



[www.cnrs.fr](http://www.cnrs.fr)

## Le Comité d'Orientation pour le Calcul Intensif (COCIN)

M. Daydé

Dr du COCIN

Délégué Scientifique INS2I en charge

HPC / Grille / Cloud

---

## Le CNRS en deux mots



P. 03

- | Environ **34,000** personnes dont **25,600** permanents, **11,400** chercheurs et **14,200** ingénieurs, techniciens and administratifs
- | Budget 2010 de l'ordre de **3** milliards d'euros
- | Environ **1,100** unités de recherche
- | **25,500** articles / an dans des journaux internationaux en moyenne
- | **16** Prix Nobels et **11** Médailles Fields



# THE TEN CNRS INSTITUTES



---

# Calcul Intensif au CNRS



- Enjeu stratégique (EU, USA, Japon, ...)
- Transverse au sein du CNRS car au coeur des pratiques de la plupart des Instituts
- Simulation numérique, méthodes et outils, mais aussi émergence du traitement sur les grandes masses de données

---

# Panorama du HPC



- 3 centres nationaux : CINES, IDRIS, CCRT
- Hiérarchie au sens de PRACE :
  - Tiers 0 : TGCC (CURIE CEA), objectif  $> 1$  PFlops fin 2011 (actuellement 300 TFlops)
  - Tiers 1 : centres nationaux (CINES, IDRIS)
  - Tiers 2 : CC INP2P3, méso-centres, ....
- GENCI finance et coordonne les investissements dans les centres nationaux (et est propriétaire des machines)



# Le Comité d'Orientation pour le Calcul Intensif (COCIN)

---

## Rôle et missions du COCIN



- Créé en Décembre 2010
- Réflexion collective sur les besoins, la structuration et les évolutions en calcul intensif au CNRS
- Prospective sur les besoins des différentes communautés, proposition de maintenance et de développement coordonné des moyens / ressources liées au calcul intensif, en particulier pour l'IDRIS.
- Dix personnalités scientifiques désignées par chacun des instituts du CNRS plus le Directeur de l'IDRIS.
- Le président et directeur désignés par le Président du CNRS



# Composition / Fonctionnement

- Président : Ph. Baptiste (INS2I)
- Dr : M. Daydé (INS2I)
- Membres : M. Asch (INSMI), D. Boutigny (IN2P3), G. Deléage (INSB), D. Girou (IDRIS), C. Le Van (INSHS), L. Lellouch (INP), C. Pouchan (INC), S. Thiebault (INEE), D. Veynante (INSIS), JP Vilotte (INSU)
- Réunion tous les 1er mercredi du mois

---

# Quelques sujets abordés



- S'emparer des sujets du moment : IDRIS, Maison de la Simulation, EQUIPEX, coordination de l'informatique en appui à la recherche au CNRS ....
- Etat des lieux, cartographie, besoins, perspectives du calcul intensif au sein du CNRS (en cours)
- Réflexion autour de l'IDRIS, articulation Tiers 1 et Tiers 2 (méso-centres)
- Liens / complémentarités entre HPC / grilles / cloud
- Problématique du traitement des grands volumes de données

---

# Maison de la Simulation



- Soutenir et stimuler les communautés afin d'utiliser au mieux les moyens déployés par GENCI et PRACE.
- Favoriser l'émergence en France d'une communauté du calcul intensif
- Similaire à *Advanced Institute for Computational Science (AICS, Japon)*
- Partenaires (CEA, CNRS, INRIA, UVSQ, Paris Sud)
- Bi-localisation Digitéo-Saclay et IDRIS
- USR créée début 2011
- Dr E. Audit (CEA), adjoints : A. Bui (CNRS), M. Kern (INRIA)

---

# Livre Blanc sur le Calcul Intensif au CNRS



- Lancé en Avril 2011 sur la base d'un questionnaire
- Contribution CNRS à la réflexion lancée par GENCI sur les besoins des communautés et sa stratégie pour 2012-2016 (renouvellement IDRIS en 2012)
- Panorama du calcul intensif dans les Instituts : pratiques, besoins, verrous, cartographie, aspects recherche, ...
- Certains des instituts (e.g. INC, INP) mènent régulièrement ce type de travail
- Objectif : finaliser ce travail pour fin 2011 et mener des réactualisations régulières



# Quelques éléments de réflexion

- High performance computing (Tiers 0, Tiers 1) vs high throughput computing (Tiers 1, méso-centres, grille de production)
- Stockage des données local / distant : organisation en Tiers, cloud ?
- Gestion / exploitation des grands volumes de données : I/O, traitement intensif, pré / post-traitement, visualisation, extraction d'information, fouille, mise à disposition des communautés, ...
- Hétérogénéité des architectures : multi-coeurs, GPU, ...
- Passage à l'échelle
- Résilience, support d'exécution, .....
- Evolution des algorithmes / méthodes de résolution (co-design), librairies, paradigmes de programmation
- Liens / complémentarités HPC / grilles / cloud : calcul, données



# Éléments de réflexion (suite)

- Quelques communautés ont des codes pouvant exploiter un parallélisme à haut degré (jusqu'à 10 000 cœurs) :
  - Combustion, astrophysique, géophysique,
  - Apparition récente de codes portés sur les accélérateurs type CELL ou GPU
  - Préparation des codes (nouvelles architectures, rôle des méso-centres et centres nationaux,...)
  - Besoins parfois supérieurs aux allocations d'heures
- Constitution de communautés autour de codes
- Formation, compétences

---

# Difficultés



- Préparation des codes face aux évolutions actuelles :  
passage à l'échelle, programmation, hétérogénéité des architectures
- Formation / recrutement de spécialistes HPC en prise avec les applications
- Constituer des communautés autour de codes
- Bien évaluer les coût de conception / développement des applications.

---

# Conclusion



- CNRS a en son sein tout l'écosystème du calcul intensif : PRACE, IDRIS, CC IN2P3, Institut des Grilles, recherche, interaction avec les méso-centres (Gdr Calcul, UMR), ... et toute la diversité des besoins / pratiques au CNRS (HPC, grille, cloud, ...) et caractère interdisciplinaire
- Favoriser les échanges entre communautés HPC / grilles / Cloud
- Synergie avec les acteurs du HPC (CEA, GENCI, INRIA, universités, ...)
- Promotion du calcul intensif à tous les niveaux centres nationaux et méso-centres
- Emergence rapide d'un fort besoin en matière de stockage / traitement sur les grands volumes de données