

Etat des lieux des structures de type mésocentres en France (septembre 2009)

F. Berthoud, V. Louvet

Ce rapport correspond à la mise à jour d'un premier document réalisé en février 2008 en préalable à l'organisation d'une journée nationale dédiée aux méso-centres, le 13 février 2008 à Paris. Ce document est rédigé dans le cadre de l'organisation d'une deuxième journée consacrée aux mésocentres à Paris le 24 septembre 2009.

Depuis la diffusion du précédent document, trois rapports relatifs au calcul mentionnent ces structures ¹ :

- le rapport sur la problématique de l'externalisation du temps de calcul au CNRS rédigé par des experts des réseaux Calcul et ResInfo,
- le rapport du CSCI (Comité Stratégique pour le Calcul Intensif) pour l'année 2008.
- Les rapports du colloque « Penser Petaflops », plus particulièrement des groupes de travail « Infrastructures du calcul intensif » et « Les métiers du calcul numérique : formation, recherche et débouchés ».

1 Topologie des moyens de calcul en France

1.1 Définition d'un mésocentre

Il est important de rappeler la définition retenue pour un mésocentre :

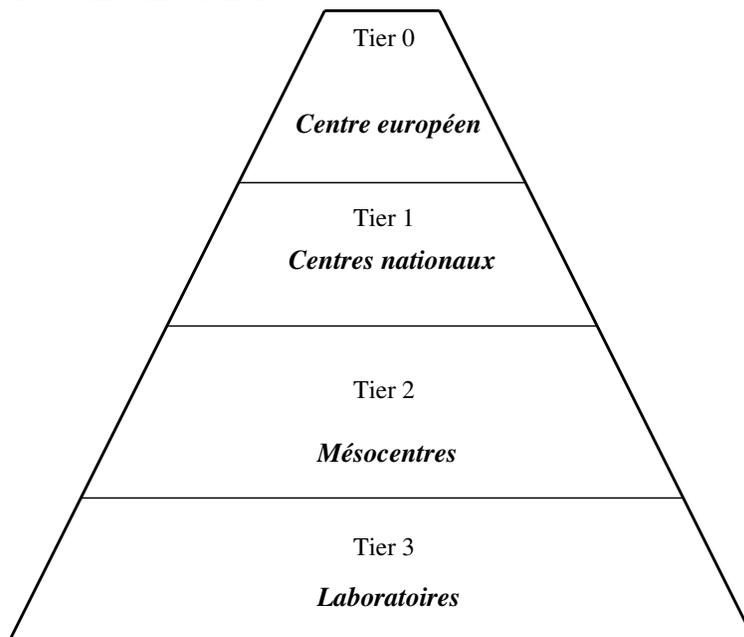
Un méso-centre pourrait être caractérisé par :

- Un ensemble de moyens humains, de ressources matérielles et logicielles à destination d'une ou plusieurs communautés scientifiques, issus de plusieurs entités (EPST, Universités, Industriels) en général d'une même région, doté de sources de financement propres, destiné à fournir un environnement scientifique et technique propice au calcul haute performance.
- C'est une structure pilotée par un comité scientifique (plus ou moins structuré) et, en principe, évaluée régulièrement.

¹Accessibles sur cette page : <http://calcul.math.cnrs.fr/spip.php?rubrique17>

1.2 Structure pyramidale

Les différents moyens de calcul à disposition des chercheurs sont communément répartis selon une structure pyramidale en fonction de leur puissance et de leur couverture locale, régionale, nationale ou internationale.



Le rapport du CSCI préconise d'assurer un bon équilibre de cette pyramide afin que chacun des niveaux soit une étape naturelle vers le niveau supérieur et permettre ainsi le développement du calcul intensif.

2 Caractérisation des mésocentres

La caractérisation matérielle des moyens de calcul intermédiaires est désormais bien décrite ² (voir paragraphe suivant). L'angle des moyens humains, des pratiques et des moyens financiers est beaucoup moins connu. C'est l'objet du rapport « Le calcul dans les laboratoires de recherche : pratiques et moyens » édité en février 2009.

Nous en rappelons ici quelques éléments importants dans la compréhension des spécificités des mésocentres.

2.1 Complémentarité des tiers

Les centres nationaux apportent aujourd'hui une puissance de calcul beaucoup plus importante que dans les années précédentes. Les récents investissements ont permis de mettre à disposition des utilisateurs des équipements actualisés. Ces moyens sont particulièrement adaptés aux grosses applications de production très optimisées, mais peu au développement

²voir sur le site de Calcul : <http://calcul.math.cnrs.fr/spip.php?rubrique7>

et aux tests de scalabilité. Leur accès, même si de gros efforts ont été fait dans ce domaine, manque de souplesse.

Les mésocentres ont un rôle intermédiaire essentiel : ils permettent aux chercheurs de mettre au point leurs codes de calcul, de réaliser certains types de simulations, de disposer d'outils pour pré ou post-traiter leurs données ...

2.2 Intérêts de la proximité

Une machine de calcul est un outil très complexe qui doit répondre à une problématique. Ses caractéristiques ne se limitent pas à la fréquence des processeurs ou à sa puissance crête. Les problèmes d'accès mémoire, de réseau, de types de calculs (parallèles ou non et type du parallélisme ...) sont souvent spécifiques au projet scientifique. Le déploiement des codes nécessite régulièrement des ajustements de l'environnement logiciel de la machine, voire du système lui-même. La proximité entre utilisateurs/développeurs et administrateurs système, dont les compétences sont particulièrement pointues et adaptées au contexte, est donc essentielle et indispensable.

Les personnels techniques en charge de l'administration système des machines de calcul et de leur exploitation font le lien entre les constructeurs et les chercheurs, pour répondre au mieux aux besoins. De plus la formation dans ces domaines des doctorants, mais aussi celle des chercheurs plus confirmés, se fait le plus souvent au contact de ces mêmes personnels.

Cette proximité a donc des atouts incontestables et est essentielle dans les structures de méso-informatique.

2.3 Importance du rôle des mésocentres dans la formation

Plusieurs rapports font état de la problématique de la formation dans le domaine du calcul. Les mésocentres, par leur proximité avec les utilisateurs, leur régionalité, doivent être et sont souvent moteurs en proposant, une vaste palette de formations liées au calcul à l'ensemble des personnels concernés quelque soit leur statut.

2.4 Atouts de la mutualisation

La mutualisation des moyens facilite la mise en place de ressources humaines de haut niveau pour la gestion de ces plateformes et l'optimisation en terme de puissance et temps de calcul.

La mutualisation des coûts (humains et infrastructures) permet, pour un budget donné, d'optimiser l'exploitation de ces ressources importantes, compte tenu du nombre d'utilisateurs et de leurs besoins à des périodes de temps différentes.

3 Etat des lieux de la mésoinformatique en septembre 2009

Il est tout d'abord intéressant de noter que, dans le classement top500 de juin 2009, les 10 premiers ont une performance crête de 300 à 1456 TFlops, et les 10 derniers de 21 à 37 TFlops.

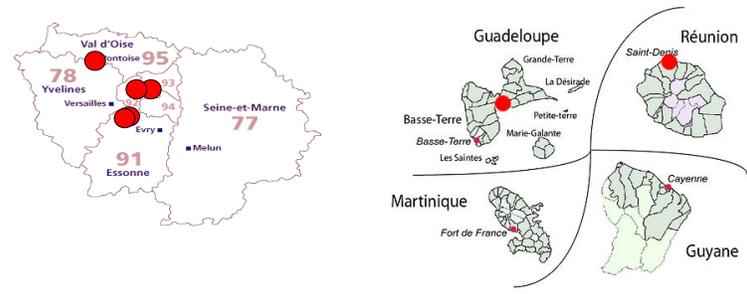
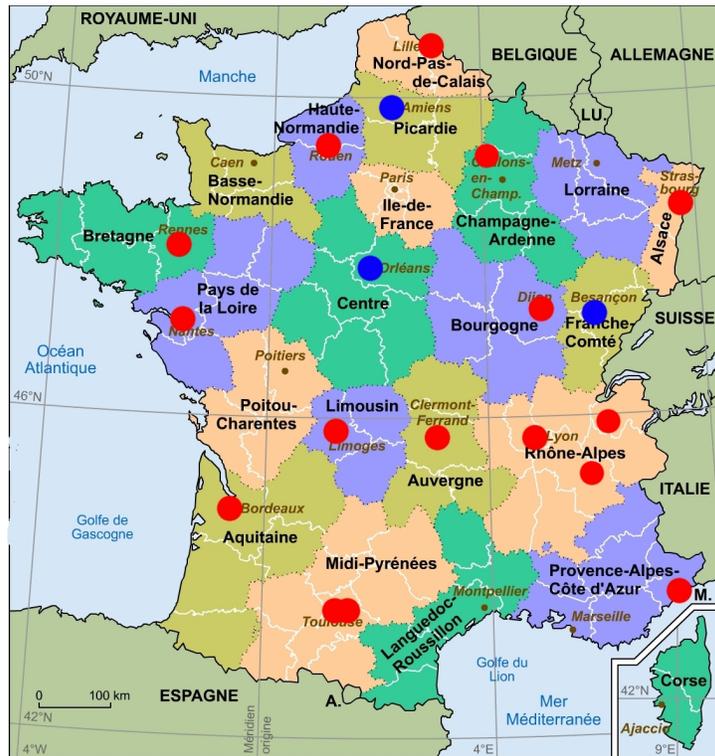
Par ailleurs, la machine Jade du CINES affiche une performance crête de 146,74 TFlops, et la blue gene de l'IDRIS de 139,26 TFlops sur ce même classement.

La puissance brute déployée dans les mésocentres est donc loin d'être négligeable puisqu'on arrive à une puissance crête totale de 182.25 TFlop/s (contre 31,45 TFlop/s en février 2008).

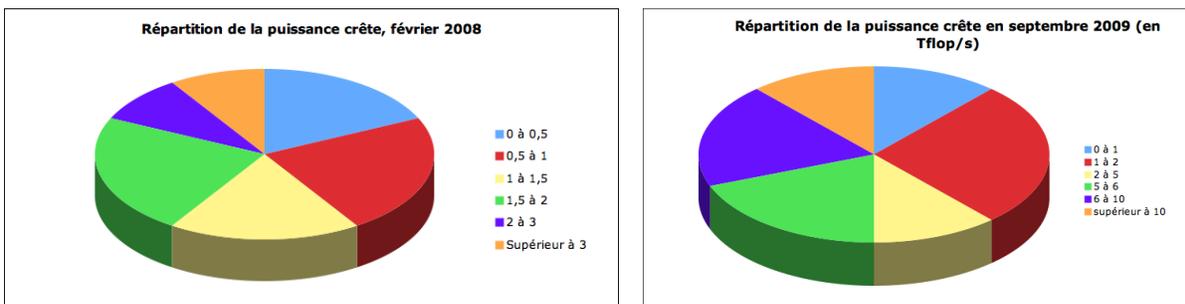
Les évolutions marquantes à noter sont :

- Une puissance moyenne en augmentation : de 1.43 TFlop/s en février 2008 à 7 TFlop/s en septembre 2009.
- Une capacité moyenne de stockage en progression : de 18.15 To en février 2008 à 113.93 To en septembre 2009.
- De nouveaux centres qui améliorent la couverture nationale.
- Des moyens humains en légère augmentation : > 70 ETP (Equivalent Temps Plein), soit une moyenne de 2.7 ETP par mésocentre, avec cependant des disparités importantes.
- Des formations proposées de plus en plus nombreuses.
- Une offre qui se diversifie : architecture type blue gene ou GPU ...

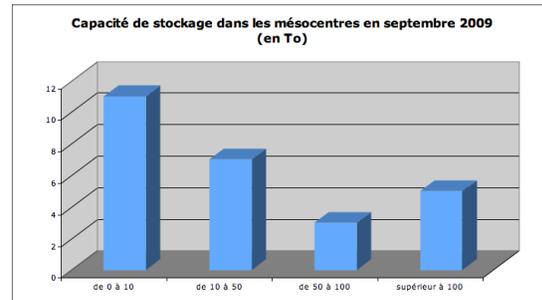
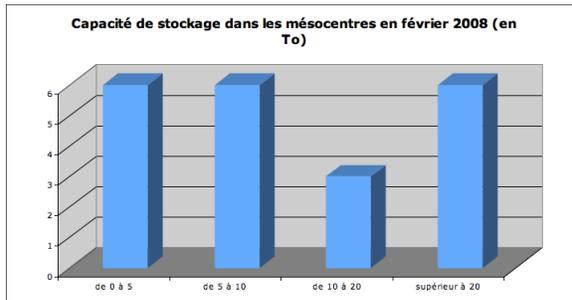
Les figures suivantes montre la localisation géographique des mésocentres en septembre 2009 (en rouge, les centres existant déjà en 2008, en bleu les nouveaux centres ou les centres nouvellement référencés).



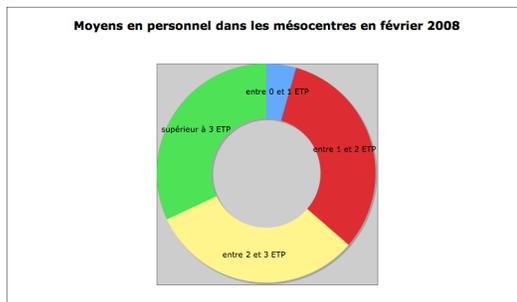
Les schémas suivant illustre l'évolution et la répartition des variations de puissances crête.



Les histogrammes suivants permettent d'illustrer l'évolution en terme de capacités de stockage.



L'évolution en terme de moyens humains est présentée par les diagrammes ci-dessous.



Le tableau suivant rappelle les principales caractéristiques de l'ensemble des mésocentres recensés. Les détails pourront être trouvés dans le rapport détaillé qui accompagne ce document.

Région	Nom, localisation	Puissance crête	Stockage	Personnel ETP
Alsace	Direction Informatique, Université de Strasbourg	9 TFlop/s	10 To	3
Aquitaine	M3PEC, Bordeaux	1.5 TFlop/s	4 To	2
Auvergne	Auvergrid, Clermont Ferrand	7.2 TFlop/s	250 To	4
Bourgogne	CRI-CCUB, Dijon	6 TFlop/s	93 To	2.5
Bretagne	GenOuest, Rennes	1.6 TFlop/s	28 To	3 permanents + 3 CDD
Centre	CCSC, Orléans	4 TFlop/s	5.4 To	1
Champagne-Ardennes	ROMEO II, Reims	5 TFlop/s	4 To	2.5
Franche-Comté	Mesocomté, Besançon	5.5 TFlop/s	24 To	3
Haute-Normandie	CRIHAN, Rouen	1.77 TFlop/s	20 To	2
Ile de France	CEMAG, Paris ENS	1.74 TFlop/s	34 To	0.3

Région	Nom, localisation	Puissance crête	Stockage	Personnel ETP
Ile de France	GRIF, Orsay + autres sites	45 TFlop/s	1700 To	12
Ile de France	Méso-centre informatique d'Ile de France sud , Orsay	1TFlop/s	5 To	1.5
Ile de France	S-CAPAD, Paris IPGP	5 TFlop/s	52.75 To	2
Ile de France	SIR-UCP, Cergy Pontoise	4 TFlop/s	8.5 To	2
Limousin	CALI, Limoges	1.3 TFlop/s	5 To	charge supplémentaire pour 2 ing.
Midi-Pyrénées	CICT-CALMIP, Toulouse	1.5 TFlop/s	6.6 To	3
Midi-Pyrénées	GenoToul, Toulouse	3 TFlop/s	50 To	4.3
Nord Pas de Calais	Calcul Intensif à l'USTL, Lille	5.5 TFlop/s	6.5 To	1
Pays de la Loire	CC IPL, Nantes	1.8 TFlop/s	3 To	1.7
Picardie	MeCS (Modélisation et Calcul Scientifique), Amiens	0.5 TFlop/s	12 To	1 (2010)
Provence Alpes Côte d'Azur	CRIMSON, Nice	8 TFlop/s	9 To	3
Rhône-Alpes	CIMENT, Grenoble	18 TFlop/s	47 To	4
Rhône-Alpes	FLCHP, Lyon	6.6 TFlop/s	140 To	3.55
Rhône-Alpes	MUST, Chambéry	5.5 TFlop/s	200 To	3
Régions d'Outre Mer	C3I (Centre Commun de Calcul Intensif), Pointe à Pitre	0.91 TFlop/s	6 To	1
Régions d'Outre Mer	CCUR (Centre de Calcul de l'Université de la Réunion), Saint Denis	0.22 TFlop/s	5 To	1