

Mésocentre Centrale Paris

5ème journées Mésocentres

Matthieu Boileau, Christophe Couture, Damien Durville,
Marc Massot, Anne-Sophie Mouronval, Laurent Series



Historique du projet

Septembre 2009

Décision d'acquisition d'un équipement mutualisé co-financé par l'Ecole et les ressources propres des laboratoires

Fin 2009 à Juillet 2010

Dialogue compétitif

Septembre 2010

Présentation du projet à la 3ème journée des mésocentres

Octobre 2010 à Janvier 2011

Installation de la machine

Février 2011

Inauguration et mise à disposition de la machine aux utilisateurs

- Contexte scientifique d'un projet fédérateur
- Organisation du mesocentre
- Présentation de la machine et de son utilisation
- Applications scientifiques

- Depuis 2004 :
 - collaborations entre EM2C et MSSMat sur les calculateurs parallèles de chacun des deux laboratoires
 - l'accès à un moyen de calcul local disponible a généré un véritable développement du calcul au sein des laboratoires
- Spécificité Centrale Paris : culture de la modélisation mathématiques et de la simulation numérique
- Essor de la communauté des mésocentres grâce notamment au groupe calcul du CNRS

C'est dans ce contexte propice que naît le projet du mesocentre de Centrale Paris avec l'ensemble de ses laboratoires.

- ▶ point clef : soutien de la direction de Centrale (sensibilisation des instances dirigeantes aux problématiques du Calcul)

Les objectifs du mésocentre

Mettre à disposition d'un environnement de Calcul Scientifique performant

- acquisition d'un système de calcul performant adapté aux différents besoins exprimés
- mise en place d'une organisation de l'exploitation et du support aux utilisateurs

Créer un véritable outil pour la recherche

- un lieu de mise au point des codes pour le passage à l'échelle (passerelle vers les centres nationaux)
- un lieu de développement d'une nouvelle génération de méthodes numériques et d'algorithmes

Créer une communauté du calcul

- amener des nouveaux acteurs de l'école au calcul scientifique

Créer un lieu d'échange

- séminaires / formations
- ouverture vers l'extérieur / interface avec le groupe Calcul

Moyens humains

- un demi-poste IR Ecole dédié au mésocentre
- participation active des ingénieurs calculs d'EM2C et de MSSMAT
- soutien du service informatique (poste à venir ?)

Comité technique opérationnel

Composition : le directeur et directeur adjoint du mésocentre, un ou deux représentants des laboratoires, les 3 ingénieurs calcul et le responsable du service informatique

- gestion des questions courantes : utilisation de la machine
- définition des orientations techniques du mésocentre
- relais des décisions prises auprès des utilisateurs

Comité scientifique (experts internes et externes Centrale Paris)

- fournir une évaluation scientifique des travaux réalisés au mesocentre

Animation scientifique

- séminaires
- formations spécifiques
- lien avec l'enseignement

- ▶ Point clef : **Le bon fonctionnement du projet dépend d'un noyau motivé associant chercheurs et ingénieurs**

- Un ordinateur SGI Altix ICE de 68 nœuds de calcul équipés de 2 processeurs hexa-cœurs soit **816 cœurs de calcul** dont :
 - 8 nœuds de 48 Go et 4 nœuds de 72 Go soit **144 cœurs** dédiées aux applications séquentielles et parallèles à mémoire partagée
 - 56 nœuds de 24 Go de mémoire soit **672 cœurs** dédiées aux applications parallèles à mémoire distribuée

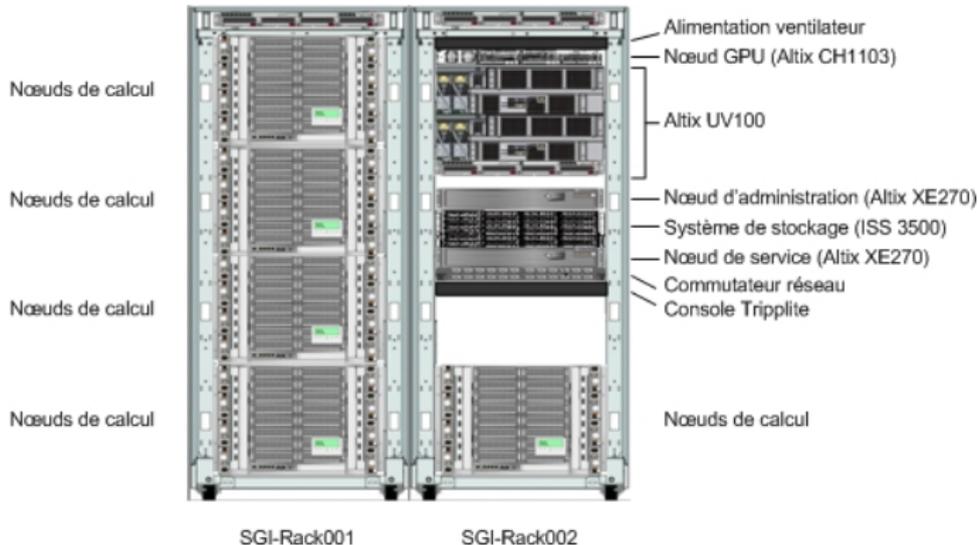
- Un nœud de calcul SMP de 8 processeurs hexa-cœurs soit **48 cœurs de calcul** de **256 Go** de mémoire et équipé de **4 cartes graphiques** dédié :
 - aux applications séquentielles et parallèles à mémoire partagée nécessitant beaucoup de mémoire
 - à la visualisation (déportée) des résultats de calcul

- Un nœud GPU de 2 processeurs quadri-cœurs équipé d'une **carte avec GPU compatible avec CUDA** dédié :
 - aux tests d'application utilisant le calcul sur GPU

- Un serveur de stockage SGI de **30 To en RAID 5** permettant à chaque utilisateur de disposer :
 - d'un répertoire de connexion d'un quota de 10 Go
 - d'un répertoire de travail d'un quota de 100 Go
- Un nœud de service permettant :
 - la connexion depuis un poste de travail
 - l'édition de sources, la compilation des programmes
 - le lancement et le suivi des jobs
- Un nœud d'administration permettant :
 - l'administration de la machine (création des comptes, mise à jour, sauvegarde d'image système)

Architecture de la machine igloo

Vue d'ensemble

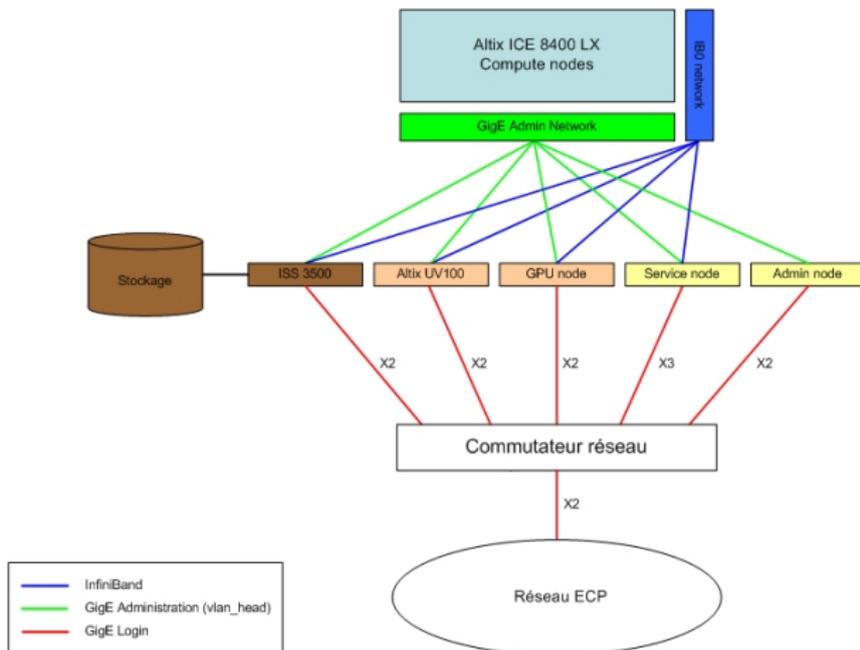


Architecture de la machine igloo

Aspect réseau

Les noeuds de calculs sont interconnectés via :

- un réseau InfiniBand pour les calculs et les Entrées/Sorties
- un réseau Gigabit pour les connexions et loadadministration



Le paramétrage de PBSPro est un point clef de la bonne utilisation de la machine pour respecter les objectifs du projet

Caractéristiques du paramétrage

- Limiter le temps d'exécution des jobs pour :
 - favoriser la rotation des jobs
 - inciter les utilisateurs à implémenter une procédure de reprise dans leur code
- Dédier une part importante de la machine aux jobs parallèles à memoire distribuée pour :
 - favoriser les jobs parallèles
 - inciter les utilisateurs à paralléliser leur code
- Dédier des nœuds pour les tests en journée :
 - favoriser le developpement des codes

Près d'une cinquantaine de produits installés sur la machine :

- bibliothèques : mkl, fftw, boost...
- applications : siesta, abinit, freefem...
- outils de debugage et de profilage : intel-idb, intel-itac, valgrind...
- outils de visualisation : paraview, tecplot, ensight...
- utilitaires : cmake, git, doxygen...

La commande **module** permet aux utilisateurs de paramétrer leur environnement en fonction de logiciels qu'ils souhaitent utiliser.

- Exemple pour la bibliothèque boost :

```
> module avail boost

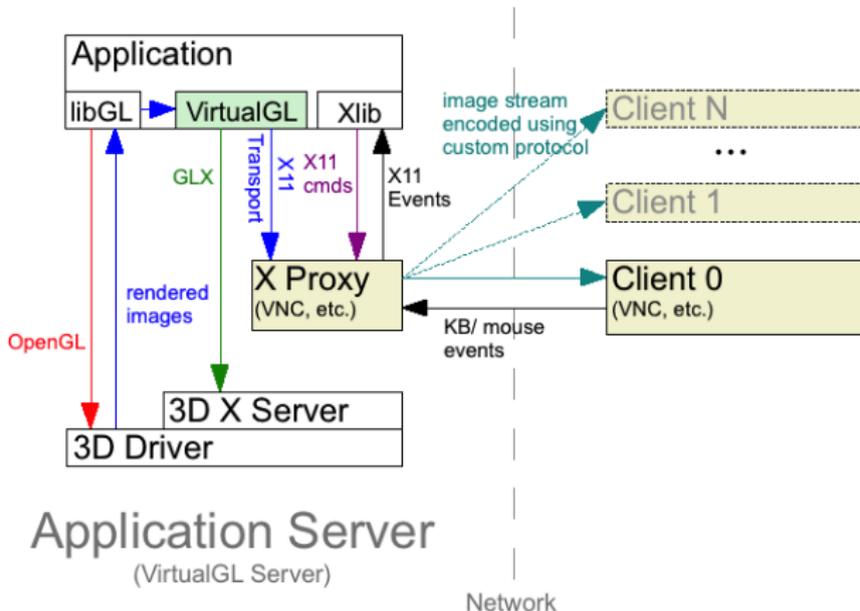
----- /opt/san/sgi/modulefiles/bibliotheques -----
boost/1.46.1-intel11.1      boost/1.49.0-intel12.1(default)
boost/1.49.0-intel11.1    boost/1.51.0-intel12.1

> module load boost/1.51.0-intel12.1
```

Utilisation de la machine

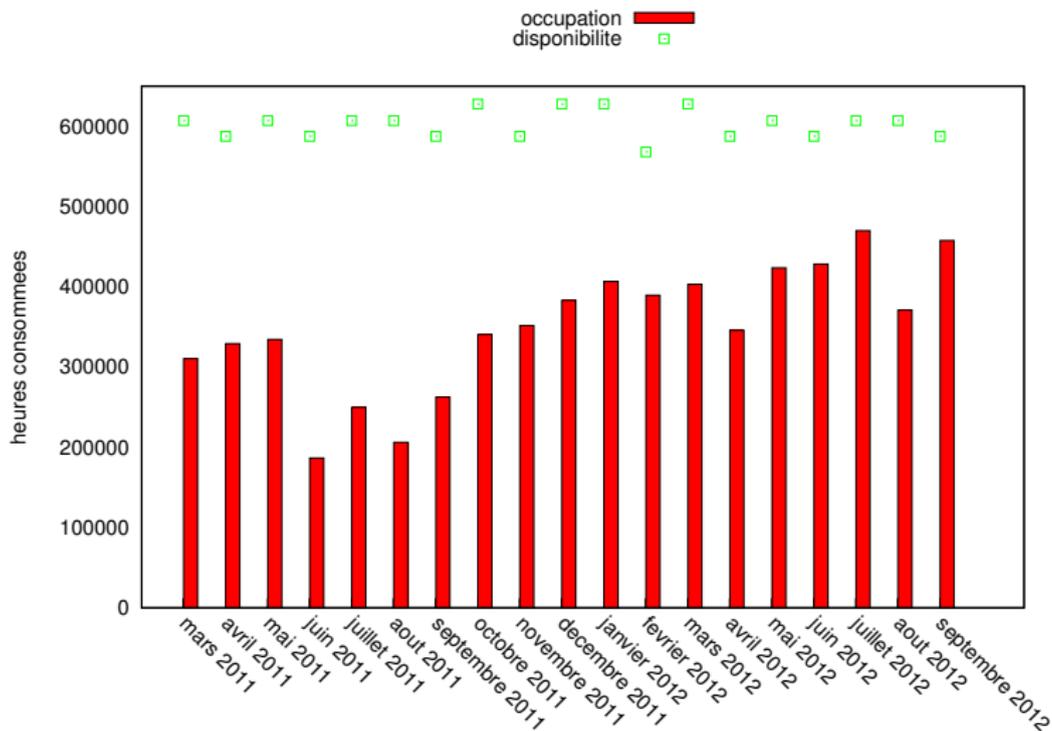
Visualisation à distance

La visualisation à distance mise en oeuvre utilise le logiciel VirtualGL permettant d'effectuer le rendu 3D sur le serveur (nœud SMP).

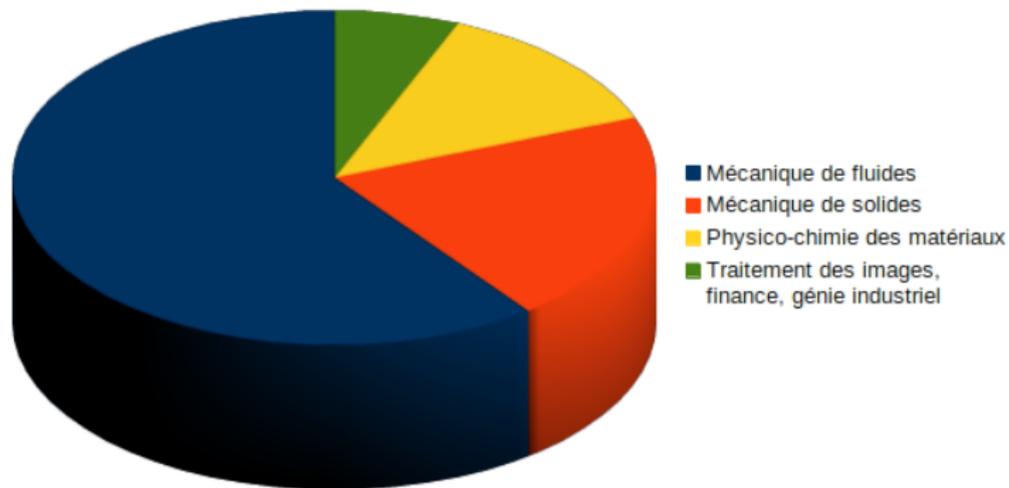


VNC (Virtual Network Computing) est un système permettant de contrôler une machine à distance

Bilan d'exploitation



Heures consommées par mois de l'Altix Ice



Ventilation par domaine d'application

Exemple d'application

Projet Multifil

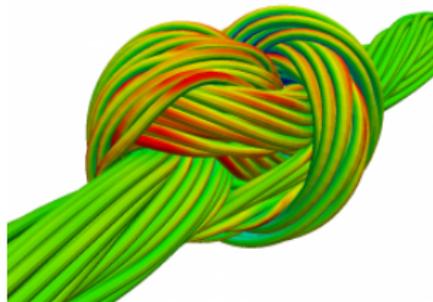
Contexte

- simulation numérique du comportement mécanique de milieux enchevêtrés
- applications : renforts câblés, renforts textiles pour les composites

Code éléments finis Multifil

- code spécialement dédié aux milieux enchevêtrés
- modèle de poutre en grandes transformations
- prise en compte des interactions de contact-frottement

Simulation du serrage d'un noeud
réalisé sur un fil torsadé de 48 fibres



Simulation du serrage d'un noeud
réalisé sur un fil torsadé de 48 fibres

Caractéristiques du modèle

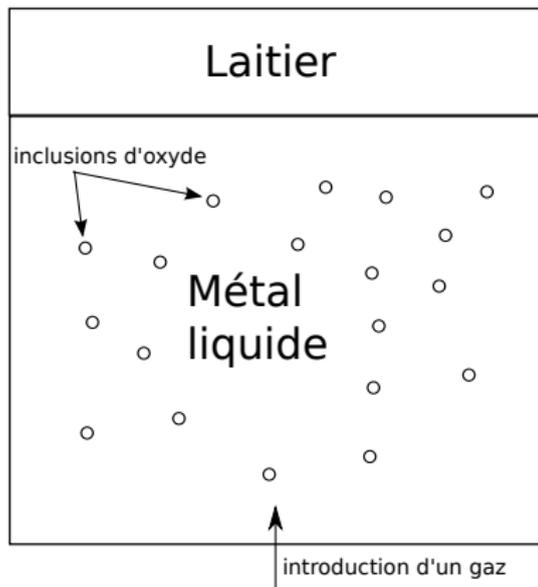
- 150 000 ddl
 - 80 000 éléments de contact
 - 48 heures sur 4 cœurs (openmp)
-
- Etude d'une parallélisation MPI envisagée

Exemple d'application

Traitement en poche des aciers

Simulation numérique directe du transport et du dépôt de particules solides au voisinage d'une surface libre

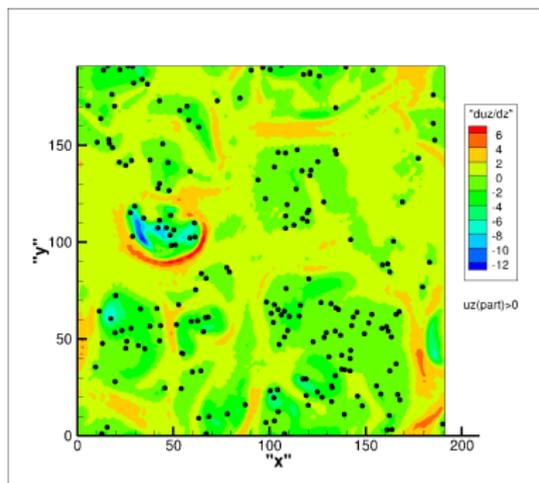
Arunvady Xayasenh - thèse de doctorat de Centrale Paris, 2011



Modèle couplant la simulation numérique directe de la turbulence au voisinage de l'interface et le suivi lagrangien des inclusions en suspension.

Exemple d'application

Traitement en poche des aciers



Carte de l'accélération normale du fluide

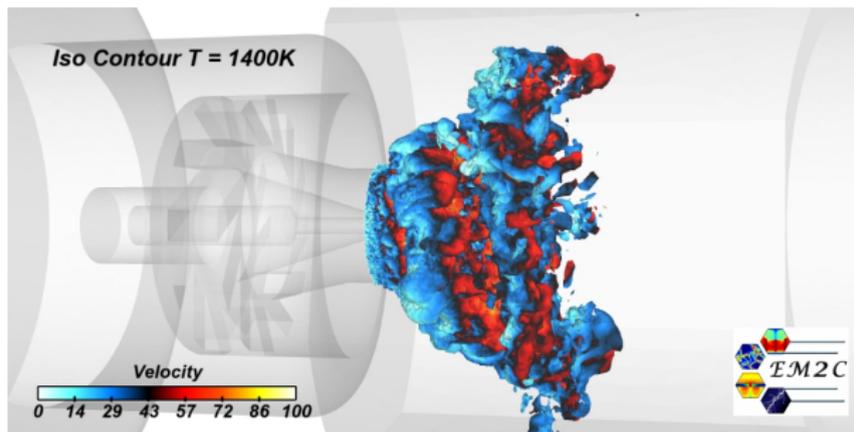
- Etude paramétrique afin d'estimer le flux d'inclusions "déposées" par turbulence à l'interface métal liquide/laitier en fonction des caractéristiques des inclusions (diamètre, densité) et de l'écoulement du métal liquide
- Chaque simulation a consommé près de deux semaines de calcul sur 12 cœurs (OpenMP).

Exemple d'application

Modélisation de la combustion turbulente

Code AVBP - Simulation aux grandes échelles de la chambre de combustion expérimentale Molecules (TU Darmstadt) équipée d'un injecteur aéronautique Turboméca (Groupe Safran)

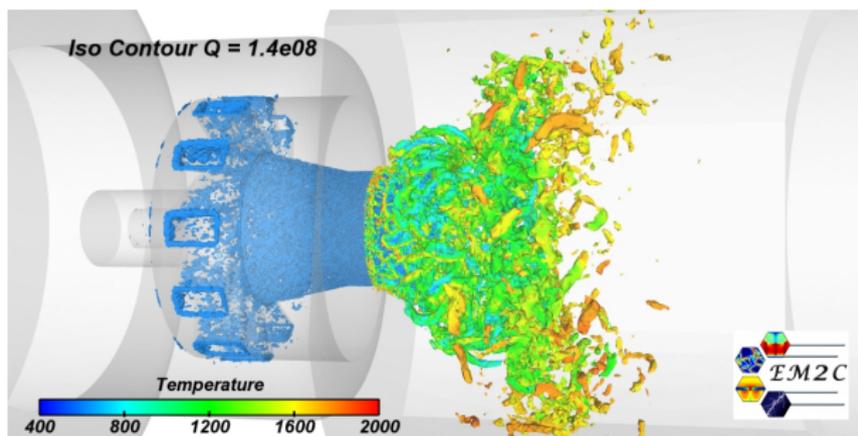
Pierre Auzillon - thèse de doctorat de Centrale Paris, 2011



Visualisation du front de flamme

Exemple d'application

Modélisation de la combustion turbulente



Visualisation des structures turbulentes

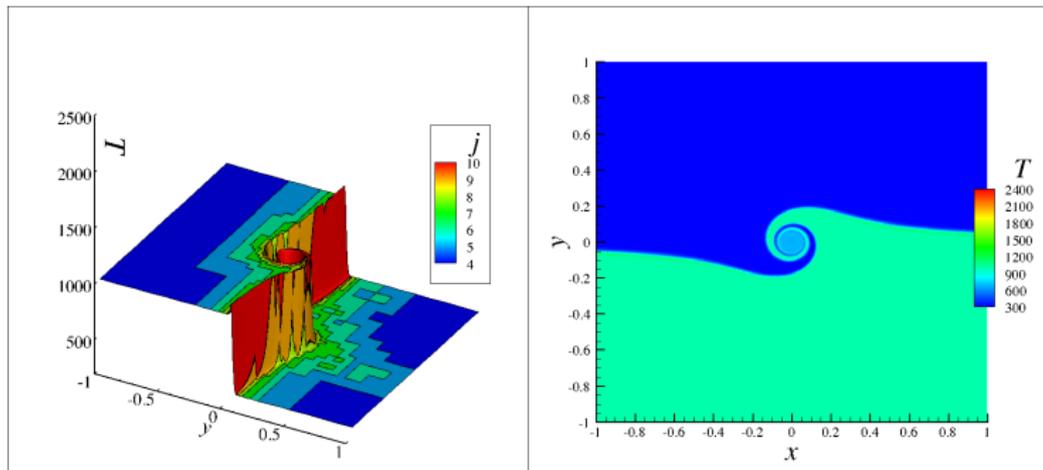
- Maillage de 40 millions de tétraèdres
- Mise en place du calcul sur **igloo** (> 10 000 heures)
- Production sur 2048 au Cines (250 000 heures)
- Post traitement sur **igloo** (1 To de données)

Exemple de projet

Simulation de fronts multi-échelles avec adaptation espace/temps et contrôle d'erreur

Code MBARETE - Multi-résolution adaptative en espace - séparation d'opérateur en temps avec intégrateurs d'ordre élevé en temps pour chaque sous bloc explicite/implicite et adaptation dynamique du pas de temps

Max Duarte - thèse de doctorat de Centrale Paris, 2011 - ANR Séchelles - Prix de Thèse GAMNI/SMAI 2011 et Prix Européen meilleur thèse ECCOMAS 2011



Auto-allumage d'une flamme de diffusion - Finest Grid = 1024^2 - $\frac{\text{Active Grid}}{\text{Finest Grid}} \lesssim 7\%$

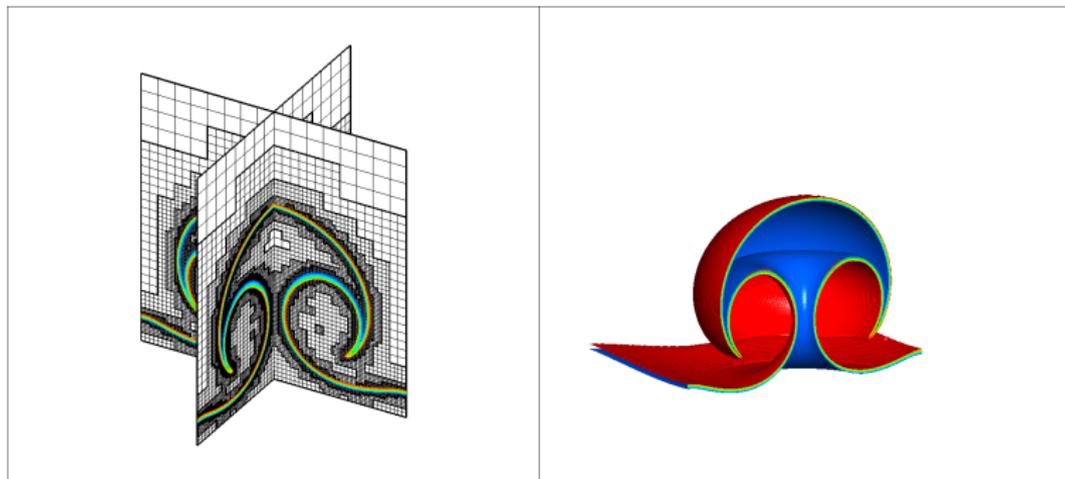
CPU time $\sim 3\text{h}28 \Rightarrow \sim 0\text{h}09$ - CPU time finest grid/constant splitting $10^{-7} \sim 11\text{h}15$

Exemple de projet

Simulation de fronts multi-échelles avec adaptation espace/temps et contrôle d'erreur

Code MBARETE - Multi-résolution adaptative en espace - séparation d'opérateur en temps avec intégrateurs d'ordre élevé en temps pour chaque sous bloc explicite/implicite et adaptation dynamique du pas de temps

Max Duarte - thèse de doctorat de Centrale Paris, 2011 - ANR Séchelles - Prix de Thèse GAMNI/SMAI 2011 et Prix Européen meilleur thèse ECCOMAS 2011



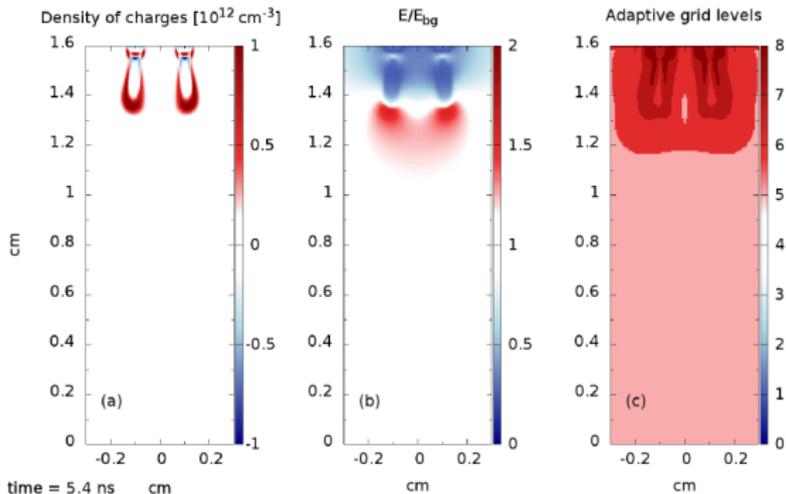
Interaction flamme prémélange - tourbillon Finest Grid = 256^3 - $\frac{\text{Active Grid}}{\text{Finest Grid}} \lesssim 15\%$ CPU time $\sim 14\text{h}$

Exemple de projet

Simulation de fronts multi-échelles avec adaptation espace/temps et contrôle d'erreur

Code MBARETE - Multi-résolution adaptative en espace - séparation d'opérateur en temps avec intégrateurs d'ordre élevé en temps pour chaque sous bloc explicite/implicite et adaptation dynamique du pas de temps

Max Duarte - thèse de doctorat de Centrale Paris, 2011 - ANR Séchelles - Prix de Thèse GAMNI/SMAI 2011 et Prix Européen meilleur thèse ECCOMAS 2011



Interaction of streamer - $E = 48 \text{ kV/cm}$ Finest Grid = 7680x8192 - CPU time $\sim 45\text{h}$

Exemple de projet

Simulation de fronts multi-échelles avec adaptation espace/temps et contrôle d'erreur

Des mathématiques aux sciences appliquées

- Transfert de développements mathématiques vers des sciences appliquées dans le cadre du méso-centre
- Applications en ingénierie bio-médicale (ICJ Lyon), physique des plasmas, écoulements réactifs - allumage annulaire - SAFRAN avec des avancées dans les domaines applicatifs

Nécessité de passer au calcul parallèle distribué / hybride

- Reprise des développements en changeant la structure de données (T. Dumont)
- Ouverture sur des questions de fond (parallélisation de solveurs de Poisson sur grilles adaptées dans le cadre multi-résolution (thématiques communes)
- Collaboration du méso-centre sur ces questions

Une extension en discussion

- ajout de 12 ou de 22 nœuds de 24 Go à l'étude
- ajout de nœud équipé de GPU dédié au calcul à l'étude

Le renouvellement de la machine

- veille technologique du comité technique
- pose la question de l'hébergement (conteneur, hébergeur...)

Une machine bien dimensionnée et une très bonne utilisation

- Cohérence avec les objectifs du méso-centre
- Tous les laboratoires impliqués - réussite au niveau du Centre de Recherche de l'ECP
- Administration et gestionnaire de queue très bien faits
- **Combinaison entre un travail équipe abouti et un axe stratégique au niveau de la direction ECP**

De belles avancées scientifiques

- Création d'une communauté du calcul autour des méthodes numériques et calcul intensif
- Element important de la vie scientifique de l'ECP - Création 2013 de la Fédération Mathématique de l'ECP, unité FR CNRS avec l'INSMI s'appuie sur le Mésocentre - Lien Université Paris-Saclay - Maison de la Simulation - FMJH

Percolation lente vers l'enseignement aux Elèves Ingénieurs

- En cours mais cela demande dans un premier temps l'établissement de la communauté calcul
- Cf. CSCI - Calcul intensif dans les Ecoles Ingénieur - plutôt au niveau doctoral pour le moment

Consolidation des développements et algorithmes performants

- Timide car la première étape était la structuration de la communauté au sein de l'ECP et relations avec l'extérieur
- Lien avec la MdS - améliorer la connexion avec les nouvelles architectures
- Connexion importante avec la communauté de maths via le lien Fédération de Maths / Mésocentre

Investissement humain

- Investissement massif d'un groupe cohérent dans le projet, mais IR Calcul très (trop?) impliqués - ASR en cours de recrutement - focaliser vers les aspects scientifiques dans le cadre des points ci-dessus