles rapidement, tout en assurant leur confidentialité. Le partage de ces informations amène également à trouver les références communes aux équipes pour en assurer la lisibilité : les métadonnées deviennent la référence commune.

En dernier lieu, l'archivage des données initiales ou intermédiaires présente pour la recherche un intérêt grandissant. Le CINES développe de nouveaux services pour répondre à ces demandes, et le département prend en charge la partie opérationnelle des infrastructures associées.

L'environnement « data-centrique » du serveur de fichiers est le point de convergence de ces services.



Les services réseaux et la sécurité

a Le réseau évolue

Le réseau interne, dont l'architecture a été mise en place il y a quelques années, est très stable.

Les deux principales améliorations ont consisté premièrement en la mise en place d'un réseau dédié haut débit pour les flux de données scientifiques. Ce réseau Ethernet est structuré autour d'un switch 10Gb/s.

Un autre réseau de données 10Gb/s Ethernet permet de fédérer les flux de données des serveurs virtualisés.

En 2011, dans le cadre du projet européen Prace, le CINES a mis en place, grâce à RENATER, un accès au réseau dédié 10Gb/s vers les partenaires du projet. Le réseau permet les échanges de données à haut débit entre les sites de stockage et les sites de calcul européens.

b Les services

Le département a la charge des services réseaux classiques : service de fichiers, service d'authentification, DNS, messagerie électronique, listes de diffusion, NTP.

En 2011, le service d'authentification basé sur OpenLdap a été totalement ré-architecturé. La quasi-totalité des serveurs et services sont dorénavant, sous le contrôle de LDAP. Cette opération lourde a été menée en application des principes issus du groupe projet ITIL (cf. 4.4).

Le service de mail a été totalement remanié en 2012 pour offrir de nouveaux services indispensables à la mobilité et au travail collaboratif (calendriers partagés, réservation de salles, synchronisation avec les smartphones,...). L'environnement Zimbra a été retenu et déployé.

La majorité des services ont été redéployés sur des serveurs et du stockage virtualisés.



c La sécurité

Le département auquel sont rattachés deux RSSI (Responsables Sécurité Systèmes d'Information), assure la sécurité des environnements informatiques du centre.

La gestion opérationnelle quotidienne est assurée par les spécialistes du département. Elle concerne les aspects classiques de la sécurité : veille concernant les failles potentielles, mises à jour régulières des composants logiciels à risque, surveillance fine des composants et des services. Des sessions d'information/formations sont proposées régulièrement.

Les rares incidents de sécurité donnent lieu à un rapport d'incident, à l'application immédiate de correctifs.

Les enjeux majeurs liés à la protection du patrimoine scientifique et technique nous amènent à engager deux actions prioritaires pour l'année 2013 :

✚ La mise en place d'un nouveau firewall pour accompagner dans les meilleures conditions les évolutions techniques liées à la mobilité des données dans un contexte de « Big Data »

✚ La définition d'une Politique de Sécurité des Systèmes d'Information en rapport avec ces nouveaux objectifs.

L'exploitation



a La supervision et le suivi des services

L'équipe exploitation a pour mission première de surveiller les services informatiques et d'intervenir en cas de problème. Elle dispose de plusieurs outils logiciels qui ont été choisis et paramétrés en fonction des besoins locaux.

En 2010, l'architecture de supervision a été complètement modernisée. La supervision de toutes les infrastructures informatiques et réseau est basée sur Nagios (www.nagios.org). C'est la version bundle « EyesOfNetwork » (EON) qui a été déployée.

La supervision est centralisée autour du serveur EON. Tous les serveurs, services, boîtiers réseaux, sont monitorés par ce serveur. La surveillance de Jade est déléguée à un serveur Nagios dédié qui est client du premier. La charge est ainsi répartie. Actuellement 15000 services sont suivis sur environ 3000 machines du centre.

Des scripts ont été développés afin de produire des indicateurs complexes adaptés à l'activité d'un serveur HPC. Ces indicateurs permettent de produire des taux de disponibilité de cette machine.

Le bundle EON offre également le reporting CACTI qui a été mis en place de manière homogène sur tous les serveurs et les boîtiers du réseau.

La surveillance des infrastructures électriques et de climatisation des salles machines, qui est habituellement réalisée au travers d'outils spécialisés type GTC, a été complétée par une série de capteurs environnementaux qui permettent un suivi de température et d'hygrométrie dans les salles machines.

b Le centre de services : SVP

Le département assure le niveau 1 du centre de services. Il est le premier interlocuteur des utilisateurs qui sollicitent le CINES. 5 personnes assurent à tour de rôle la permanence de 7h à 18h du lundi au jeudi.

Toutes les demandes qui sont adressées par mail, par téléphone, ou par fax sont enregistrées dans un outil de gestion d'incidents. Les tickets sont traités par SVP si le périmètre concerne la hotline. Sinon, ils sont orientés vers les équipes de spécialistes. Chaque incident est documenté et le dialogue avec l'utilisateur est tracé grâce au logiciel.

Le CINES dispose d'un centre de services depuis de nombreuses années, mais a modernisé les méthodes de travail et le logiciel de soutien de cette activité, en application des recommandations du groupe de travail ITIL (cf. 4.4)

Le logiciel choisi sur la base d'un dossier technique, et après tests, est le logiciel OTRS. (www.otrs.com). Il propose 6 processus certifiés ITIL.

Ce logiciel traite tout nouveau mail comme un ticket. Il a été mis en relation directe avec les outils de supervision. Toutes les alertes de supervision, habituellement signalées par mail, donnent lieu automatiquement à la création d'un ticket d'incident.

La mise en place de ce mécanisme a eu deux conséquences immédiates : la réduction nette de perte d'informations et de tickets et l'augmentation de la réactivité sur panne.

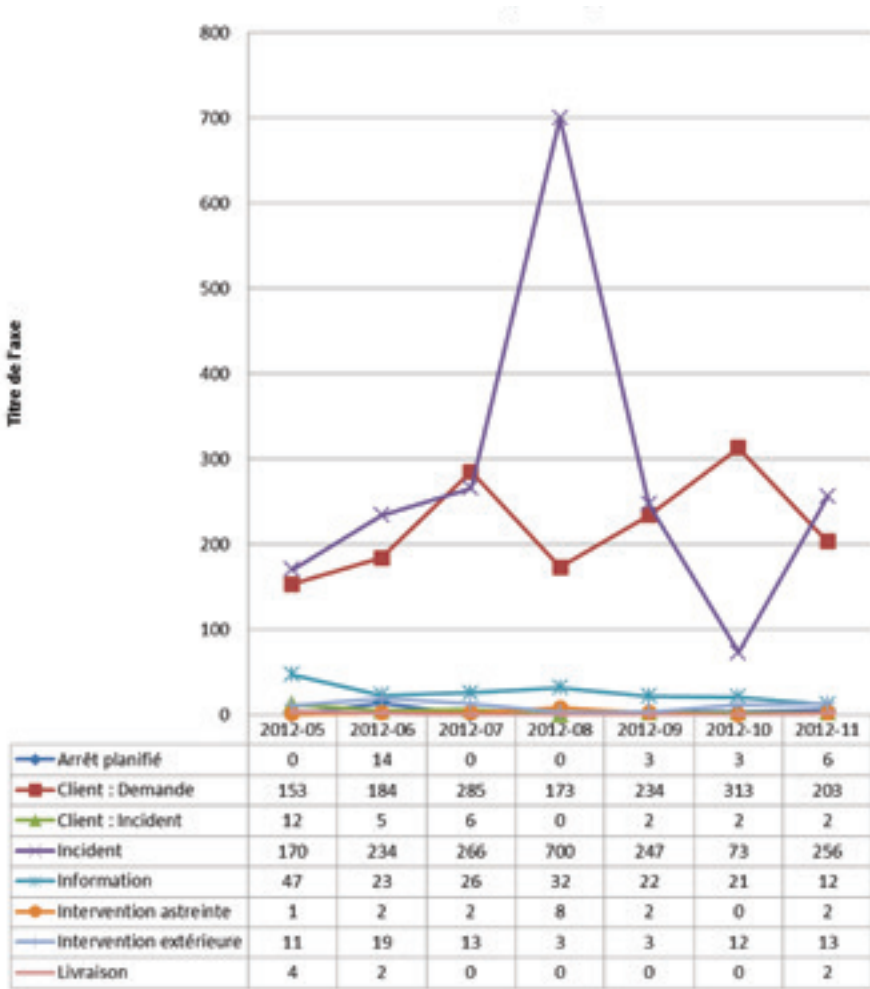
Voir schéma ci-contre : quelques indicateurs généraux relatifs à cette activité.

✚ Nombre d'incidents/mois= 278 en moyenne

✚ Nombre de demandes utilisateurs/mois=230 en moyenne

En 2013, ce service sera également accessible par une interface web et accompagné d'une FAQ.

Nombre de tickets par type



c L'astreinte

Le CINES offre un service 24h/24 365j/an. La surveillance des services est assurée en journée de 7h à 18h par l'équipe exploitation. La nuit et le week-end, ainsi que tous les jours fériés et pendant les vacances, une astreinte est organisée pour intervenir dans les meilleurs délais en cas de problème. Les personnels d'astreinte sont amenés à prendre en charge des problèmes relatifs aux infrastructures électriques ou climatisation, ainsi que des dysfonctionnements informatiques.

Les modalités d'intervention de l'astreinte ont été revues en 2011 afin d'optimiser les conditions d'astreinte, et la présence sur site des personnels. Depuis début 2012, tous les personnels du département participent à l'astreinte. La réorganisation et de l'astreinte et l'automatisation poussée de la supervision, ont permis de dégager un ETP (Equivalent Temps Plein) ingénieur.

Quelques chiffres :

- ✚ Nombre d'appels d'astreinte en 2012 : 1 par semaine en moyenne
- ✚ Nombre d'interventions d'astreinte en 2012 (11 mois) : 27
- ✚ Nombre d'interventions informatiques : 3/27
- ✚ Nombre d'interventions infrastructures (électricité, climatisation et autre) : 24/27

Chaque intervention fait l'objet d'un rapport d'incident. Le suivi et la correction sont assurés par les équipes responsables. Les pannes majeures donnent lieu à une ou plusieurs réunion(s) spécifiques afin d'améliorer le contexte global.

d ITIL

Le CINES a entrepris une réflexion basée sur les bonnes pratiques ITIL V2 il y a quelques années. L'objectif visé était d'améliorer le service rendu aux utilisateurs. Un groupe de travail inter-départements, qui compte 8 personnes, a été mis en place et est piloté par un chef de projet issu du département.

Une formation « ITIL Foundations » a été proposée et suivie par une forte proportion des personnels. La première étape a consisté en la définition de la CMDB (Configuration Management Data Base), et en la description de tous les processus de création, de modification ou de suppression qui concernent cette base : la Gestion des Configurations. La seconde étape fut la description et la mise en place de la Gestion des Incidents.

Cette étude a conduit à la description d'un référentiel unique autour duquel sont construits tous les processus de gestion métiers.

L'approche globale a conduit à ré-architecturer les composantes logicielles autour de la CMDB.

Les composantes principales de cette organisation :

- ✚ LDAP : héberge une partie de la CMDB (les utilisateurs), référence l'usage des ressources.
- ✚ OTRS : le logiciel de traitement des incidents et des demandes utilisateurs. Il contient également l'autre partie de la CMDB.
- ✚ Nagios : surveille tous les éléments de la CMDB (serveurs, services, réseaux,...) en fonction des niveaux de service. Il génère les alertes qui seront automatiquement traitées comme incidents par OTRS.
- ✚ BONITA est utilisé pour la mise en place opérationnelle des processus. (description des processus et des circuits de validation)

e L'hébergement pour des tiers

L'activité d'hébergement se situe dans une volonté de mutualiser les investissements lourds qui sont réalisés pour les environnements scientifiques. Le CINES a débuté cette activité dès 2002 en hébergeant le nœud régional (NR) de RENATER, ainsi que les équipements réseau de la Région Languedoc-Roussillon et ceux du réseau métropolitain Haut Débit Montpellier Nord (HDMON).

Depuis 2008, cette activité a été confirmée par le ministère de tutelle du CINES. Plusieurs autres organismes dépendant du Ministère de l'Enseignement et de la Recherche ont signé un contrat d'hébergement avec le CINES.

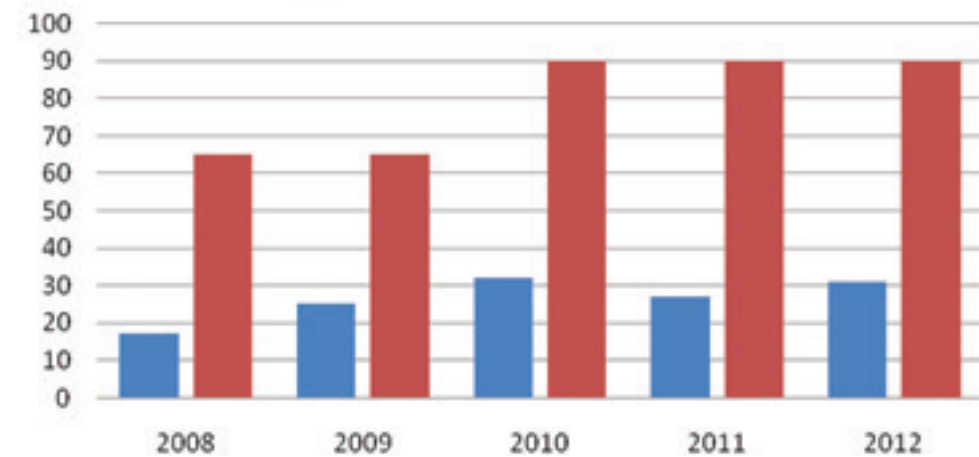
Le CINES assure la fourniture d'électricité et de

climatisation dans une salle machine sécurisée. Le client prend en charge l'administration de ses environnements informatiques. Les accès sont contrôlés. Un comité de liaison assure le suivi technique et contractuel. Un rapport trimestriel d'exploitation est fourni à chaque client.

f L'activité en chiffres

En 2012, 8 organismes ont pu bénéficier de l'hébergement dans les locaux du CINES. Cette activité devrait s'étendre dans les prochaines années, en permettant de mutualiser les équipements existants et en s'appuyant sur les infrastructures importantes déployées pour le calcul haute performance.

Hébergement nombre de baies



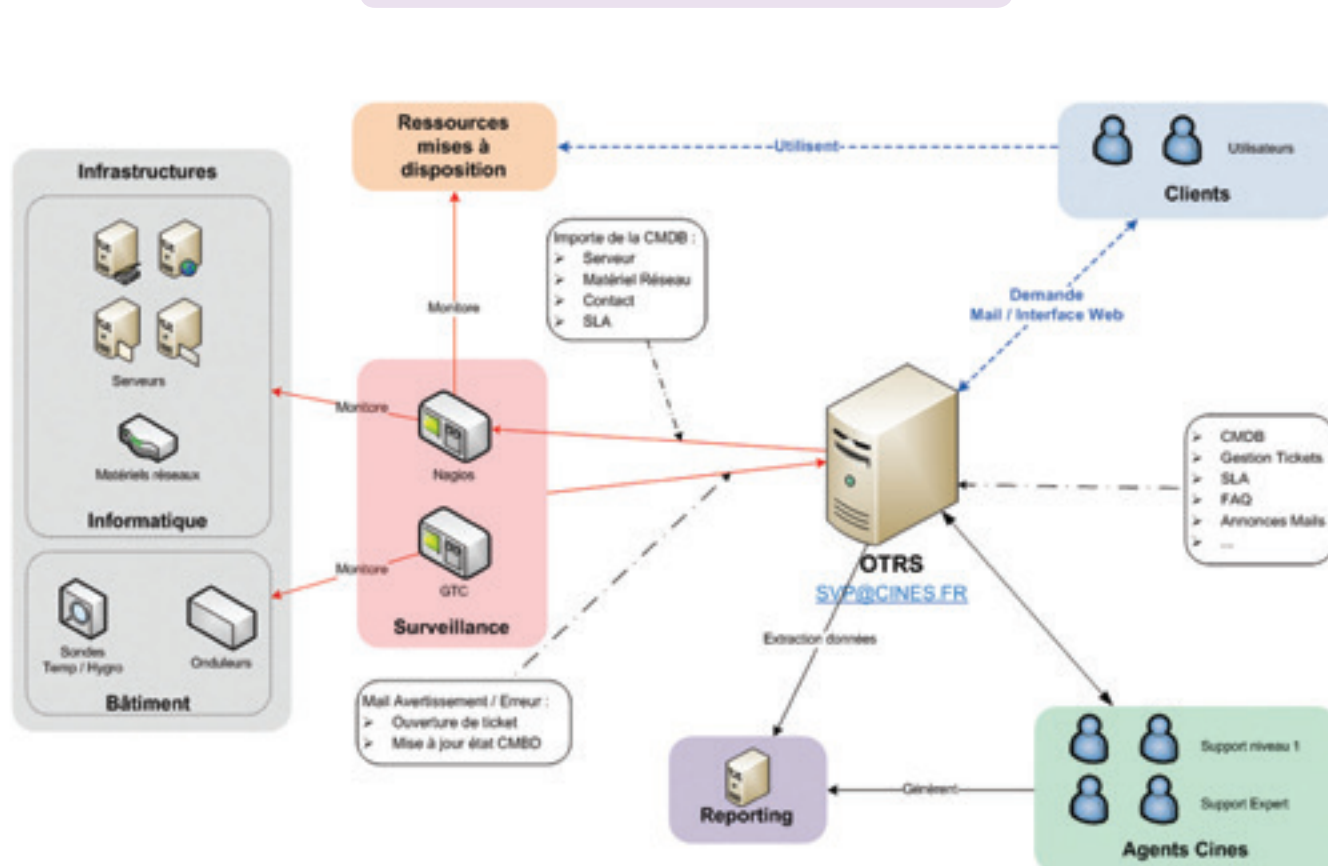
Bilan et Perspectives

Le mouvement profond qui a permis le renouvellement et l'amélioration générale des services proposés en interne ou aux utilisateurs du centre n'aurait pas été possible sans une équipe compétente, motivée et solidaire.

Pour les années qui viennent, l'équipe devra relever de nouveaux défis parmi lesquels on

peut citer : la montée en puissance du service de données, convergence du calcul scientifique et de l'archivage, les deux missions du CINES ; et un focus particulier sur la SMSI (Système de Management de Sécurité Informatique), afin de garantir la sécurité du patrimoine scientifique dont le centre est dépositaire.

ITIL : Gestion des événements au CINES



Système de données

Exploitation

Réseau, hébergement et Infrastructure

LES SERVICES D'ARCHIVAGE PERENNE DU CINES

42

Les nouveaux utilisateurs...	44
Archive ouverte HAL	44
TGE Adonis (projet pilote avec le CRDO Paris et le CRDO Aix	44
Bibliothèque de l'université Pierre...	44
Bibliothèque CUJAS	44
École française d'Extrême Orient (EFEO)	44
Bibliothèque Interuniversitaire de Santé	45
Bibliothèque Sainte-Geneviève (BSG)	45

LES CHIFFRES CLÉS

45

Évolution de la volumétrie	45
Répartition des objets archivés...	47
Facile_ outil de validation du format...	48

LA DÉMARCHE QUALITÉ

50

En route vers la certification	50
Formalisation de l'activité...	51
En route vers la certification	52

PARTAGE DE ...

53

Mise à disposition de la documentation	53
Participation à divers groupes ...	54
Publications	54
Formations	54

PERSPECTIVES...

55

L'archivage intermédiaire...	55
L'implication dans les projets européens	55



Le service d'archivage pérenne du CINES

Au cours des années 2000, les missions historiques du CINES ont évolué pour s'étendre au domaine de l'archivage numérique pérenne. Le ministère de l'Enseignement supérieur et de la Recherche a confié au CINES la mise en place d'un service d'archivage à long terme des données sous forme numérique destiné à la préservation du patrimoine scientifique. Cette problématique étant au cœur des préoccupations actuelles, plusieurs décisions ont depuis renforcé cette mission :

✚ Arrêté du 7 août 2006 relatif aux modalités de dépôt, de signalement, de reproduction, de diffusion et de conservation des thèses ou des travaux présentés en soutenance en vue d'un doctorat ; qui désigne le CINES comme centre de conservation des thèses électroniques françaises.

✚ Convention du 2 mai 2007 (faisant suite à celle du 15 octobre 2003) relative à la mise en ligne et l'archivage pérenne des données numérisées dans le cadre du programme Persée.

✚ Lettre de cadrage du 12 février 2008 recentrant les activités du CINES autour de deux missions stratégiques : le calcul intensif et l'archivage pérenne.

Une équipe de douze personnes aux profils divers et complémentaires travaillent dans ce service : un chef de projet, une archiviste, des experts en formats de fichiers, des développeurs informatiques, des administrateurs des systèmes et infrastructures, des techniciens hautement qualifiés (pour une surveillance 24h/24 et 7j/7), etc. Cette solution est proposée à l'ensemble de la communauté de l'enseignement supérieur et de la recherche dès lors que le corpus de documents à archiver est reconnu d'intérêt national.

Après trois années de conception et de développement, une première version du système PAC - Plateforme d'Archivage au CINES (PAC v1.0) - a été mise en service au printemps 2007 avec comme axe initial l'archivage des thèses électroniques (transmises par l'ABES). Toutefois, l'infrastructure matérielle ne permettant pas de gérer le volume de données prévu pour de nouveaux projets d'archives (revues numérisées en sciences humaines et sociales du programme PERSEE), un appel d'offres a été lancé pour l'acquisition d'une plateforme capable de gérer de larges volumes (plus de 40 Téraoctets utiles).

C'est donc une deuxième version du système PAC (PAC v2.0) qui est exploitée depuis le printemps 2008. Elle privilégie toujours la même approche généraliste, évitant ainsi le traitement des projets

d'archives au cas par cas, et permet de mutualiser la plateforme pour tous les projets d'archives. Elle respecte les normes et pratiques du domaine comme l'OAIS, le standard d'échange de données pour l'archivage (SEDA) ou le Dublin Core.

Le système PAC a été conçu comme un ensemble de trois serveurs logiques, s'inspirant du modèle proposé par la norme ISO 14 721 (OAIS – Open Archival Information System) :

✚ un serveur de versement auquel le producteur des documents pourra transmettre ses archives ;

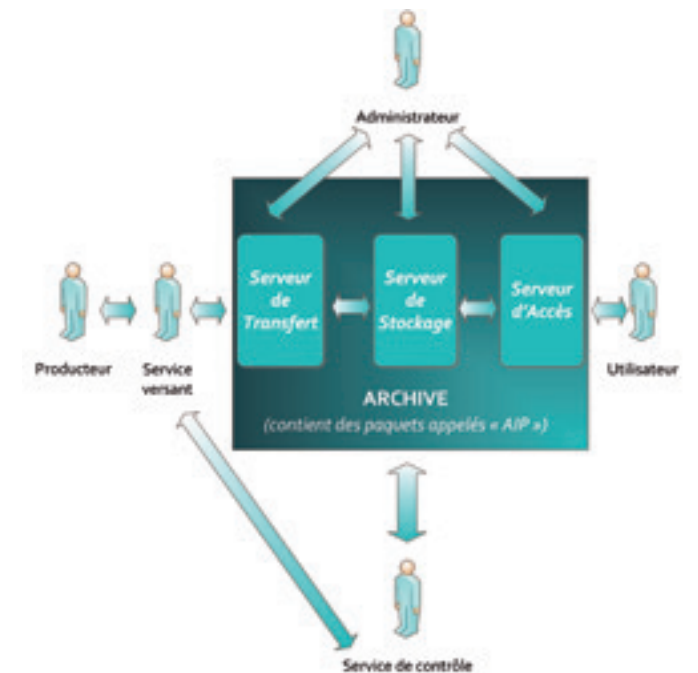
✚ un serveur de stockage où sont conservés les documents ;

✚ un serveur d'accès au travers duquel les utilisateurs autorisés (producteurs par exemple) peuvent consulter les archives et en obtenir une copie.

Le serveur de transfert réalise des opérations ayant attrait au cœur de métier de l'archivage électronique. En effet, il fournit notamment une assurance qualité sur les archives (contrôle de la conformité des métadonnées, contrôle des formats de fichiers, récupération automatique des schémas et DTD qui définissent les métadonnées métier, etc.). C'est la raison pour laquelle le développement de cette brique logicielle de PAC (appelée « module d'entrée » ou « module de versement ») est effectuée par les équipes du CINES.

Ce serveur logique (développé en Java et utilisant une base de données MySQL) est couplé d'une part au progiciel Arcsys (de la société Infotel) et d'autre part au module ArCheck. Arcsys gère le stockage des fichiers et fournit une interface d'accès basique aux données. Le module ArCheck (développé en interne) effectue pour sa part des opérations métier complémentaires en vérifiant la qualité de toutes les archives au moins une fois par an.

D'un point de vue matériel, la plateforme actuelle utilise des serveurs d'application SUN X4150, des baies de stockage SUN Storagetek ST6140-4G (en



double exemplaire sur les supports disques) et des bandothèques IBM 3584 (un exemplaire des archives dans chaque robothèque).

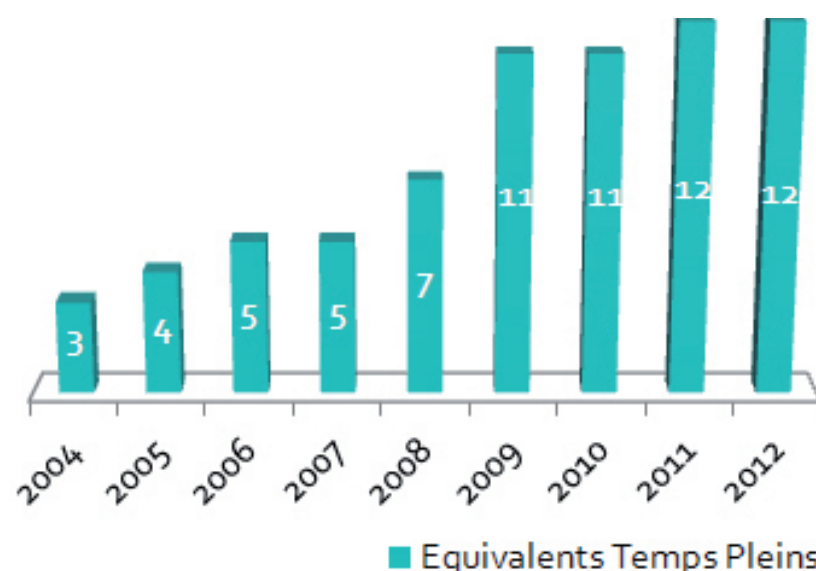


Les premiers documents archivés dans la plateforme furent les thèses électroniques délivrées en France et collectées par l'ABES – Agence Bibliographique de l'Enseignement Supérieur – auprès des bibliothèques universitaires par l'intermédiaire de l'outil STAR. Depuis 2007, plus de 8000 thèses ont été archivées dans la plateforme.



Les revues en sciences humaines et sociales numérisées dans le cadre du programme PERSEE ont constitué le deuxième projet d'archivage du CINES. Les premières revues ont été archivées dès 2008. La volumétrie de ce projet est relativement considérable puisqu'elle atteint 10,5 To au début de l'année 2012.

Évolution du nombre de collaborateurs



Les nouveaux utilisateurs du service PAC

2010 et 2011 ont été des années charnières pour le service PAC. De nombreux projets ont débuté en 2010 et sont passés en phase d'archivage en 2011. Aussi, en deux ans, le nombre d'établissements archivant au CINES (« services versants ») a plus que quadruplé.

a Archive ouverte HAL (CCSD)



Après 6 mois de mise en œuvre, les premiers documents de l'archive ouverte HAL (Hyper Article en Ligne) gérée par le CCSD (Centre pour la Communication Scientifique Directe) ont été archivés dans PAC le 29 mars 2010. Il s'agit d'articles scientifiques pluridisciplinaires de niveau recherche ainsi que de thèses françaises ou étrangères disponibles en accès libre sur le portail HAL.

c Bibliothèque de l'Université Pierre et Marie Curie (BUPMC)



La Bibliothèque Universitaire Pierre et Marie Curie est entrée en production le 1er février 2011 pour archiver durablement ses ouvrages numérisés sur les neurosciences, la biologie, la géologie et la physique-chimie.

d Bibliothèque Cujas



La bibliothèque inter universitaire de droit CUJAS, en tant que pôle de référence, couvre l'ensemble des sciences juridiques. Elle archive au CINES depuis fin février 2011 les ouvrages fondamentaux en droit qu'elle numérise ainsi que certains documents pédagogiques.

e École Française d'Extrême Orient (EFEO)



Depuis le 15 mars 2011, l'École française d'Extrême-Orient archive dans PAC un important fonds numérisé de photographies, plaques de

verre et plaques stéréoscopiques, etc. prises depuis le 19ème siècle dans le cadre de missions patrimoniales en Extrême-Orient. Environ 10 000 images sur près de 85 000 ont déjà été archivées.

f Bibliothèque InterUniversitaire de Santé (BIUSanté)



La Bibliothèque Inter Universitaire Santé (BIUSanté) possède un fonds d'ouvrages spécialisés sur l'histoire de la médecine et de l'art dentaire. Elle a entrepris depuis quelques années la numérisation de ses collections, mises à disposition dans la bibliothèque numérique Medic@, et

dispose également d'une banque d'images et de revues spécialisées numérisées. L'ensemble de ces documents sont archivés au CINES depuis juillet 2011 et représentent déjà près de 3 To de données (au 31 décembre 2011).

g Bibliothèque Sainte Geneviève (BSG)

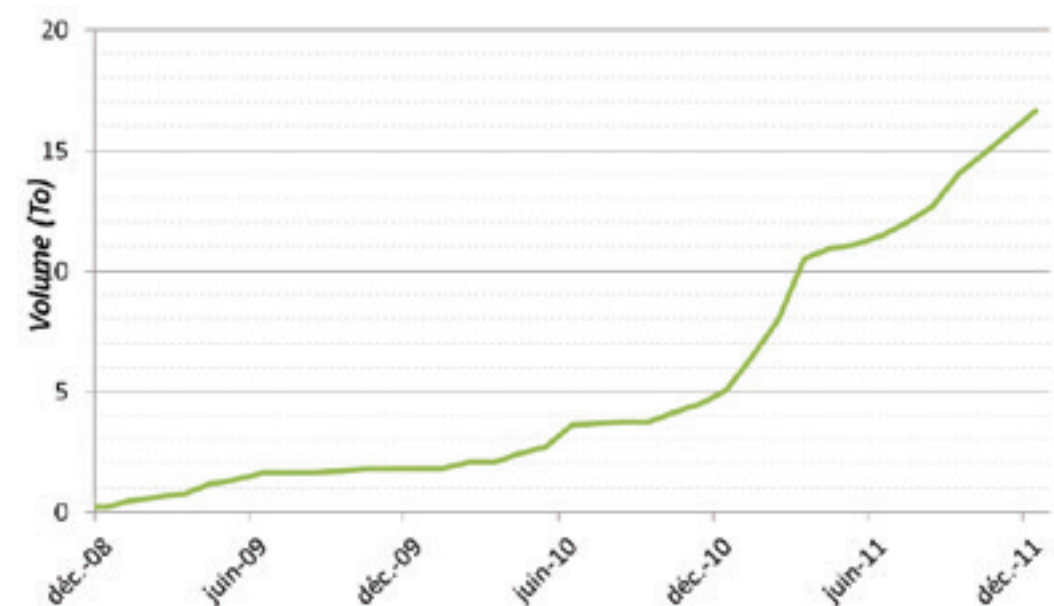


La Bibliothèque inter universitaire Sainte Geneviève a entrepris la numérisation de ses collections d'ouvrages anciens et les archive dans PAC depuis août 2011. A terme, cela concerne près de 3 To de données.

Les chiffres clés

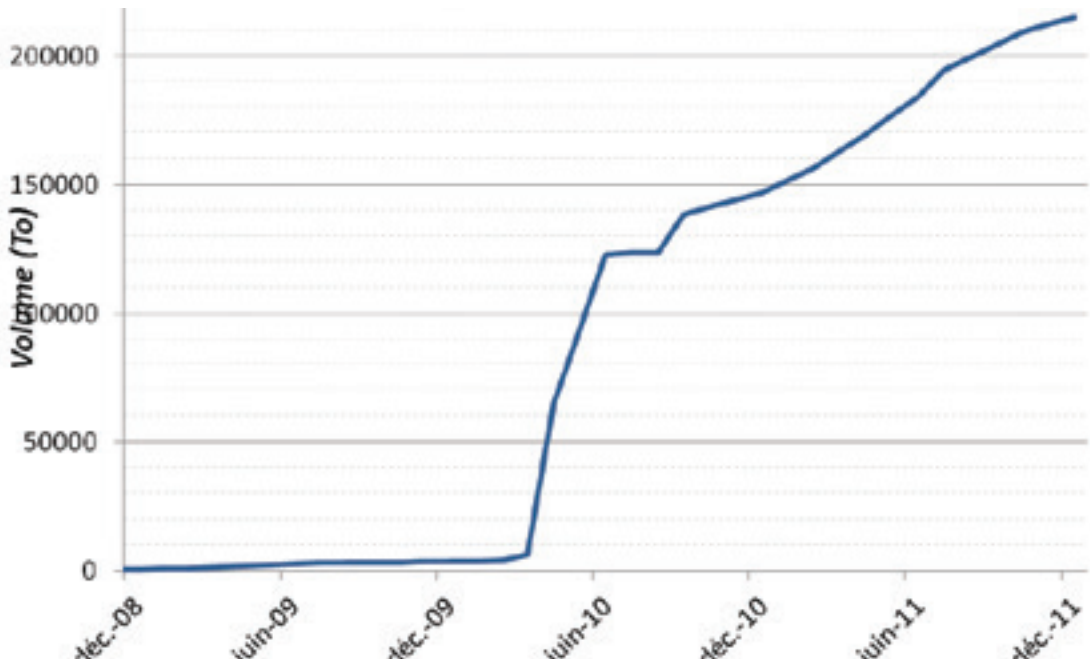
a Évolution de la volumétrie

Évolution de la volumétrie



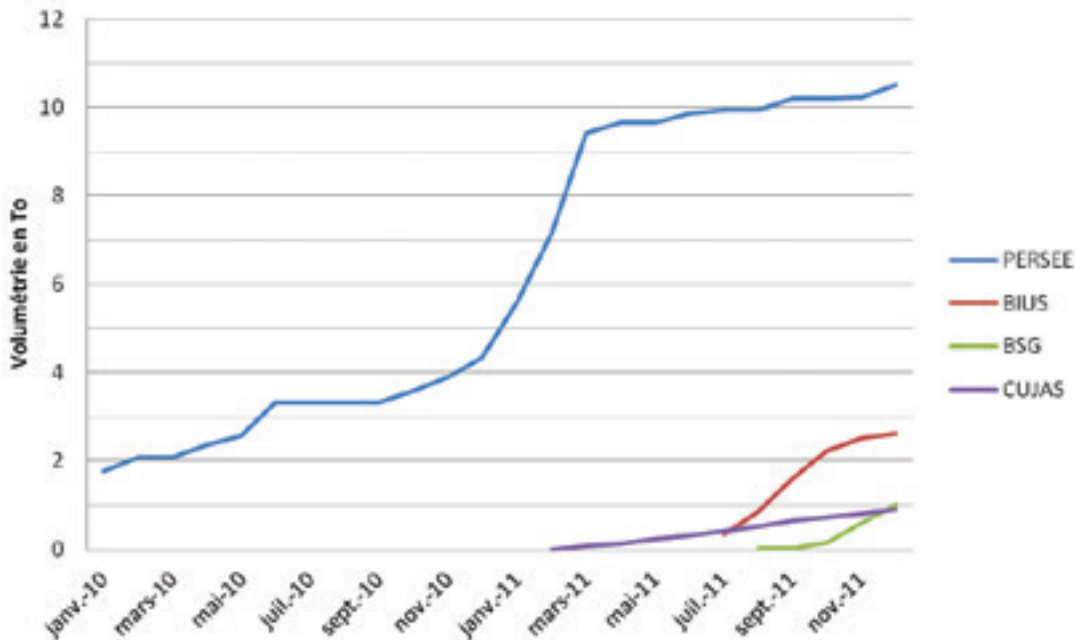
L'augmentation plus importante de la volumétrie à partir de 2010 provient de l'archivage des revues numérisées en sciences humaines et sociales de PERSEE ainsi que celui de plusieurs bibliothèques universitaires numérisées ayant commencé à archiver entre 2010 et 2011.

Évolution du nombre total de documents archivés

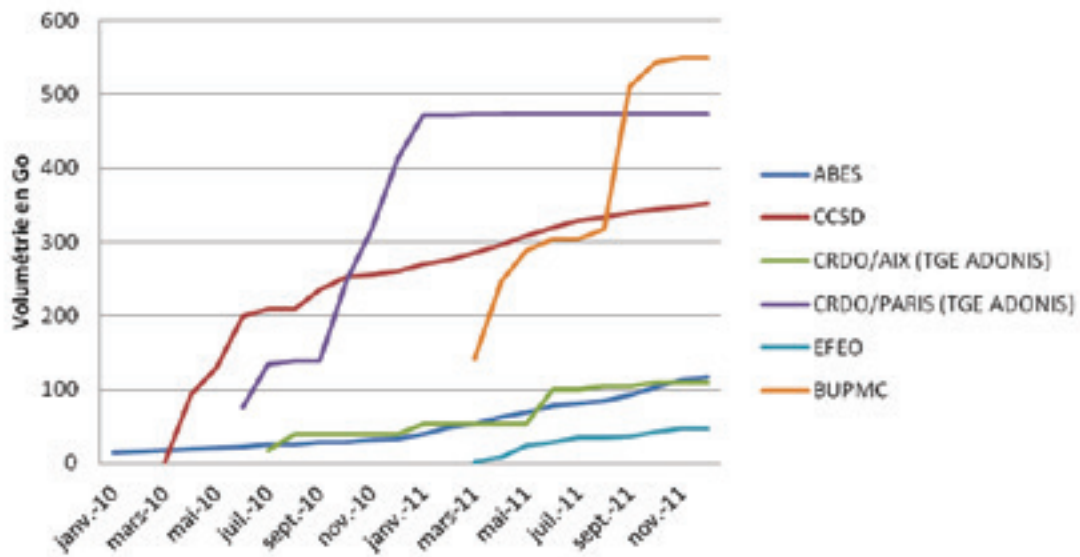


Le pic de documents versés en 2010 s'explique principalement par l'archivage des articles scientifiques du portail HAL.

Évolution de la volumétrie par service versant



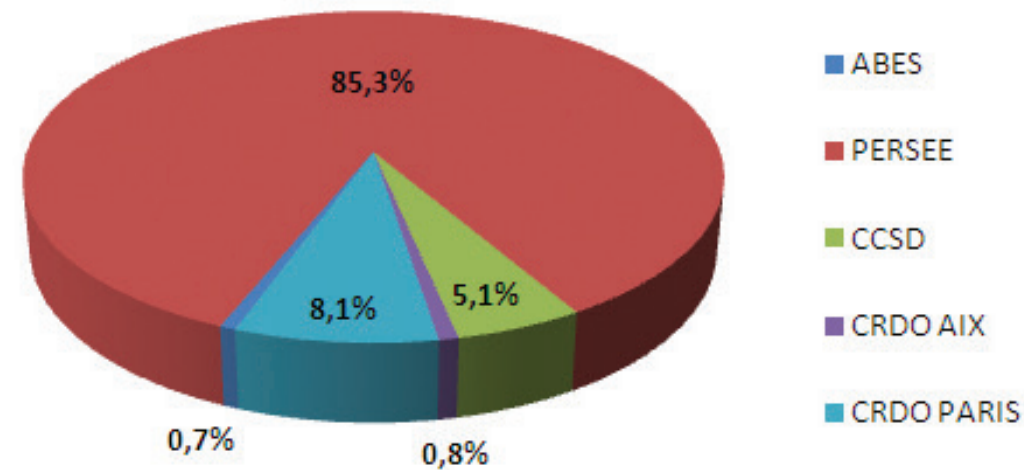
Évolution de la volumétrie par service versant



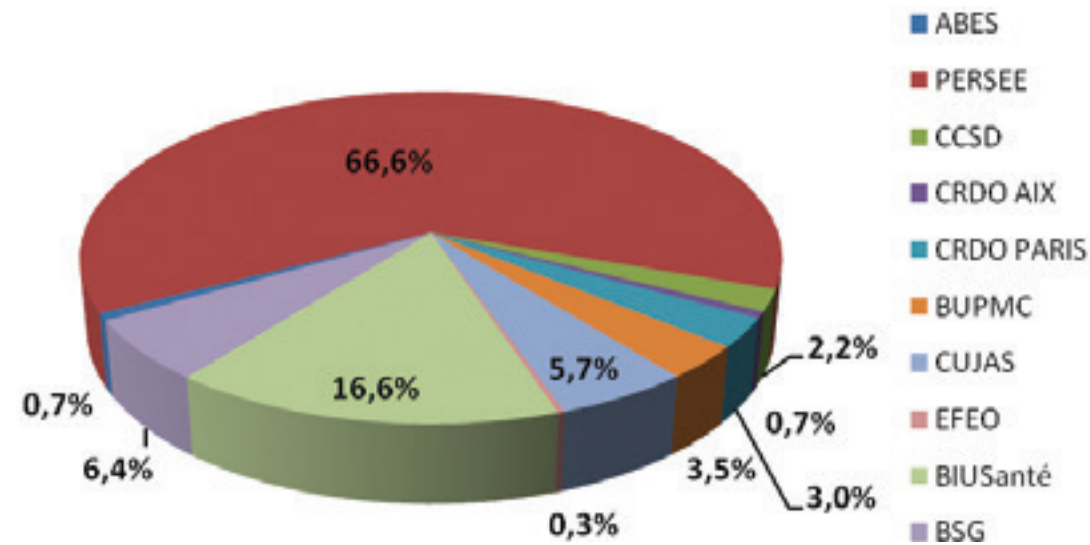
b Répartition des objets archivés en 2010 et 2011

Services versants	Nombre d'objets archivés		Volume archivé (en Go)		Nombre total objets archivés ²	Volume total archivé (en Go) ¹
	2010	2011	2010	2011		
ABES (thèses)	+ 1 665	+ 4 896	+ 20	+ 84	7 807	117
Persée (revues SHS numérisées)	+ 2 585	+ 7 275	+ 1 729	+ 6 190	12 165	10 513
CCSD (archive ouverte HAL)	+ 136 096	+ 34 555	+ 261	+ 91	170 651	352
CRDO AIX (données orales)	+ 143	+ 211	+ 40	+ 69	354	109
CRDO PARIS (données orales)	+ 2 859	+ 767	+ 412	+ 62	3 626	474
BUPMC (ouvrages numérisés)	/	+ 1 148	/	+ 550	1 148	550
CUJAS (ouvrages numérisés)	/	+ 135	/	+ 900	135	900
EFEO (images numérisées)	/	+ 10 043	/	+ 47	10 043	47
BIUSanté (ouvrages numérisés)	/	+ 9 027	/	+ 2 622	9 027	2 622
BSG (ouvrages numérisés)	/	+ 121	/	+ 1 008	121	1 008
TOTAL	+ 143 348	+ 68 178	+ 2 462	+ 11 623	215 077	16 692

Évolution de la volumétrie par service versant (04/01/2011 - total : 5,1 To)



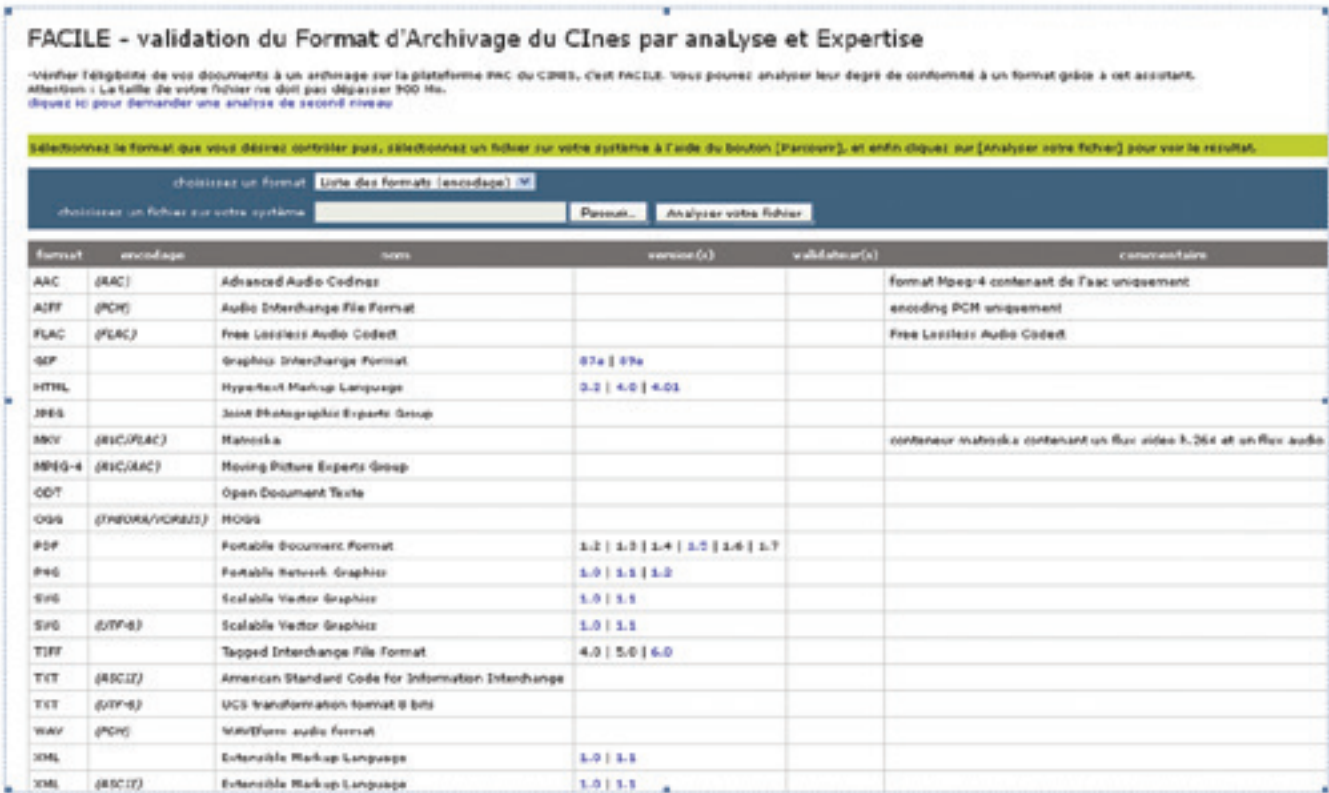
Évolution de la volumétrie par service versant (04/01/2011 - total : 16,9 To)



C FACILE – outil de validation du Format d’Archivage du CInes par anaLyse et Expertise

Développé par le CINES et accessible librement sur internet, FACILE permet de valider la qualité du format des fichiers à archiver. Il intègre les mêmes outils et la même liste de formats supportés par la plateforme d’archivage du CINES. Les requêtes FACILE correspondent au nombre d’analyses de fichiers faites via l’interface disponible sur : <http://facile.cines.fr/>

Lorsque les utilisateurs disposent d’un fichier qui n’est pas dans un format archivable (i.e. l’outil FACILE ne valide pas le format du fichier), ils peuvent bénéficier d’une assistance « expertise formats ». Ce service les aide à convertir le(s) fichier(s) dans un format archivable. On parle de « demande d’assistance FACILE niveau 2 ». Les demandes d’assistance de ce genre ont plus que triplé entre 2010 et 2011.



Lorsque les utilisateurs disposent d’un fichier qui n’est pas dans un format archivable (i.e. l’outil FACILE ne valide pas le format du fichier), ils peuvent bénéficier d’une assistance « expertise formats ». Ce service les aide à convertir le(s) fichier(s) dans un format archivable. On parle de « demande d’assistance FACILE niveau 2 ». Les demandes d’assistance de ce genre ont plus que triplé entre 2010 et 2011.

	Nombre de requêtes FACILE	Nombre d'appels FACILE niveau 2
TOTAL 2010	136 084 analyses de fichiers	81 demandes d'assistance
TOTAL 2011	146 715 analyses de fichiers	250 demandes d'assistance

La démarche qualité

a En route vers la certification...

Pour attester de la qualité de ses services, le CINES s’est lancé depuis trois ans dans une démarche de certification de son service d’archivage pérenne. Cette démarche passe d’abord par le choix de la certification la plus pertinente dans le cadre de son activité, sachant que le domaine est assez récent et que les réflexions sont encore en cours.

Un premier audit avait été réalisé par les consultants de la société Ourouk en 2009 afin de faire un bilan des points forts et des points faibles du service. Un certain nombre de préconisations ont ainsi été formulées puis intégrées dans le calendrier des activités 2010 sous la forme de plans d’actions.

Audit 2009

Audit 2010-2011

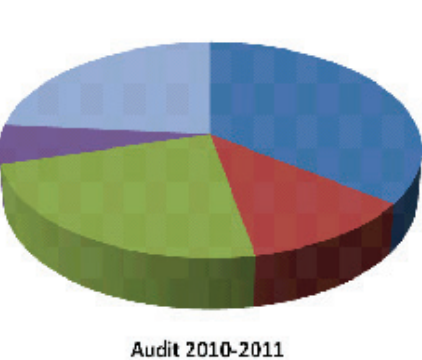
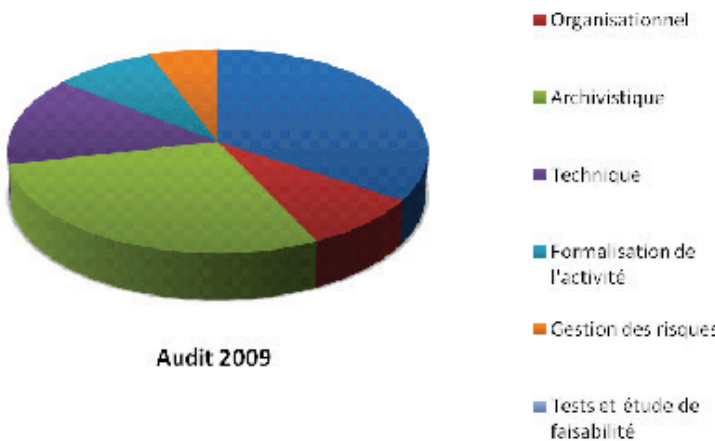


Une des préconisations de l’audit de 2009 était de s’impliquer davantage dans la rédaction des normes. C’est chose faite avec la participation du CINES au projet européen APARSEN qui consiste à tester la future norme ISO 16 363 – référentiel de certification de services d’archivage électronique. Six services d’archivage (trois en Europe et trois en Amérique du Nord), dont le CINES, ont été sollicités afin d’évaluer le projet de norme. L’évaluation du référentiel d’audit s’est déroulée en deux temps : un audit interne, qui a mobilisé trois auditeurs (charge de travail de soixante

Jours/Homme), et un audit externe, basé sur le rapport d’audit interne et réalisé par une équipe de spécialistes internationaux. Ces audits ont tout d’abord permis de faire un premier bilan des améliorations apportées suite aux préconisations du cabinet Ourouk en 2009. Ils ont également mis en lumière les points à conforter, principalement dans les domaines stratégiques et financiers, archivistiques, organisationnels, techniques et au niveau des phases de tests. Des axes stratégiques ont été définis et les plans d’actions correspondants ont été insérés dans le calendrier des activités 2011.

Audit 2009

Audit 2010-2011



Suite au projet pilote avec le Très Grand Équipement Adonis qui concerne des archives publiques et afin de se conformer à la réglementation en vigueur, le CINES a demandé un agrément auprès du Service Interministériel des Archives de France (SIAF) pour la conservation d’archives publiques courantes et intermédiaires. Cet agrément est soumis à l’examen d’un dossier décrivant l’ensemble des procédures mises en œuvre pour l’archivage et une visite sur site d’une dizaine d’experts nationaux. Il a été accordé au CINES en décembre 2010 pour une durée de trois ans renouvelable.



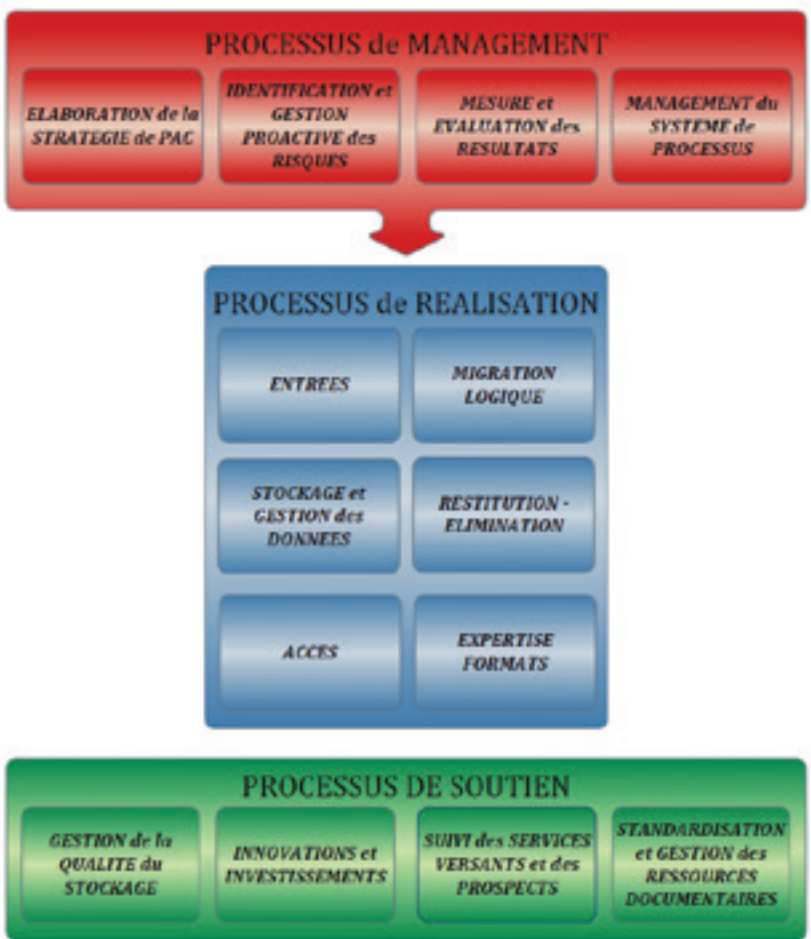
Le CINES a également reçu, en mars 2011, l’accréditation DSA, Data Seal of Approval, qui est attribuée aux centres de préservation numérique ayant mis en place des procédures d’assurance qualité afin de garantir l’accessibilité et l’intelligibilité des informations qui leur sont confiées.

b Formalisation de l’activité et amélioration de la qualité de la documentation

Dans un souci de professionnalisation et d’amélioration permanente du service proposé, le CINES a entrepris ces dernières années plusieurs démarches de formalisation de son

activité, notamment en améliorant la qualité de la documentation sur le service PAC. Après la réalisation en 2009 d’une étude des risques liés à l’activité d’archivage pérenne de PAC,

Cartographie générale des processus métiers de PAC



un plan de gestion de ces risques a été mis en place en 2010. Il fait l'objet d'une révision semestrielle impliquant toute l'équipe. Chaque risque identifié est réévalué et son plan d'actions est mis à jour. Pour chaque risque, un membre de l'équipe est chargé de coordonner les actions qui en limiteront l'impact.

Une démarche de formalisation des processus métiers liés à son activité d'archivage pérenne a également été entreprise au sein du service PAC en 2010. Un an plus tard, quatorze processus ont été identifiés, répartis en trois catégories : processus de management, de réalisation et de soutien. Chaque processus est composé de sous-processus, eux-mêmes composés d'activités ; ce qui représente plus de trois cents objets au total. Dans une volonté de transparence, toutes les cartographies sont accessibles librement sur le site web du CINES.

Le CINES a partagé son retour d'expérience sur cette question de la formalisation de processus métiers à de nombreuses reprises parmi lesquelles l'International Conference on Preservation of Digital Objects (IPRES 2010) qui s'est tenue à Vienne (Autriche) en septembre 2010 et les journées QuaRES (huitième école inter-organismes sur la qualité en recherche et en enseignement supérieur), également en septembre 2010.

De manière générale, afin de répondre aux préconisations des différents audits, l'ensemble de la documentation utilisée par le service a été améliorée :

- ✚ les étapes pour la mise en œuvre d'un projet d'archivage ont été listées et précisées sur un document fourni à tout nouveau service versant ;
- ✚ des modèles de documents ont été définis ;
- ✚ la convention encadrant les modalités de l'archivage au CINES avec les services versants a été revue d'une part pour être davantage en adéquation avec les termes juridiques de rigueur, et d'autre part pour prendre en compte les exigences de la norme française NF Z42-013. Elle a fait l'objet d'une relecture et d'une validation par une avocate spécialisée en droit de l'information et de la propriété intellectuelle ;
- ✚ etc.

La documentation associée aux objets archivés dans la plateforme a aussi été enrichie. De nouvelles métadonnées sur les documents archivés ont été créées pour prendre en compte des informations sur le contexte de l'archivage (structure des paquets archivés, manière de renseigner les métadonnées, etc.) et la description du fonds dans son ensemble, de son producteur et du service versant. Toutes ces informations sont regroupées sous la forme d'un fichier XML (ppdi.xml) par projet d'archives et conservé dans la plateforme d'archivage PAC.

C En route vers la certification...

Sur le plan technique, la plateforme d'archivage PAC a fait l'objet de plusieurs évolutions afin d'améliorer le service rendu et de répondre aux besoins des nouveaux services versants.

La plateforme PAC offre un niveau de sécurité élevé. Le plan de gestion des risques montre que les risques liés à la sécurité ont un degré d'exposition acceptable. Le CINES continue de travailler à l'amélioration de la sécurité notamment par l'intégration de la norme ISO 27001.

Dans le cadre du projet mené avec le Très Grand Équipement Adonis, le CRDO (Centre de Ressources de Données Orales), centre pilote, a demandé au CINES la prise en charge de formats de fichiers audio et vidéo supplémentaires dans PAC. Après validation par nos experts de la pérennité de ces formats, les modifications nécessaires ont été effectuées : recherche et intégration d'un contrôleur pour ces formats, tests, validation de l'ensemble du processus.

Afin de garantir une interopérabilité avec les services d'archives publics chargés de l'archivage définitif, un module d'export des objets archivés a été développé en 2010 conformément au Standard d'Échange de Données pour l'Archivage (développé par le Service Interministériel des Archives de France). Cet outil génère automatiquement des paquets d'archives avec un bordereau de métadonnées respectant les schémas XML du SEDA.

Par ailleurs, le CINES s'est doté d'un moyen supplémentaire pour garantir de manière certaine l'intégrité des documents archivés dans PAC. Le module ArCheck a été développé afin de renforcer les contrôles d'intégrité à tous les niveaux. Ceux-ci

sont paramétrables et automatisés. Les actions et leurs résultats sont journalisés pour être exploités dans des rapports notamment.

En 2011, d'autres améliorations ont été apportées au niveau des phases de tests (refonte complète du banc d'essais interne) et de la plateforme elle-même. Le module d'entrées a été retravaillé pour permettre une gestion plus fine des erreurs et des spécifications ont été rédigées pour le rendre conforme à 100% au Standard d'Échange de Données pour l'Archivage. Le module d'accès qui, en l'absence de besoin jusqu'à présent, n'avait pas été développé, a fait l'objet de spécifications fonctionnelles qui seront suivies d'une mise en

œuvre en 2012/2013. Enfin, la société Infotel, fournisseur de notre logiciel d'archivage, a réalisé à notre demande un audit technique afin d'optimiser Arcsys, l'outil que nous utilisons.

Un projet de réplication sur site distant des données archivées dans PAC est en cours de réalisation avec la Bibliothèque nationale de France (BnF) ; ce qui assurera au CINES une sauvegarde distante de plus de huit cents kilomètres. Le scénario retenu, actuellement en phase de test, prend la forme d'une réplication croisée : le CINES envoie des archives à la BnF pour être conservées sur ses machines, et la BnF fait de même en nous envoyant ses données. La possibilité d'une réplication vers l'IN2P3 est également à l'étude.

Partage de l'expérience et de l'expertise du CINES

avec les communautés (enseignement supérieur,

recherche, archivistique, services publics, etc.)

a Mise à disposition de la documentation

Dans une perspective d'échange et de partage de son expertise, le CINES met à disposition de tous une large partie de sa documentation de référence sur son site Internet.

C'est à ce titre-là notamment que la rubrique du site web consacrée à l'archivage numérique pérenne a été entièrement revue en 2010 : plus structurée, plus de contenu, plus illustrée, plus pédagogique. L'expertise et les services proposés par PAC y sont expliqués et détaillés en français et en anglais. L'article de Wikipedia consacré au CINES a également été mis à jour pour prendre en compte la mission d'archivage numérique pérenne. D'autre part, certains des logiciels développés par le CINES sont mis à disposition de la communauté comme l'outil de contrôle des formats de fichiers avant archivage (formatValidator). Ce module dont le code source est librement accessible (sur le site internet de l'Adullact, association qui promeut le logiciel libre : <http://adullact.org/>), a été intégré dans le logiciel d'archivage libre As@lae de l'Adullact.

L'application FACILE, développée également par le CINES et accessible sur son site web (<http://facile.cines.fr/>), permet de tester en ligne l'éligibilité d'un fichier à l'archivage dans PAC. Cet outil rencontre un vif succès auprès de la communauté, étudiants et bibliothèques universitaires essentiellement, qui contrôlent notamment grâce à cela la validité de leurs thèses au format PDF. Lorsqu'un fichier est déclaré non conforme par FACILE, l'utilisateur peut faire une demande d'assistance (FACILE 2) auprès du CINES pour rendre son fichier archivable. Ces demandes d'assistance ont quasiment doublé entre 2010 et 2011. Une thèse sur cinq qui est archivée dans PAC a préalablement fait l'objet d'une assistance via FACILE 2.

b Participation à divers groupes de réflexion et de travail

Étant un des premiers services d'archivage électronique opérationnel en France, le CINES partage son expertise au travers de groupes de réflexion et de travail ayant des approches diverses. Dans le domaine de l'information-documentation, il participe à la Commission Archivage Électronique de l'Association des Archivistes Français et à la Bibliothèque Scientifique Numérique (BSN) dans lequel il est responsable du segment « Préservation à long terme ». Il pilote le groupe PIN (Pérennisation de l'Information Numérique) de l'Association Aristote et collabore avec l'AMUE pour la diffusion de bonnes pratiques en archivage électronique au sein de la communauté enseignement supérieur et recherche. En 2011, deux membres de l'équipe PAC ont par ailleurs intégré le comité de pilotage sur l'évolution du Standard d'Échange de Données pour l'Archivage (SEDA), coordonné par le Service Interministériel des Archives de France (SIAF). Sur le plan purement technique, il appartient au club des utilisateurs d'Arcsys (progiciel d'archivage développé par la société Infotel, filiale de Sun-Oracle). Le CINES est également représenté au niveau international dans le DataSeal of Approval, organisme européen qui évalue la qualité de systèmes d'archivage électronique.

A des fins d'échanges d'expertise et d'expérience, des rencontres ont régulièrement lieu avec d'autres institutions sur la question de l'archivage électronique, comme la visite en 2010 de représentants d'EDF R&D dans le cadre de leur projet de réalisation d'une plateforme d'archivage électronique.

c Publications

En tant que service pionnier en France dans le domaine de l'archivage numérique pérenne, le CINES est sollicité pour la rédaction de publications de référence. Ainsi, suite à une commande du TGE-Adonis et du Service Interministériel des Archives de France, la cellule d'expertise formats du CINES a rédigé en 2010 un « Guide méthodologique pour le choix de formats numériques pérennes dans un contexte de données orales et visuelles ». Ce guide fournit une étude détaillée des spécificités des formats audio et

vidéo et propose une sélection de formats retenus comme pérennes au regard de critères prédéfinis. Mis à jour en 2011, il est accessible librement à l'adresse suivante : <http://www.archivesdefrance.culture.gouv.fr/static/4923>.

Une autre étude a été commandée au CINES en 2011 par le Service Interministériel des Archives de France (SIAF). Actuellement en cours, elle concerne les formats PDF et prévoit l'analyse des différentes versions du format PDF et le test de plusieurs logiciels de conversion et de validation de formats de fichier. Les résultats seront publiés courant 2012 au sein d'un guide méthodologique.

En 2010, le CINES a contribué à la rédaction du rapport sur le Schéma Numérique des Bibliothèques (SNB) en partageant notamment son expérience en matière de préservation du patrimoine numérique. Ce rapport, confié par Mme Christine Albanel, Ministre de la Culture à M. Bruno Racine, directeur de la Bibliothèque nationale de France, met en évidence l'intérêt d'une collaboration entre les ministères concernés (culture et communication, éducation Nationale, enseignement supérieur, affaires étrangères) et les collectivités territoriales pour l'adaptation des bibliothèques publiques aux nouveaux usages numériques. Il identifie également les partenariats public-privé possibles.

d Formations

Dans le cadre du knowledge management, plusieurs formations ont été effectuées en interne pour les membres de l'équipe d'une part sur l'approfondissement de concepts archivistiques dont le modèle conceptuel OAIS (norme ISO 14721), d'autre part sur les pratiques et les règles mises en place dans PAC : gestion des métadonnées, des relations entre documents, etc.

Le CINES assure également un certain nombre d'interventions dans des cursus de formation longue durée (ENSSIB), lors de formations professionnelles (pour le compte de l'Association des Archivistes Français ou du groupe PIN notamment) ou encore à l'occasion de séminaires ou journées d'études (IPRES, Journées AMUE, etc.).

Perspectives d'évolutions

a L'archivage intermédiaire des données scientifiques

A l'heure actuelle, les données scientifiques représentent un enjeu majeur pour la recherche. Or la production de ces données étant essentiellement numérique, leur conservation soulève plusieurs problèmes tant par leur degré de technicité que par la spécificité des formats de fichiers utilisés. De par sa double compétence en archivage électronique et en calcul intensif, le CINES a souhaité orienter ses réflexions sur la question de la conservation des données/archives scientifiques.

Dans cette perspective, il collabore depuis 2010 avec le CNES (Centre National d'Études Spatiales) sur le prototypage du framework BEST. Ce dernier permettra la description de données dans des formats propriétaires ou définis au cas par cas par les producteurs de données.

Par ailleurs, en travaillant avec les communautés enseignement supérieur et recherche, il a été constaté qu'il est souvent difficile de sélectionner en amont les données scientifiques qui doivent être archivées de façon définitive. Aussi le CINES travaille à la mise en place (2010-2012) d'un service « d'archivage intermédiaire de données scientifiques » (projet ISAAC) pour des communautés d'utilisateurs intéressées. Ce service correspond à un stockage sécurisé des données, comportant obligatoirement un jeu minimal de métadonnées descriptives associées, afin de faciliter la recherche et la compréhension, pour une période déterminée (trois ou quatre ans) et offre une possibilité d'accès partagé à ces données pour la communauté. Au terme de ce laps de temps, les données sélectionnées seront soit versées en archive définitive, soit restituées à l'utilisateur pour destruction ou simple stockage dans son laboratoire. Ce n'est pas une solution proposée pour du stockage prolongé des données mais bien une approche pour faciliter la préparation d'une archive pérenne.

b L'implication dans les projets européens

En tant que centre national disposant d'une expertise dans le domaine du calcul intensif (HPC) et de la préservation à long terme de données numériques, le CINES participe à des programmes européens tels que : PRACE-1IP, PRACE-2IP ou EUDAT.

Le CINES a fait bénéficier le projet PRACE-1IP de son expertise en matière d'infrastructures de stockage et d'archivage en rédigeant un livre blanc sur les systèmes existants pour préserver les données scientifiques.

Les problématiques inhérentes à la préservation à long terme des données sont reprises dans le projet EUDAT (European Data Infrastructure). En effet, les communautés scientifiques ont exprimé auprès des centres de calcul européens le besoin de préserver et de partager les données qu'elles produisent chaque année. L'objectif du projet EUDAT est de mettre en place une grille européenne d'archivage de ces données scientifiques. La problématique est donc multiple : il s'agit à la fois de documenter, stocker, partager et répliquer des données de taille importante. D'ici à 2015, le CINES devrait être le nœud français de cette grille européenne de préservation à long terme de données scientifiques.



LE DÉPARTEMENT

58

- Ses missions 58
- Ressources humaines 58
- La gestion financière, le contrôle de gestion 58
- Les marchés 58

ORGANISATION DU CINES

59

- Le conseil d'administration 59
- Le comité hygiène et sécurité 59
- Le conseil du centre 59

Système de données

Exploitation

Réseau, hébergement et Infrastructure

Le département

a Ses missions

Les missions du Département Administration et Ressources Humaines (DARH) sont définies comme suit :

- ✚ Les ressources humaines,
- ✚ les conventions, marchés et travaux,
- ✚ la gestion financière, le contrôle de gestion.

Le Service du DARH est aussi chargé de veiller à l'efficacité de la circulation de l'information au sein des différents départements du CINES, ainsi qu'à l'extérieur, avec une mission de coordination et de suivi des actions entreprises par l'établissement.

Le département, sous la responsabilité du Secrétaire Général, compte un effectif de 12 personnes, réparties en 9 titulaires, et 3 personnes en CDD.

b Ressources humaines

Début 2010 le nombre total d'agents du CINES était de 50. Fin 2012, il était de 52. Cette évolution s'explique du fait du recours à 9 contractuels. L'effectif global des postes sur budget État est de 45 postes.

La répartition par catégories des personnels est stable et s'apprécie comme suit :

- ✚ Agents de catégorie A : 66,6 %,
- ✚ Agents de catégorie B : 22,2 %
- ✚ Agents de catégorie C : 11,11 %

Entre 2010 et fin 2012, 17 personnes ont quitté l'établissement :

- ✚ 10 par fin de contrat,
- ✚ 6 par départ à la retraite,
- ✚ 1 par concours.

Dans la même période 19 recrutements ont été effectués :

- ✚ 1 par voie de mutation,
- ✚ 12 sur des contrats à durée déterminée,
- ✚ 6 par concours.

2 agents ont bénéficié d'une promotion de grade.

c La gestion financière, le contrôle de gestion

Le Service du contrôle de gestion et de gestion financière a pour mission :

- ✚ de procéder aux mandatement et titres de recettes,
- ✚ de gérer la trésorerie de l'établissement,
- ✚ d'élaborer des outils de gestion (tableaux de bord) permettant le contrôle et le suivi budgétaire,
- ✚ de proposer des solutions pour optimiser les coûts et frais de fonctionnement,
- ✚ d'apporter un conseil à la Direction et aux différents Départements afin d'assurer une bonne visibilité,
- ✚ de réaliser, avec le concours de l'agent comptable, l'analyse budgétaire et financière (bilans comptables en particulier).

d Les marchés

Cette mission, donne au DARH les moyens de pouvoir organiser les procédures de marchés publics. Sa mission s'exerce auprès des autres départements dans le cadre d'une aide à la bonne détermination des besoins et la définition de la procédure la plus adaptée. Il est chargé d'élaborer les pièces contractuelles et de contrôler la régularité des procédures. Enfin, il s'occupe de la vérification des rapports de présentation soumis au contrôle préalable. Il assure enfin, le suivi d'exécution des marchés.

Organisation du CINES

Le CINES est doté par ses statuts d'un Conseil d'Administration (référence: décret 99) d'un Conseil Scientifique, d'un CHS, et, par son règlement intérieur, d'une instance paritaire, le Conseil de Centre.

a Le Conseil d'Administration

Le conseil d'administration comprend :

✚ Quatre membres de droit :

Le directeur chargé de la recherche au ministère chargé de l'enseignement supérieur et de la recherche ou son représentant, le directeur chargé des bibliothèques au ministère chargé de l'enseignement supérieur et de la recherche ou son représentant, le responsable de l'organe chargé de l'expertise scientifique et technique au ministère chargé de l'enseignement supérieur et de la recherche ou son représentant, le directeur général du Centre national de la recherche scientifique ou son représentant.

✚ Cinq personnalités nommées par arrêté du ministre chargé de l'enseignement supérieur et de la recherche, dont deux présidents d'université désignés sur proposition de la conférence des présidents d'université ;

✚ Trois membres élus par et parmi les personnels du Centre Informatique National de l'Enseignement Supérieur.

Le Directeur du CINES et le Contrôleur financier assistent aux réunions, avec voix consultatives. Le Secrétaire Général du CINES et l'agent comptable participent également aux séances.

Le conseil d'administration se réunit au moins deux fois par an, sur convocation de son président qui fixe l'ordre du jour.

En outre, il peut se réunir en séance extraordinaire, à la demande du ministre chargé de l'enseignement supérieur et de la recherche ou de la majorité de ses membres.

b Le Comité Hygiène et Sécurité

Le Comité Hygiène et Sécurité (CHS) du CINES a été créé en 2010. Il est composé de onze personnes. Le directeur du CINES préside les réunions du CHS. Il est assisté par le secrétaire général du CINES, de la responsable du département services informatiques et infrastructures (DS2I) et de l'agent chargé de la mise en œuvre (ACMO). Le personnel est représenté par cinq membres élus. L'inspecteur du travail et le médecin du travail font aussi parti de ce comité. Durant la période 2010-2012, le CHS s'est réuni cinq fois.

Il a depuis été remplacé par le Comité d'Hygiène, de Sécurité et des Conditions de Travail (CHSCT), créée en 2012, par arrêté du 12 décembre 2012.

c Le Conseil de Centre

C'est une structure d'information, de concertation et de communication entre la direction et les personnels du CINES.

Il a été remplacé par le Comité technique (CT), crée en avril 2012. Celui-ci n'est plus paritaire. Il est régi par le décret N°2011-184 du 15 février 2011, pour sa composition, le mode de désignation de ses membres et ses attributions. Le CT est un organe consultatif composé de 3 représentants du personnel nommés pour une durée de 4 ans par les syndicats représentatifs, du Directeur (président du CT) et du Secrétaire Général en sa qualité de chargé des ressources humaines. Les représentants du personnel ont chacun un suppléant désigné.

La Formation

La gestion de l'amélioration des compétences est devenue de plus en plus indispensable pour anticiper les besoins du CINES impliqué dans un élargissement constant de ses missions. Cette évolution permet en outre de faciliter les évolutions de carrière professionnelle des agents. La formation est un moyen de parvenir à atteindre ces résultats et objectifs. Les moyens sont pluriels et divers. Il peut s'agir de développer les compétences des agents, notamment le renforcement des expertises, ou leur permettre de soutenir l'intégration des nouveaux embauchés. Durant la période 2010-2012, le CINES a été engagé dans des projets de réorganisation et de créations d'activités. Des actions de formation ont donc été organisées afin d'accompagner ces changements et ainsi faciliter l'adaptation aux changements. Notre politique de formation est donc extrêmement liée au développement du CINES, de ses projets et de sa stratégie. Elle constitue une des réponses aux conditions de leur mise en œuvre.

Année 2010

	SUJET	DATES	NBRE DE JOURS	NBRE DE PARTICIPANTS
CIRen 44	SHIBBOLETH	25 - 28 janvier	4	39
CIRen 45	IPV6	21 au 24 septembre	4	14
CIRen 46	EVO	06-oct	1	20
CIRen 47	METROLOGIE	19 au 21 octobre	3	9
CIRen 48	DNS Sec	24 au 26 novembre	3	20

Année 2011

	SUJET	DATES	NBRE DE JOURS	NBRE DE PARTICIPANTS
CIRen 49	SHIBBOLETH	19-20 avril	2	20
CIRen 50a	EVO	7-8 juin	2	19
CIRen 50b	EVO	9-10 juin	2	14
CIRen 51	IPV6	19 au 23 septembre	4	18

a Les formations dites « CIREN »

Les formations « CIREN » sont organisées conjointement entre REATER et le CINES et sont abritées dans les locaux du Cines. Auparavant gratuites, elles sont aujourd'hui facturées 100 € pour couvrir les transports et hébergements des formateurs ainsi que les repas du midi des stagiaires. Certes il n'y a pas de frais de salles et mise à disposition des matériels informatiques, mais l'équilibre financier est ainsi assuré sur la période des trois années 2010-2012.

- ✚ 12 formations ont eu lieu sur la période.
- ✚ Le nombre de participants varie de 9 à 39.
- ✚ Le nombre de jours d'une formation de 1 à 4.
- ✚ Au total on compte 575 jours/participants entre 2010 et 2012

Année 2012

	SUJET	DATES	NBRE DE JOURS	NBRE DE PARTICIPANTS
CIRen 52	EVO	5 au 6 juin	2	15
CIRen 53a	SHIBBOLETH	19 au 20 juin	2	14
CIRen 35b	SHIBBOLETH	21 au 22 juin	2	10

b Les formations de type « expert en archivage »

Département Archivage et Diffusion

Les formations sous-citées correspondent à des cursus de niveau avancé, de type « expert en archivage électronique ». Ces formations ont été faites en interne :

- ✚ Expertise sur les formats de fichiers audio et vidéo
- ✚ Formation SIAF
- ✚ Formation Arcsys : Arcsys Software
- ✚ Formation DPC --> Lorène Bécard sur les «formats» de métadonnées METS et PREMIS
- ✚ Formation OAIS --> Equipe PAC
- ✚ Formation sur les relations entre archives (versions, mises à jour, parenté...) --> Equipe PAC
- ✚ Formation sur les métadonnées dans les archives --> Equipe PAC
- ✚ Formation ORI-OAI --> Equipe PAC
- ✚ Formations sur les processus métier (travail de formalisation, présentation des processus métier mis en œuvre dans le service),
- ✚ iRODS
- ✚ ESGF (outil pour la gestion mondiale des données produites par la communauté de climatologie)
- ✚ iRODS

Par ailleurs, le CINES intervient dans des cursus de formation organisés par des établissements extérieurs (Association des Archivistes Français, ENSIBB, DG des patrimoines, Ecole nationale des Chartes, PIN...).

Voici la liste des établissements partenaires (organismes de formations dans lesquelles les personnels du CINES sont intervenus) :

- ✚ AAF (Association des Archivistes Français)
- ✚ ENSSIB - Ecole Nationale Supérieure des Sciences de l'Information et des Bibliothèques)
- ✚ Ecole Nationale des Chartes (formation initiale de niveau master et formation continue)
- ✚ PIN (groupe Pérennisation des Informations Numériques de l'association Aristote)
- ✚ Ministère de la Culture et de la Communication : direction générale des patrimoines
- ✚ INP (Institut national du Patrimoine) Centre du livre et de la lecture en Poitou-Charentes

c Les formations calcul

La formation aux techniques et aux outils du calcul intensif est plus que jamais à la base de la réussite des projets scientifiques de modélisation et de simulation. Par le passé, le CINES et avant lui le CNUSC, a toujours été présent sur le terrain de la formation. Depuis 2010, le CINES a souhaité adapter son offre de formation au nouveau paysage du HPC. Ce travail est encore en cours et les deux filières présentées ici sont appelées à se développer dans les mois qui viennent. Deux axes de développement de l'activité formation ont été privilégiés : le plan national et le plan européen. Au niveau national, le travail s'est structuré autour de l'élaboration d'une nouvelle filière de formation dont l'objectif est de tirer le meilleur parti des possibilités techniques offertes par la machine Jade.

- ✚ Large scale data visualization with Visit (2 jours), dispensé une première fois en novembre 2012 (15 participants) ;
- ✚ Jade Tier-1 system workshop (1 jour),
- ✚ Initiation to parallelization using MPI (3 jours),
- ✚ Code coupling using OpenPalm (3 jours, réalisé avec le concours du CERFACS).

La mise à jour des connaissances est indispensable pour maintenir un service de qualité et proposer de nouvelles solutions à nos utilisateurs. En soutien de la veille technologique personnelle, des formations sont proposées, en fonction des besoins individuels et des projets. Ces formations sont réalisées souvent par le CINES, parfois par des prestataires extérieurs. Le département a bénéficié et produit 50 jours/homme de formation dans les 12 derniers mois.



Centre Informatique National de l'Enseignement Supérieur

950, rue de Saint-Priest - 34097 Montpellier Cedex 5 - France

Tél. : 04 67 14 14 14 - Fax : 04 67 52 37 63

www.cines.fr - services utilisateurs : svp@cines.fr

Directeur de la publication : Francis Daumas
Co-rédactrices en chef : Marion Massol et Lorène Bécharde
Contact - responsable de communication : Claude Kindou

