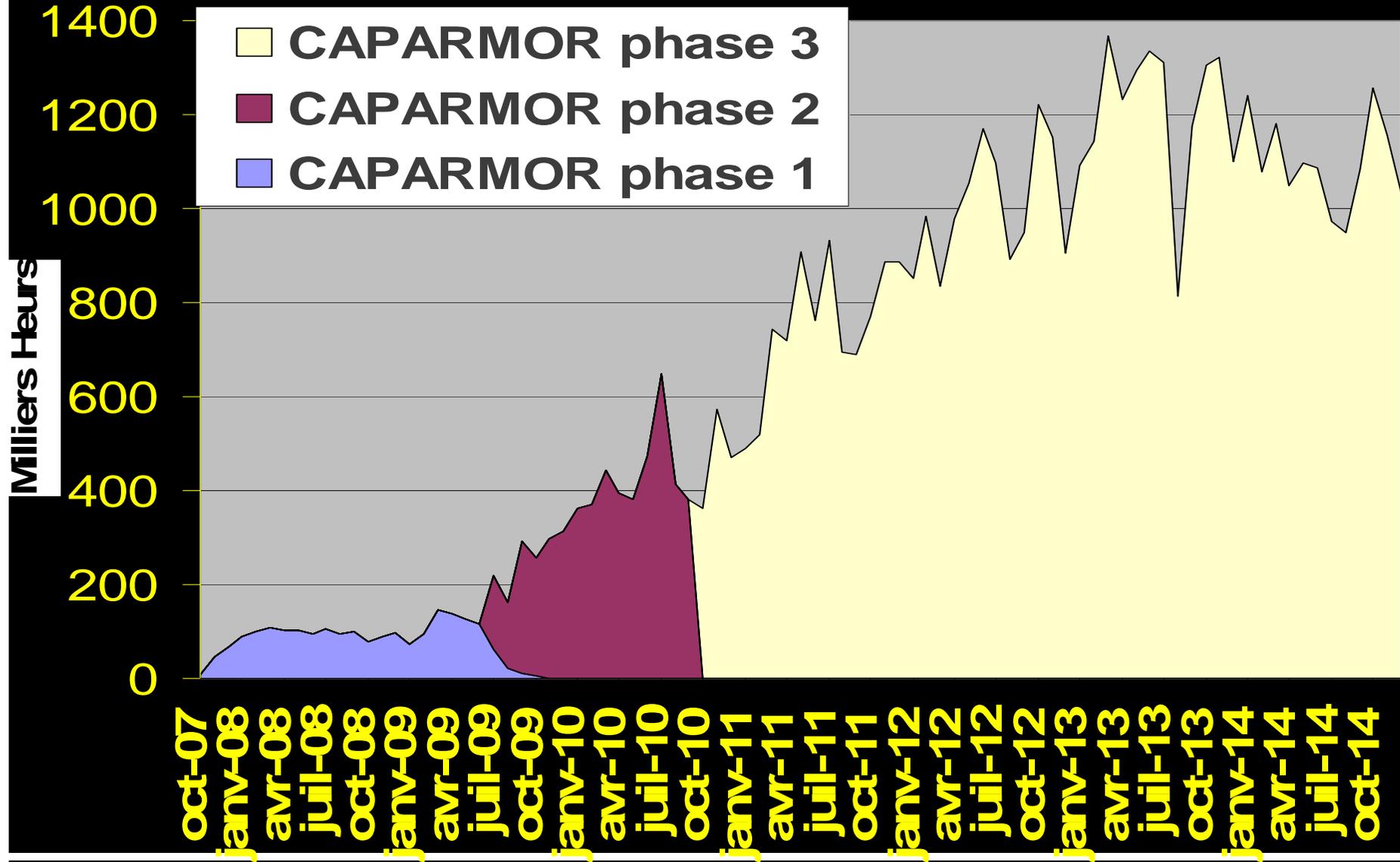


DATARMOR

13 Mars 2015

Statistiques d'usage



Datarmor: Groupe de travail

- Définition des besoins scientifiques.
- Animateur Pierre Garreau
- Composition: IFREMER, IUEM, SHOM, ENSTA Bretagne, Ecole Navale, IRD et Pôle Mer Bretagne
- Participants : des utilisateurs et des informaticiens issus de services communs ou de labos

Contenu du projet Datarmor

- Remplacement de la machine Caparmor
- Rester au niveau d'un « mésocentre » (budget constant)
- Introduire des critères de performances en accès disque (problématique BigData)
- Permettre de créer un environnement de R&D associant des chercheurs en sciences marines et des spécialistes des STIC

L'environnement du Centre de Données (thématique)

- Base de données Coriolis/physique
- CERSAT : données spatiales
- Données des campagnes océanographiques
- Système d'information halieutique
- Données d'environnement côtier
- Portail & infrastructure de données géographiques

L'environnement du Centre de Données (technique)

- SGBD relationnel ORACLE : catalogues, données ponctuelles (in-situ), méta-données (10 milliards d'enregistrements)
- Stockage disque (fichiers à plat) : données image (~700To de disque)
- Système d'archivage pérenne : sécurité des données patrimoniales (~800To d'archives)
- Cluster de portails WEB virtualisés

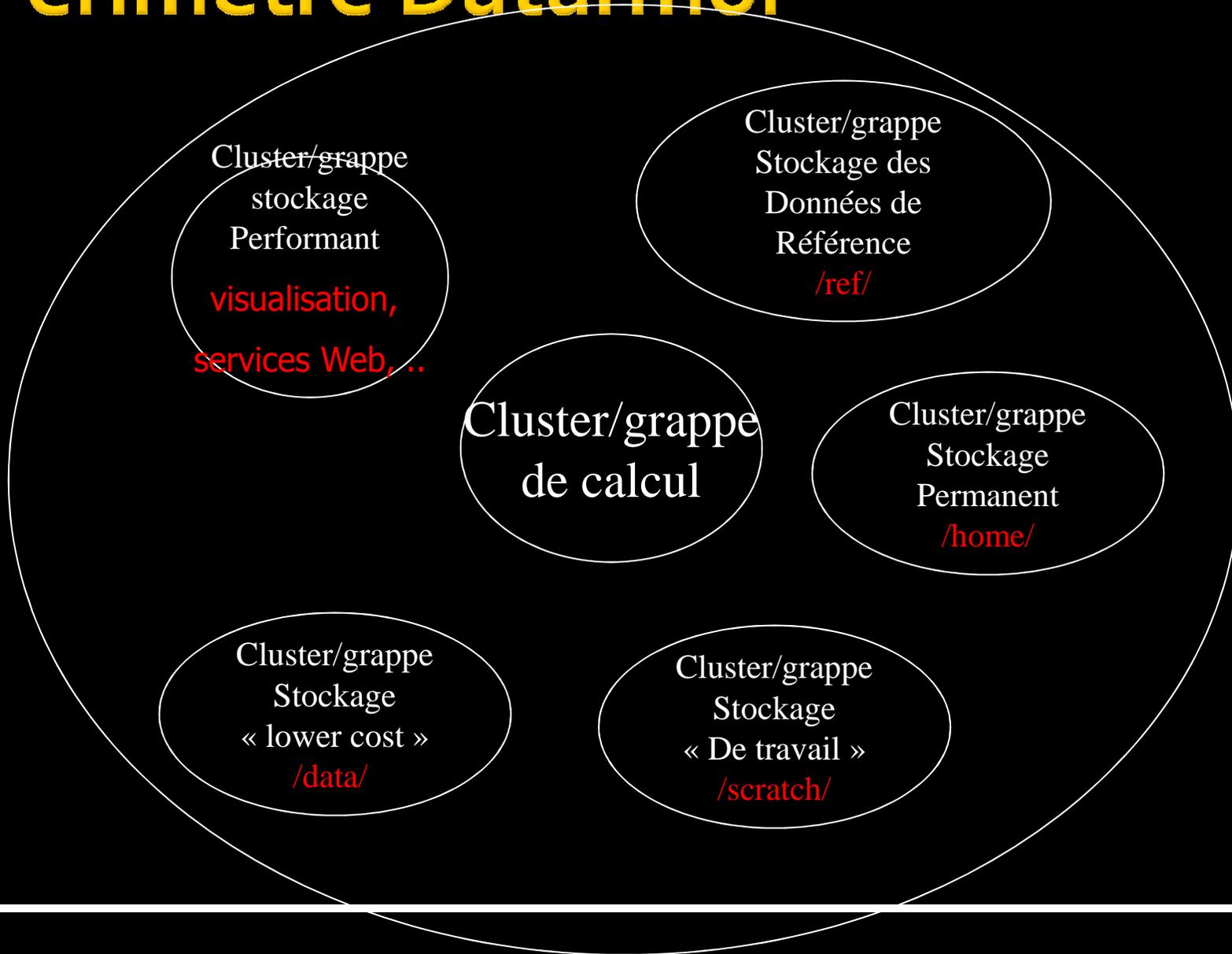
Evolution du contexte

- Augmentation régulière du volume des données due à :
 - L'automatisation des système d'observation
 - Les progrès technologiques sur les capteurs
 - La récupération de données d'opportunité
 - Pour certains traitements, les accès disque sont devenus le facteur limitant

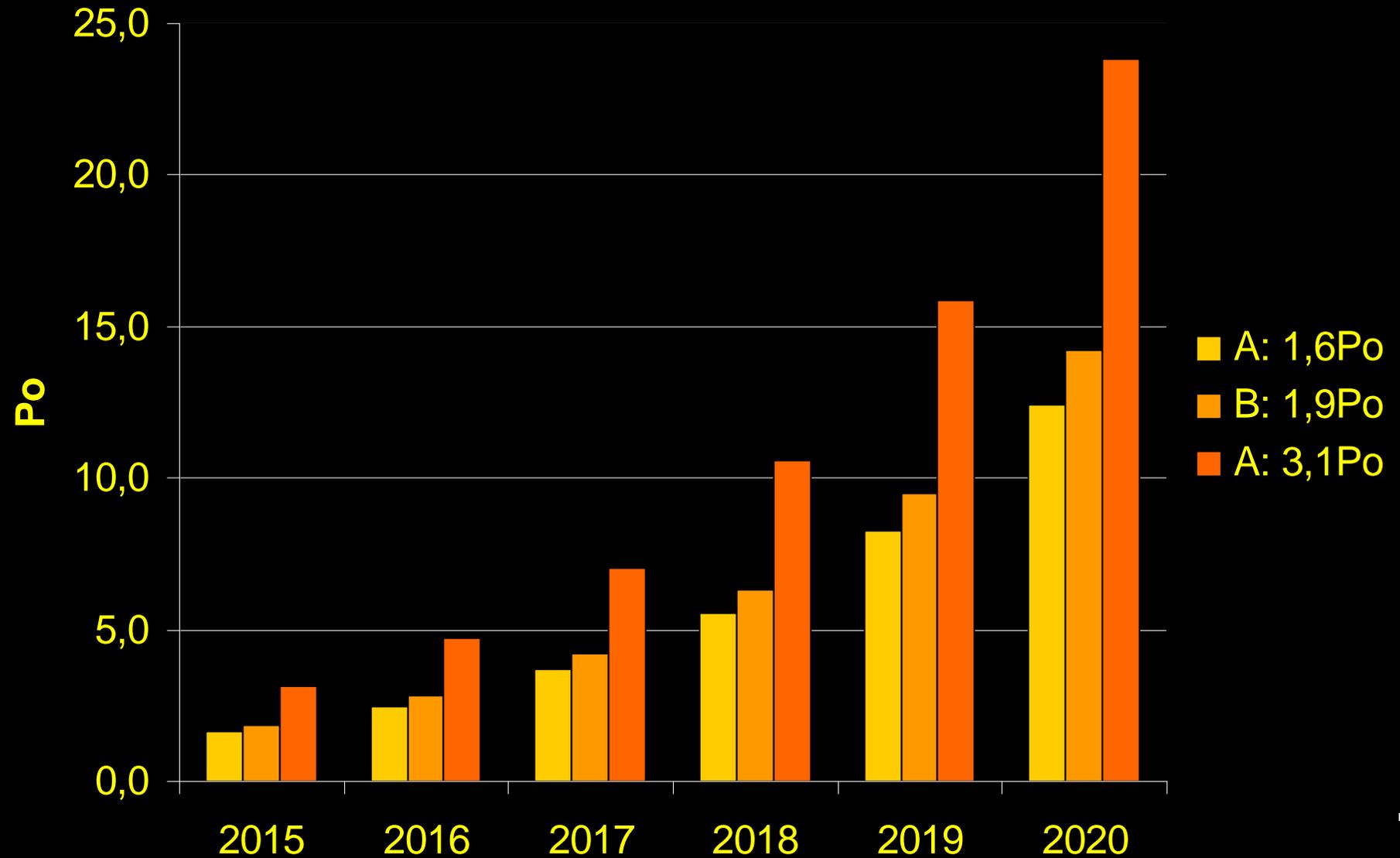
Contexte présent élargi

- Finalité « sciences marines » partagée entre les partenaires
- Proximité « réseau » des données marines
- Filière de PME sur ce domaine
- Au niveau national : pôles de données d'observation de la terre
- Pôle de données « océan »

Périmètre Datarmor



Data size estimation 2015- 2020



Les utilisateurs/bénéficiaires

- Scientifiques (pilotage scientifique)
- Professionnels (mission et financements
Ministériels ciblés)
- Startup/PME (pilotage pôle mer?) notamment
pour la partie calcul

Gouvernance prévue

- Etablir une convention d'exploitation avec les collectivités territoriales
 - But : définir les parts respectives de la recherche/activités publiques d'une part et des PME d'autre part
- Etablir des conventions avec :
 - le pôle de compétitivité Mer Bretagne et les PME
 - les établissements partenaires/utilisateurs

Gouvernance prévue

- Comité de pilotage scientifique
- définit la politique de partage des ressources (calcul/traitement, stockage, ...)
- délègue au Pôle de compétitivité la gestion d'un Comité de pilotage « entreprise » avec une autonomie de décision

Le partage des « ressources »

- Calcul/traitement :
 - essentiellement paramétrage des outils de scheduling de la configuration
- Stockage :
- Validation du statut de « données de référence »
- Arbitrage des demandes d'espaces disques
 - Espace sauvegardé
 - Espace pour diffusion Web/ftp/autres
 - Espace de travail (scratch, lower cost, ...)
 - autres

Le CPER Bretagne

- Projet Datarmor en voie d'acceptation
- En 2016 :
 - achat/installation de la machine
 - Travaux de logistique pour alimentation électrique et système de refroidissement
 - Doit permettre de démonter Caparmor en 2016
 - 2017-2020 budget annuel pour upgrade

Appel d'offres

- Pour une installation en 2016, il faut donc armer un appel d'offres dès à présent
- Appel d'offres avec « dialogue compétitif » (en 2 phases)
- On s'adresse à des « intégrateurs » capables de tout mettre en place (1 seul lot)
- Une tranche ferme, puis des tranches optionnelles

Les benchmarks

- Programmes « test » fournis aux candidats pendant l'appel d'offres
- Ils devront s'engager contractuellement sur des niveaux de performance
- → Appel à propositions de benchmarks par les utilisateurs

Appel d'offres

- Phase 1 : cahier des charges fonctionnelles
 - chiffrages approximatifs des volumes
 - laisser les candidats proposer des solutions ouvertes
- Dialogue compétitif : affiner la solution avec les meilleurs candidats
- Phase 2 : cibler des technologies et des quantifications précises avec un nombre restreint de candidats

Time Line

