Cinq années de mutualisation des ressources en calcul scientifique au PSMN de l'ENS-Lyon

Hervé Gilquin

1 octobre 2012

Sommaire de la présentation

- Historique.
- Personnels
- Infrastructures.
- Fonctionnement.
- Exemples.
- Pérennisation.

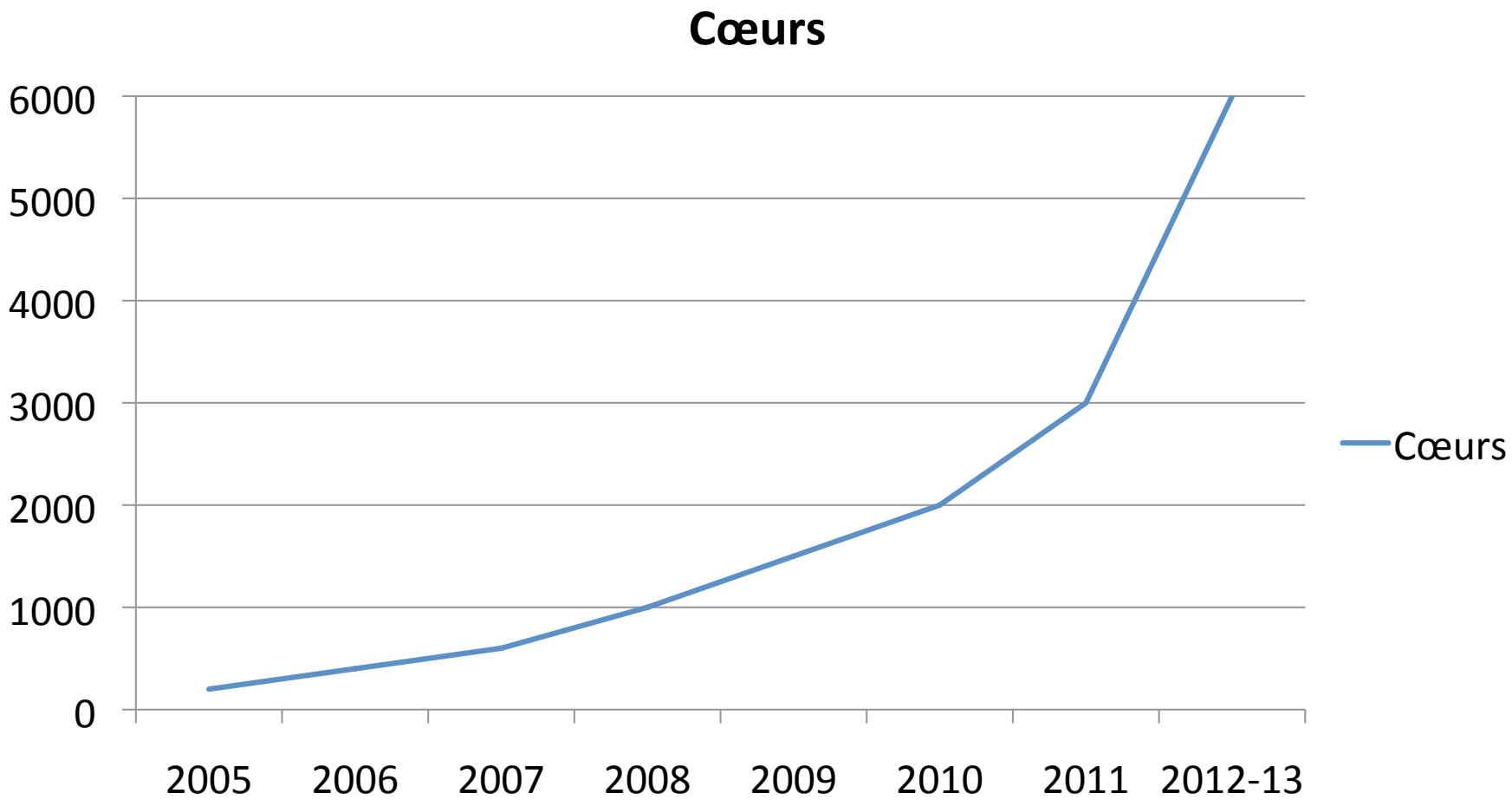
Historique

- 1993 — Création du PSMN.
- 1994 – 2002 — Structure pluri-formation.
- 2002 – 2010 — Structure Fédérative de Recherche « Fédération Lyonnaise de Calcul Haute Performance ».
- 2011 – 2015 — SFR « Fédération Lyonnaise de Modélisation et Science Numérique ».
- Depuis 2007 — Mutualisation au sein de l'ENS Lyon.
- 2007 – 2013 — PER « Calcul Intensif en Rhône-Alpes ».
- Depuis 2012 — Financement GENCI Equip@Meso (EQUIPEX)

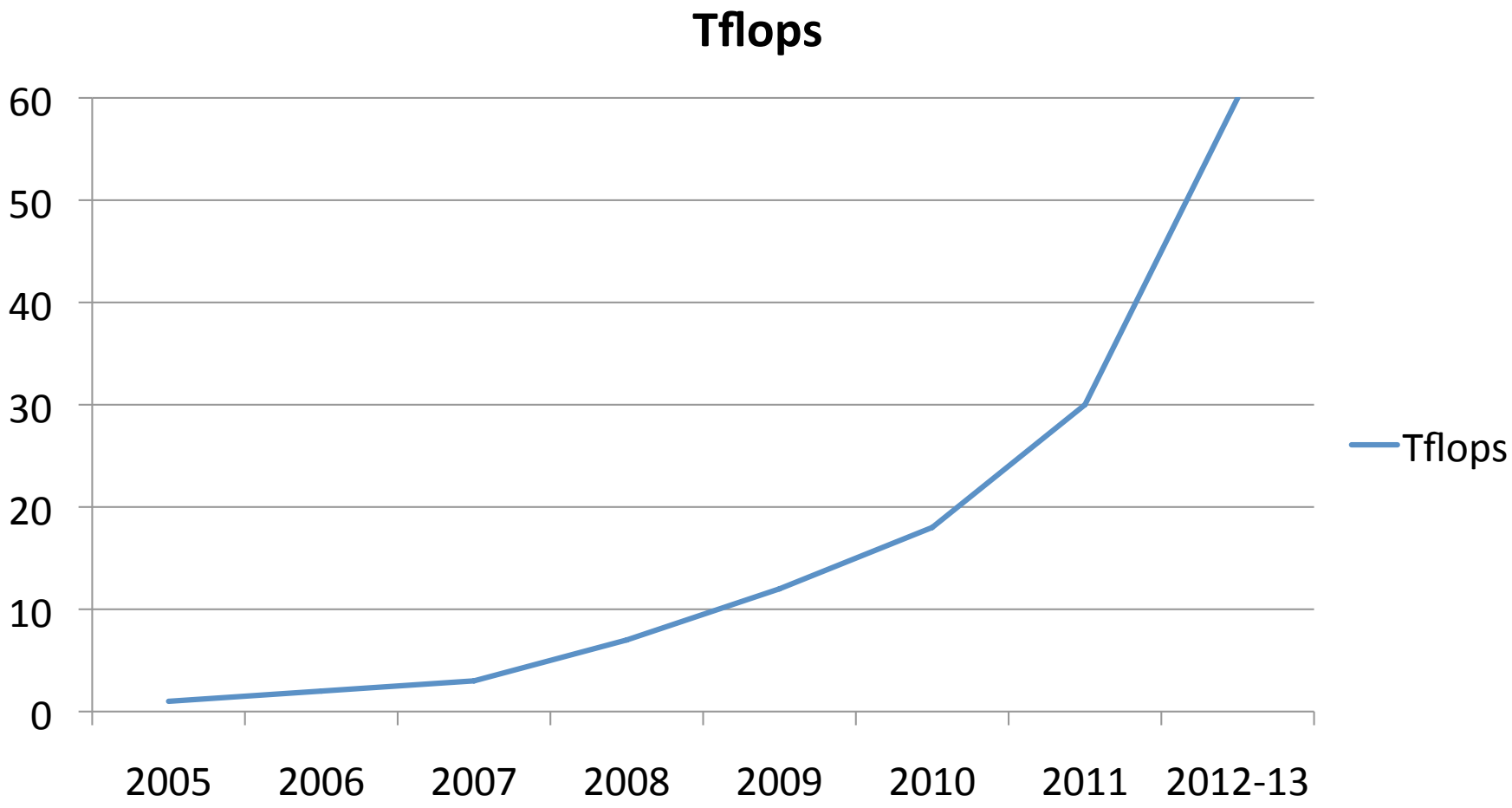
Evolution des ressources du PSMN

• 1993	5 serveurs	8 procs	3 labs	6 ut
• 2000	2 serveurs	48 procs	7 labs	20 ut
• 2005	100 serveurs	200 procs	12labs	80 ut
• 2007	150 serveurs	600 cœurs	12 labs	100 ut
• 2008	200 serveurs	1000 cœurs	15 labs	120 ut
• 2009	250 serveurs	1500 cœurs	15 labs	130 ut
• 2010	300 serveurs	2000 cœurs	15 labs	150 ut
• 2011	350 serveurs	3000 cœurs	15+ labs	150 ut
• 2012-13	500 serveurs	5-6000 cœurs	30+ labs	300 ut

Évolution du # de coeurs



Évolution de la puissance



Equipe d'exploitation

- (1993) 1 IR Calcul Scientifique 80% CNRS
- (2010) 1 IR Calcul Scientifique 50% ENS
- (1993) 1 IR Systèmes et Réseau 20% CNRS
- (2011) 1 Technicien Systèmes et Réseau ENS
- (2002) 1 secrétaire 5% ENS

- (2012-13)1 IR 40% ENS

Equipe d’exploitation

Infrastructures

- (1993-2000) Salle Info Maths (9 m2 5Kw froid).
- (2000-2004) Petite salle spécifique (20 m2 25Kw froid).
- (2004-2009) Salle Info dans bâtiment neuf (100 m2 70Kw froid).
- (2009) Réaménagement (140 m2 170Kw froid).
- (2012-2013) Aménagement tranche 1 (140 m2 170Kw froid).
- (2013-2014) Aménagement tranche 2 (140 m2 170Kw froid).
- (2014) Augmentation puissance froid (140 m2 3-400Kw froid).

Fonctionnement

Organisation

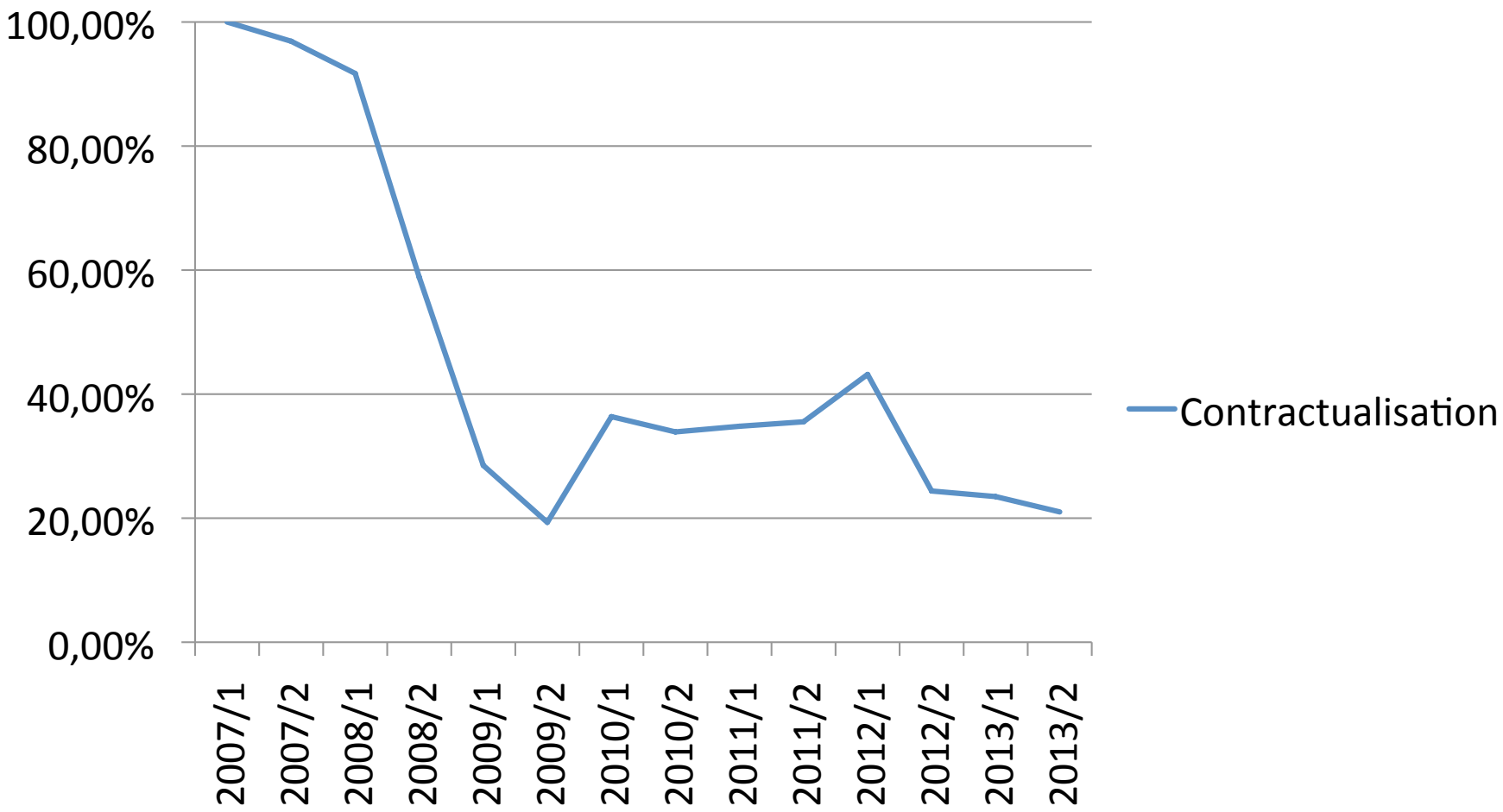
- Bureau: un représentant par équipe ou laboratoire membre.
- Ouverture de compte (représentant sous la responsabilité du directeur).
- Bureau des utilisateurs (fréquence réunion 3 fois par an).
- À la rencontre des utilisateurs:
  - 1 par mois
  - 9-10 dans l'année
- Conseil Scientifique (Coordinateur E. Lévêque).

Fonctionnement

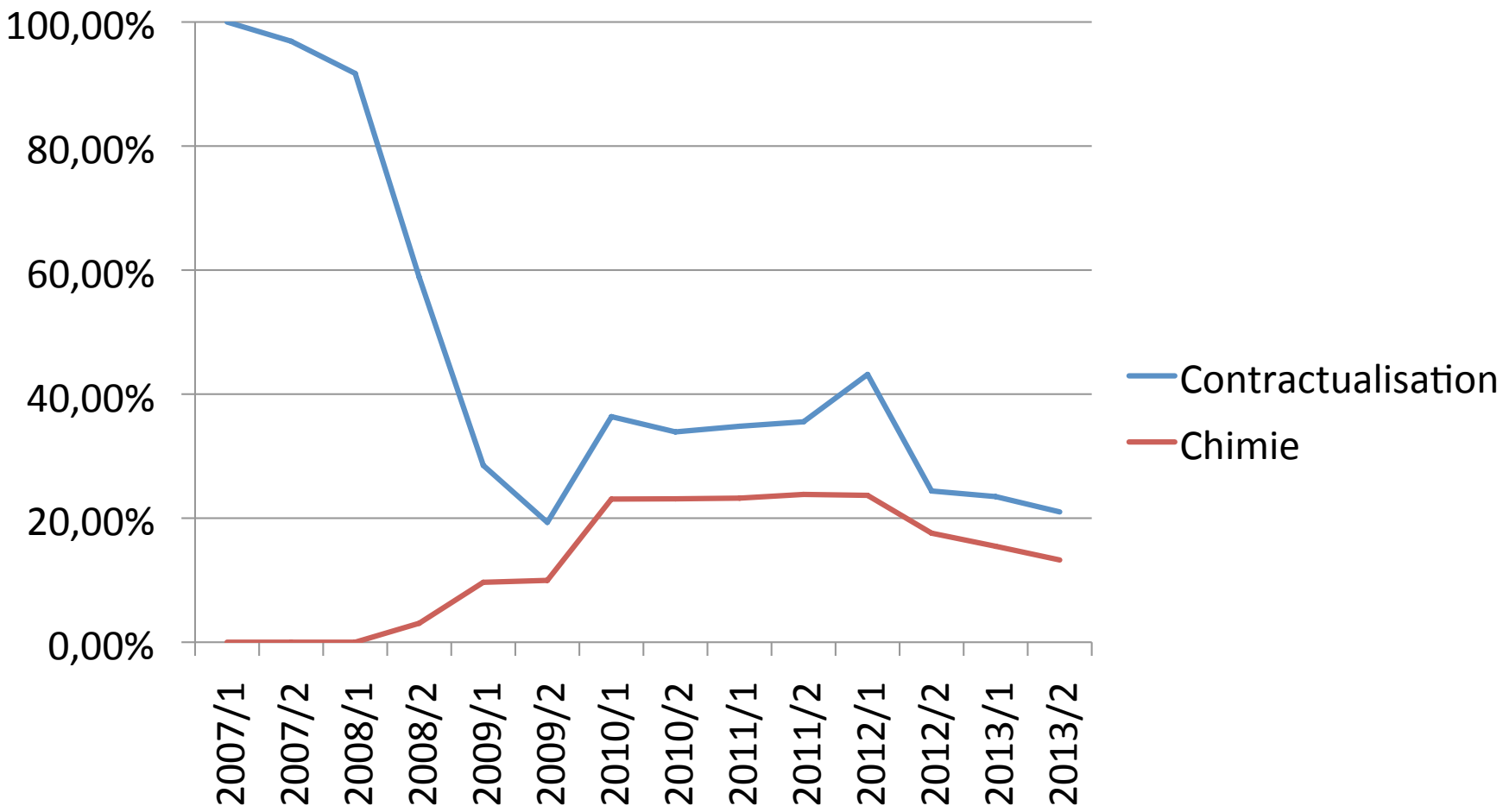
Financement

- SFR (Plan quadriennaux et quinquennaux).
- CIRA (Plan état-Région).
- Fond recherche de l'ENS-Lyon.
- Mutualisation.
  - ANR.
  - Projets Européens (ERC).
  - Contrats de recherche.
  - Contractualisation.
- Start-up.
- Equip@Meso.

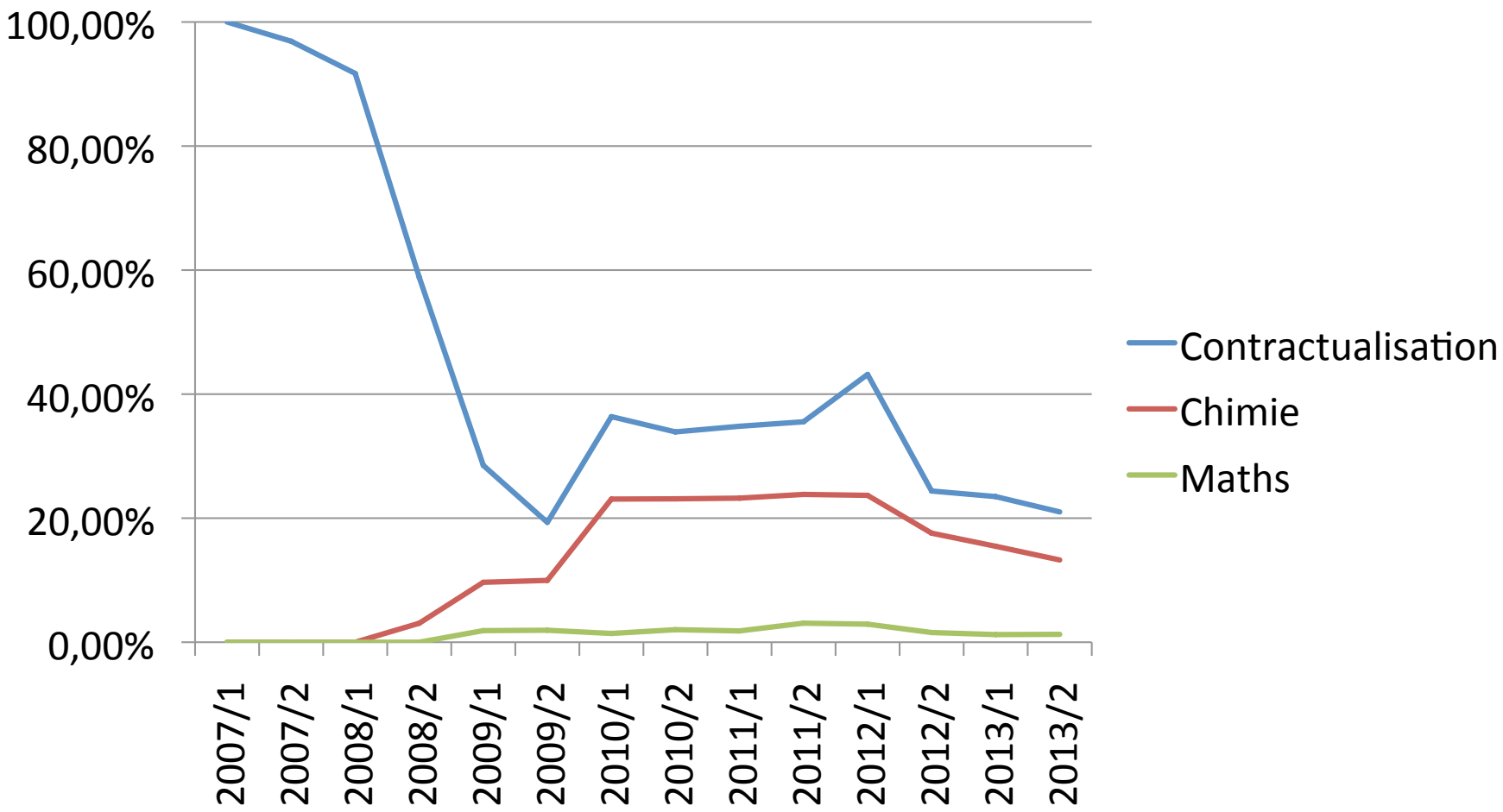
Part relative de financement



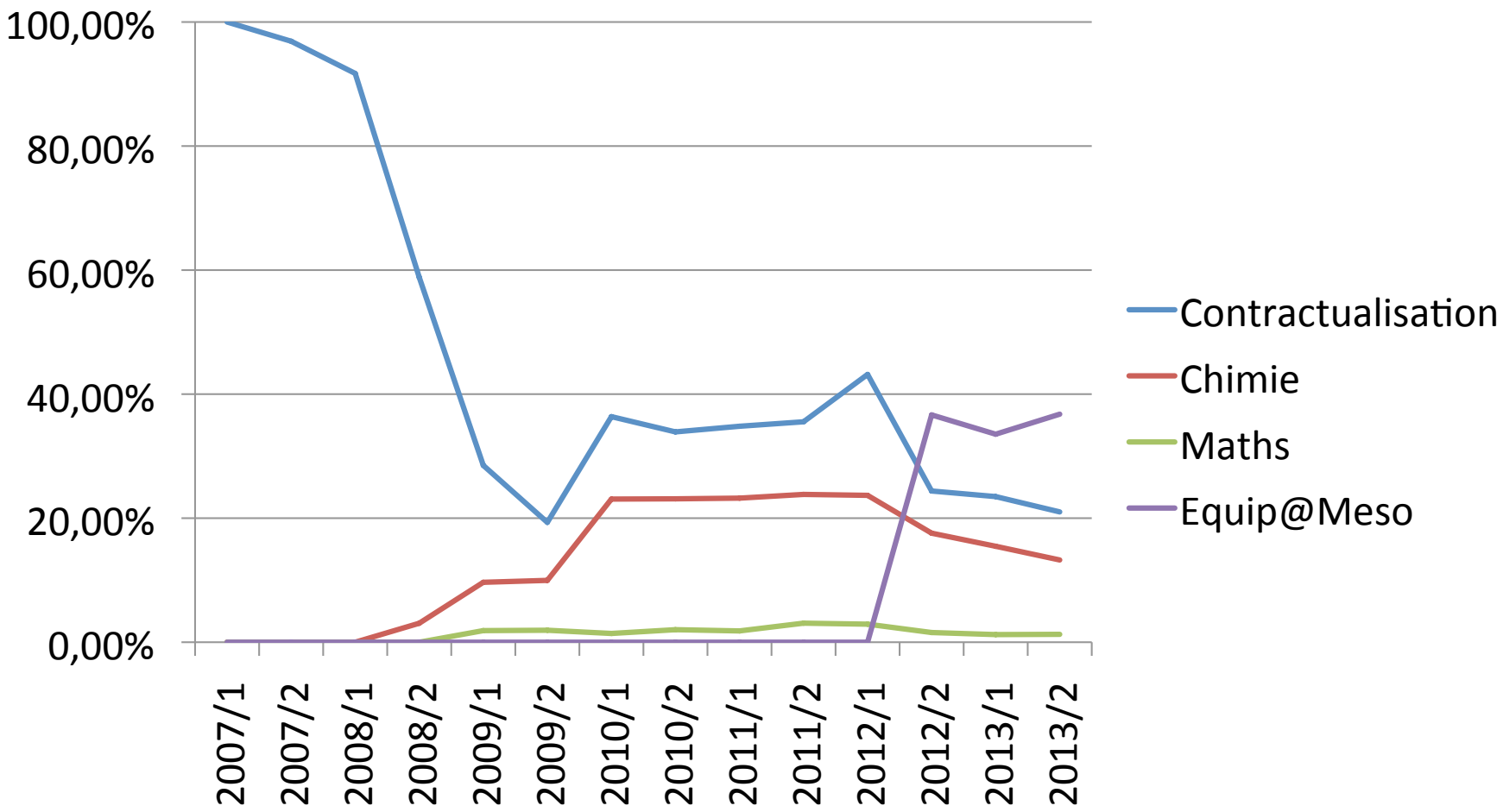
Part relative de financement



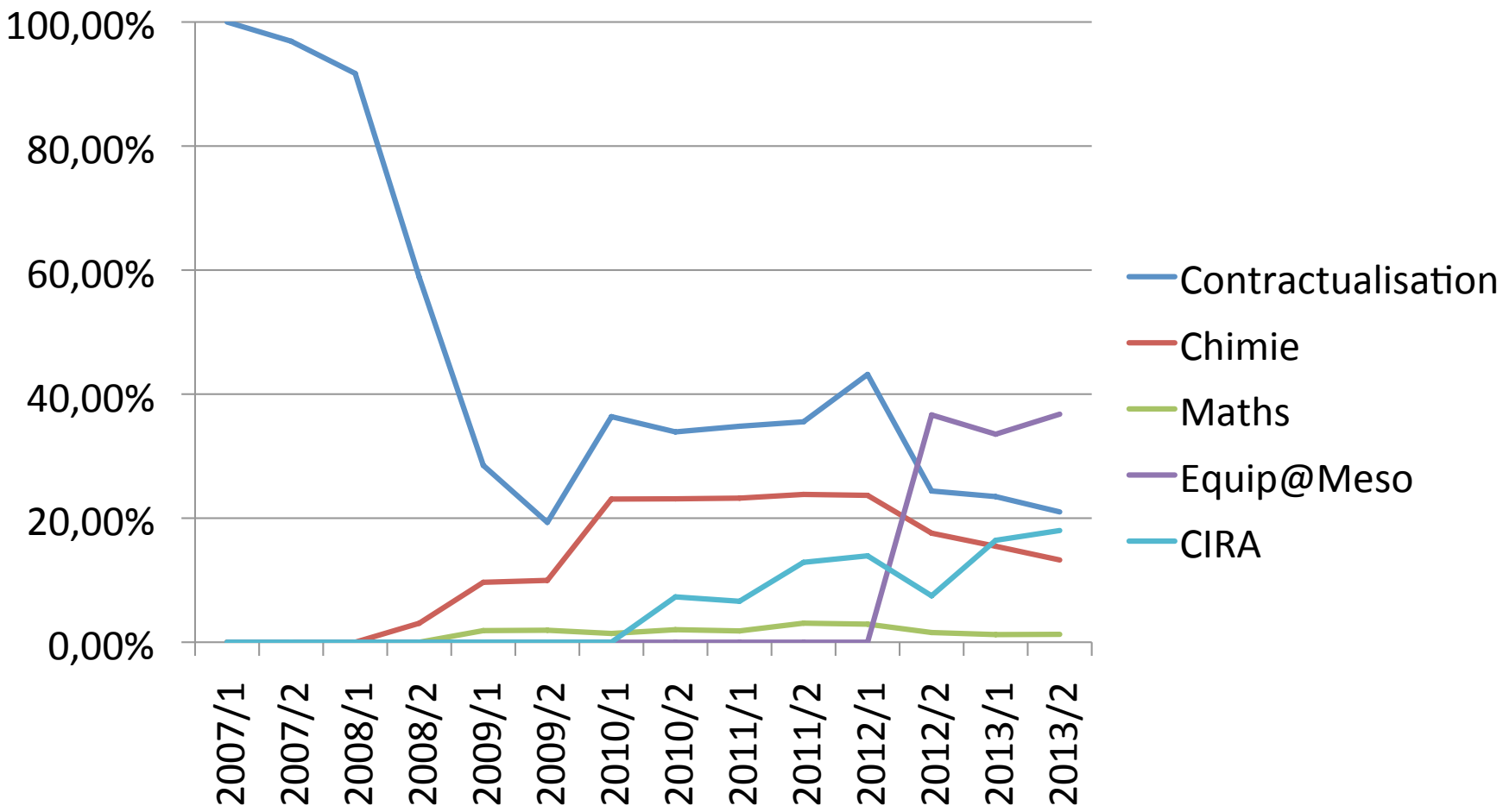
Part relative de financement



Part relative de financement



Part relative de financement



Fonctionnement

Relations internes à l'établissement

- Direction de la recherche.
- Services financiers.
- Service des marchés.
- Services techniques.

Exemples

Problématique de l'exemple 1

- Besoin de traiter des données expérimentales
  - Physique (caméras à balayage ultra rapide).
  - De 8K à 12K petits jobs par expérience.
  - Plusieurs expériences par semaine à terme.
  - Financement obtenu pour quelques serveurs (96 cœurs).
  - Evolution souhaitée vers traitement presque temps réel.
  - Lien spécifique (fibre dédiée) entre lieu de l'expérience et lieu de calcul.

Exemples

Réponse du PSMN à l'exemple 1

- Financement mutualisé (pas de serveurs spécifiques)
  - 12K jobs traités en 24h environ
  - Jusqu'à 3-400 cœurs mis à disposition pour chaque traitement.
  - Evolution vers traitement presque temps réel -> mise à disposition de plus de coeurs.
  - Lien spécifique (fibre dédiée) entre lieu de l'expérience et lieu de calcul -> en cours.

Exemples

Problématique de l'exemple 2

- Financements étagés dans le temps (chimie)
  - Besoin de réseau IB (switch 36 ports).
  - Projets sur plusieurs années.
  - Cohérence des serveurs (fréquence des procs) à un instant donné.
  - Utilisation des procs les plus performants à un instant donné.

Exemples

Réponse du PSMN à l'exemple 2

- Financement mutualisé (pas de serveurs spécifiquement dédiés)
  - Achat de clusters de 36 serveurs ou un multiple (cf switchs IB).
  - Lissage des investissements.
  - Nouvelle architecture cluster accessible tous les 18 mois.

Exemples

Problématique de l'exemple 3

- Arrivée d'une nouvelle équipe ou d'un nouveau laboratoire (biologie)
  - Financement au stade de la demande ou du projet de demande.
  - Chercheurs en cours d'installation.
  - Besoin d'accès immédiat à des serveurs et du stockage.
  - Séquenceur en cours d'installation.

Exemples

Réponse du PSMN à l'exemple 3

- La mutualisation permet
  - Aux chercheurs de commencer à travailler même sans financement.
  - D'obtenir des résultats en vue de financements éventuels.
  - De regrouper les problématiques proches.
  - Mise en commun des outils logiciels.
  - Mise en commun des achats spécifiques.

Exemples

Problématique de l'exemple 4

- Partenariat avec une Start-up (viabilité du concept)
  - Mise à disposition d'outils logiciels à d'une communauté d'utilisateur.
  - Tests de viabilité du concept.
  - Besoin d'accès à beaucoup de serveurs.
  - Financement faible dans un premier temps.
  - Développement d'outils logiciels spécifiques.

Exemples

Réponse du PSMN à l'exemple 4

- Mutualisation du financement même faible
  - Développements logiciels à la charge de la Start-up.
  - Accès à beaucoup de serveurs pour un temps cumulé limité.
  - Développements logiciels resteront utilisables par le PSMN.
  - Si viabilité prouvée poursuite du partenariat.

Exemples

Problématique de l'exemple 5

- Mise en place d'une coopération
  - Participation au développement d'un logiciel commercial.
  - Présence limitée dans le temps d'un développeur dans le laboratoire.
  - Besoin d'accès à beaucoup de serveurs pendant un temps limité.

Exemples

Réponse du PSMN à l'exemple 5

- La mutualisation a permis, en compensation d'un accès à beaucoup de serveurs pour un temps cumulé limité.
  - La mise à disposition du logiciel sur tous les serveurs du PSMN.
    - Licence illimitée
  - L' utilisation du logiciel par tous les laboratoires membres du PSMN.

Exemples

Problématique de l'exemple 6

- Besoin de quelques dizaines d'heures de calcul
  - Biologie.
  - > 48 GO de Mémoire.
  - Calcul nécessaire pour une publication.
  - Pas de ressource locale disponible.
  - Pas de financement.

Pérennisation

- Conserver la proximité et la réactivité.
- Conserver un mode de fonctionnement souple et adaptable.
- Disponibilité et écoute constante.
- Conserver l'équipe d'exploitation.
- Mutualisation des savoir-faire (des personnels).
- Aménagement des infrastructures.