« Évolution des intergiciels & outils informatiques »

Rencontres de Branville

Vincent Garonne

Laboratoire de l'Accélérateur Linéaire

22 mai 2006

Présentation du service informatique/Devdu

Sommaire

Les tendances et futures problématiques

Résumé des tendances (voir présentation M. Jouvin)

- Vers des infrastructures massivement parallèles/distribuées
  - Ressources + petites, + puissantes, e.g palm, multi-cœur
  - $\nearrow$ et omniprésence du réseau, e.g. WiFi Max
  - Données/informations massives et largement distribuées

Les possibles questions dans 10 ans...

1. Comment utiliser ces infrastructures pour les futures expériences de physique des particules, e.g. ILC ?
2. Quelles seront les nouveaux modèles de travail, techniques et outils dans ces environnements ?
3. But de cette présentation : Essayer d'amener des éléments de réponses, poser des questions et esquisser le rôle futur du LAL

Les tendances et futures problématiques

Résumé des tendances (voir présentation M. Jouvin)

- Vers des infrastructures massivement parallèles/distribuées
  - Ressources + petites, + puissantes, e.g palm, multi-cœur
  - $\nearrow$ et omniprésence du réseau, e.g. WiFi Max
  - Données/informations massives et largement distribuées

Les possibles questions dans 10 ans...

1. Comment utiliser ces infrastructures pour les futures expériences de physique des particules, e.g. ILC ?
2. Quelles seront les nouveaux modèles de travail, techniques et outils dans ces environnements ?
3. Méthode : « Pour voir le futur, il faut regarder derrière soi »

Les différentes générations de grilles

Les grilles de $1^{ére}$ génération : DataGrid, Nordugrid et GRID3

- Approche client/serveur, basé sur le toolkit Globus 2
- Identification des principales fonctionnalités et lacunes

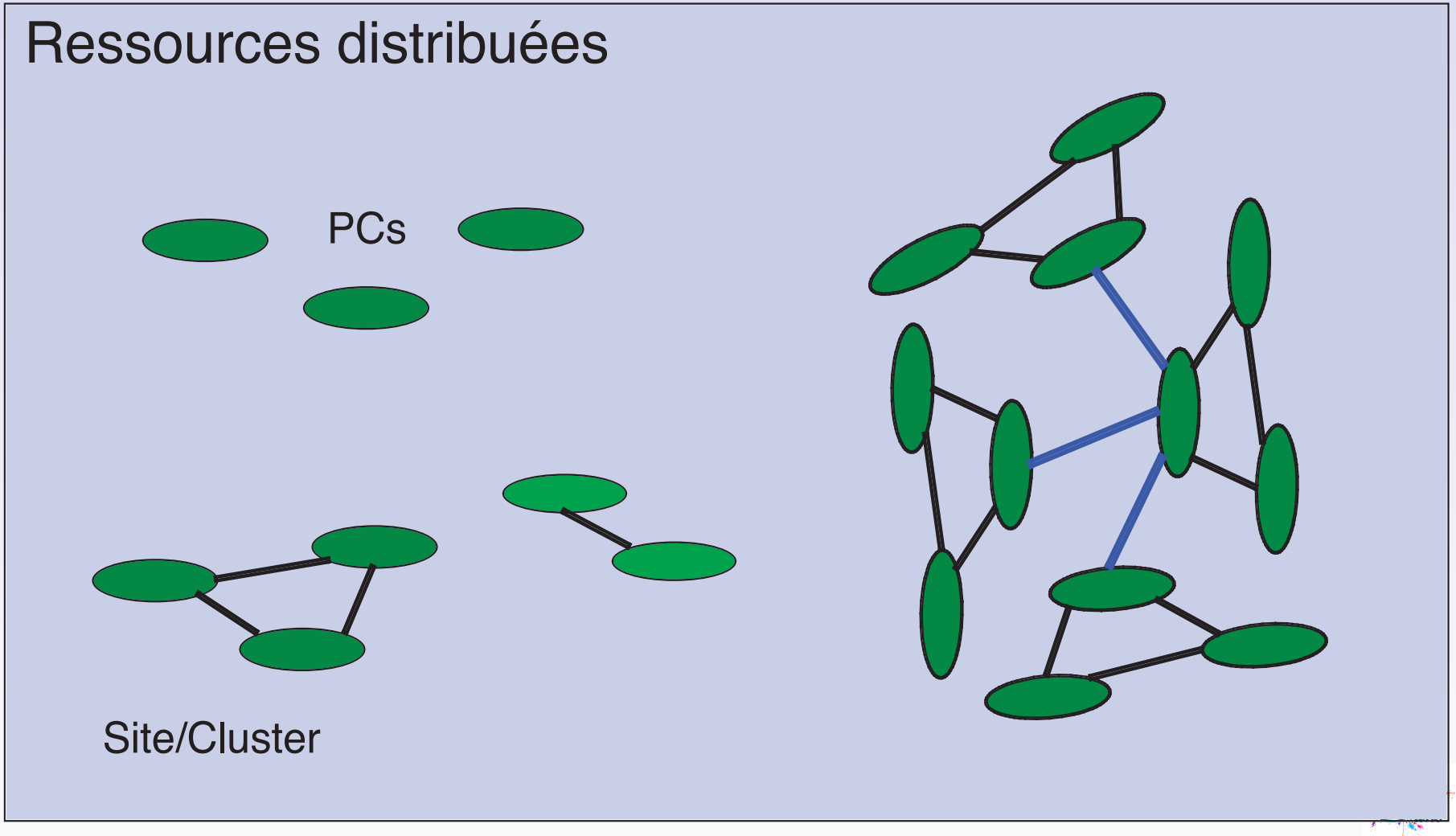
Les grilles de $2^{ème}$ génération : EGEE, ARC et OSG

- Architecture Orientée Services, basée sur OGSI/OGSA, ARDA
- Grilles de production

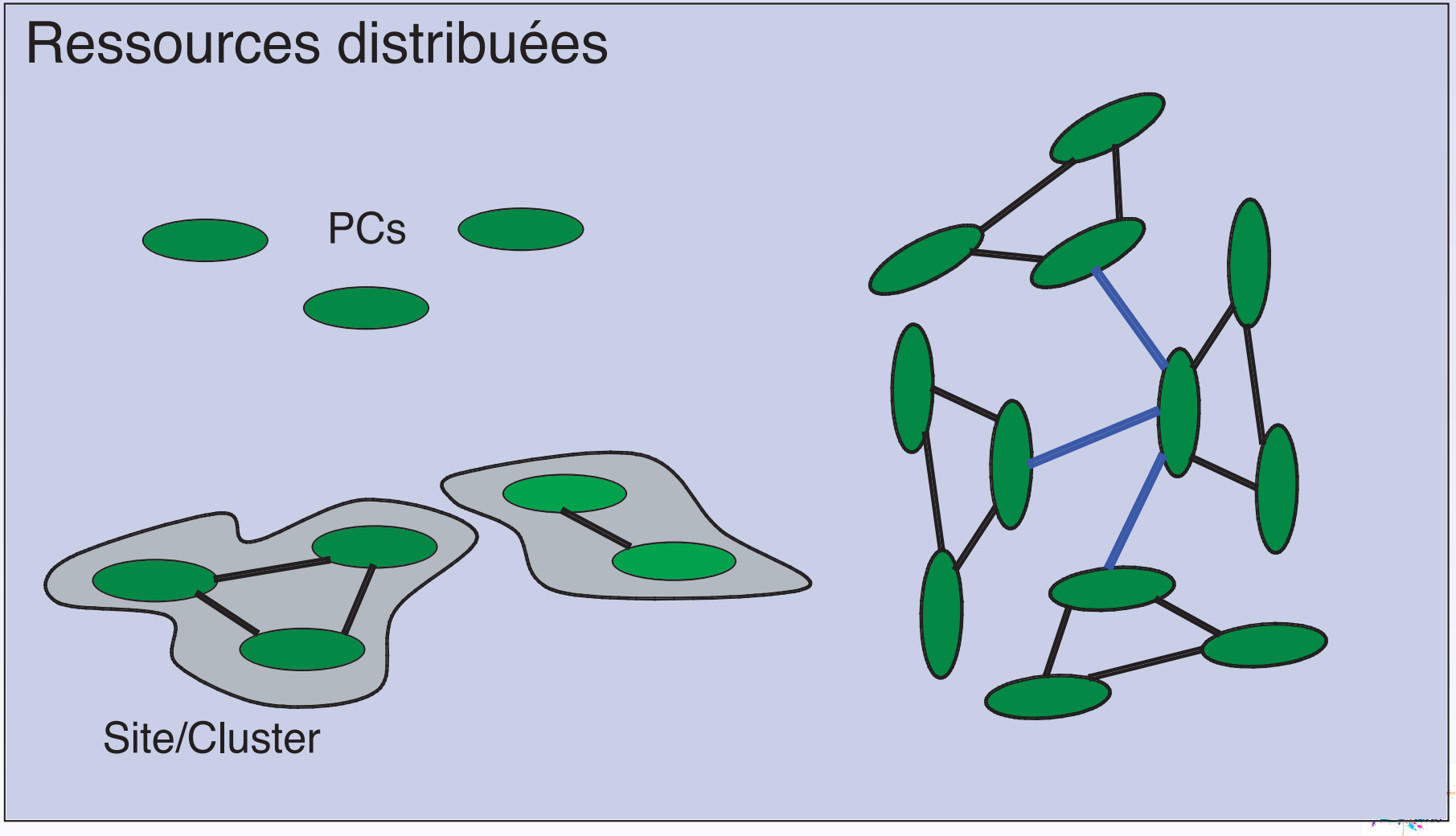
Les grilles communautaires : Alien, DIRAC et PANDA

- Développés par les expériences de physique pour leurs besoins
- Grille légère, en surcouche des grilles de production

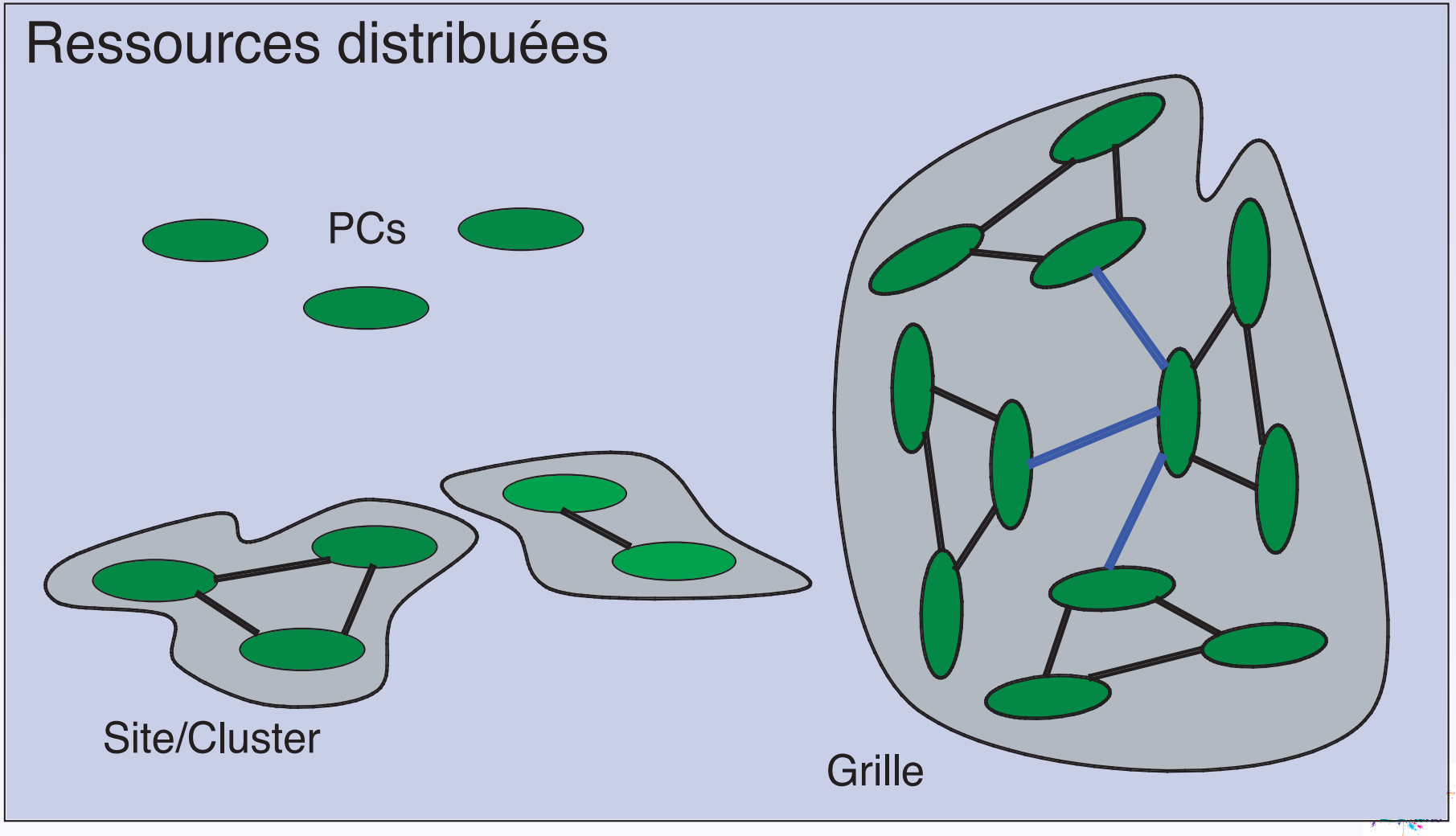
Les ressources physiques



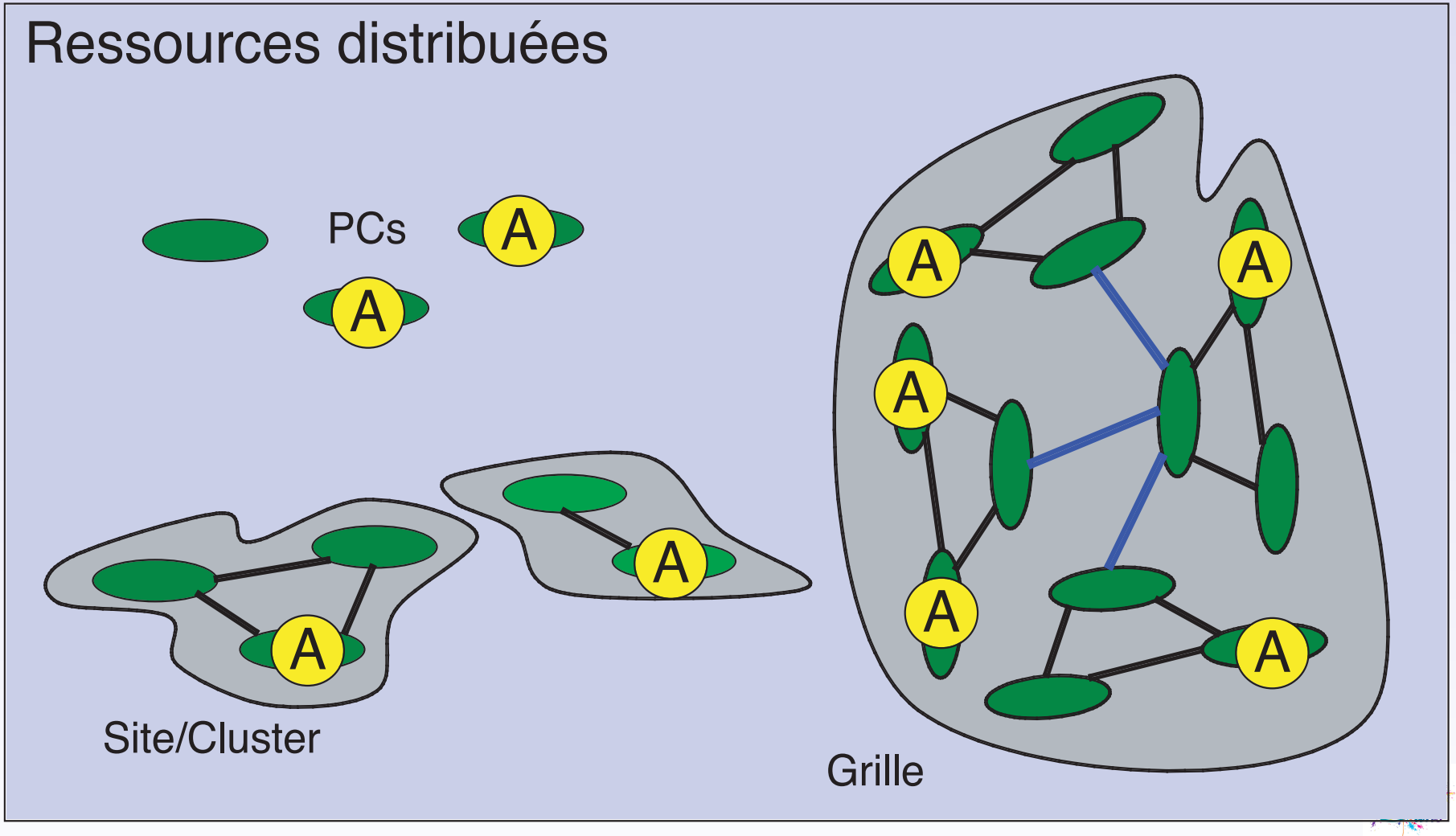
La couche logicielle locale



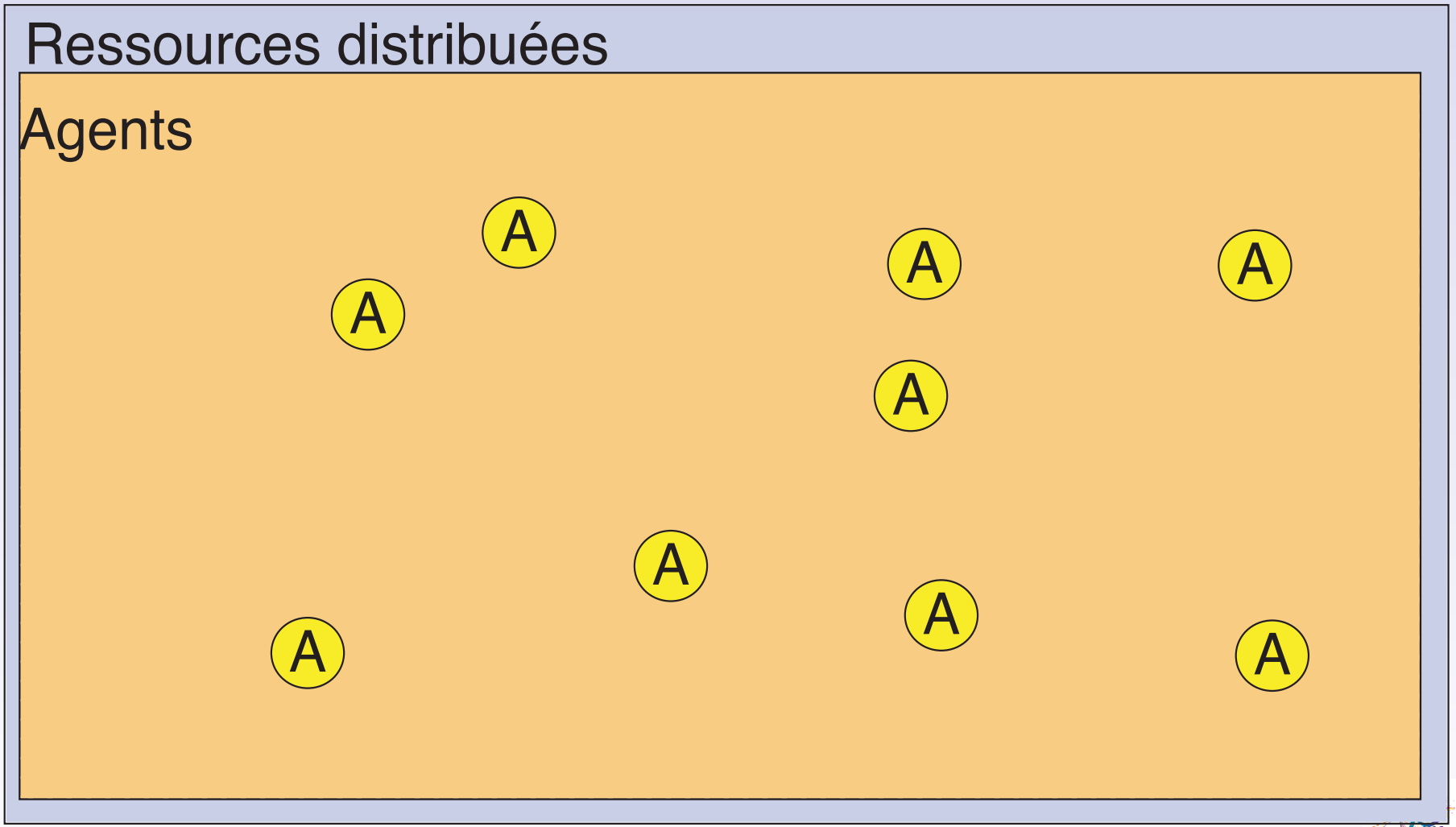
La couche logicielle globale : les grilles



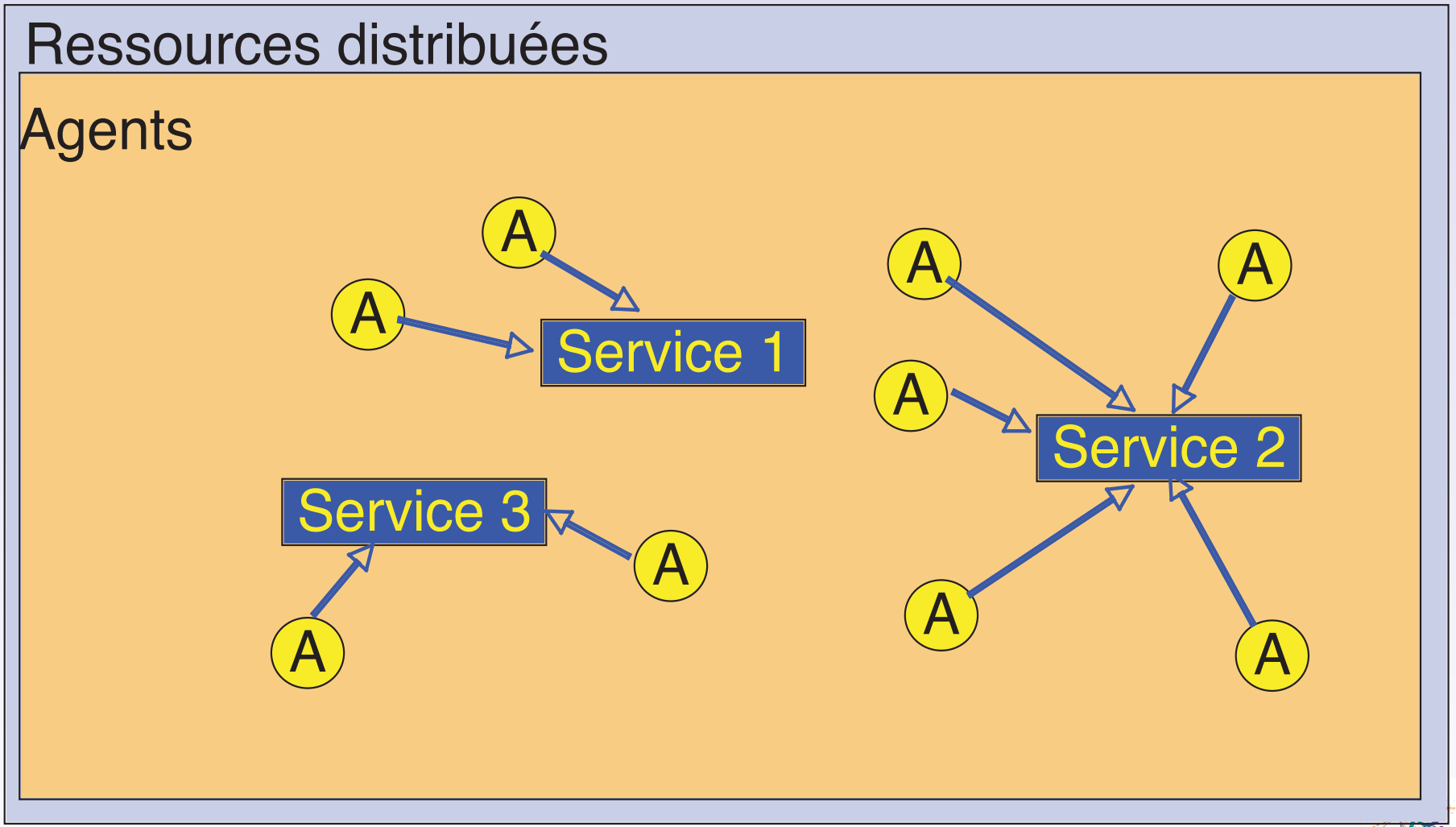
Les agents


Ressources distribuées
PCs
A
A
A
A
A
A
A
A
A
Site/Cluster
Grille

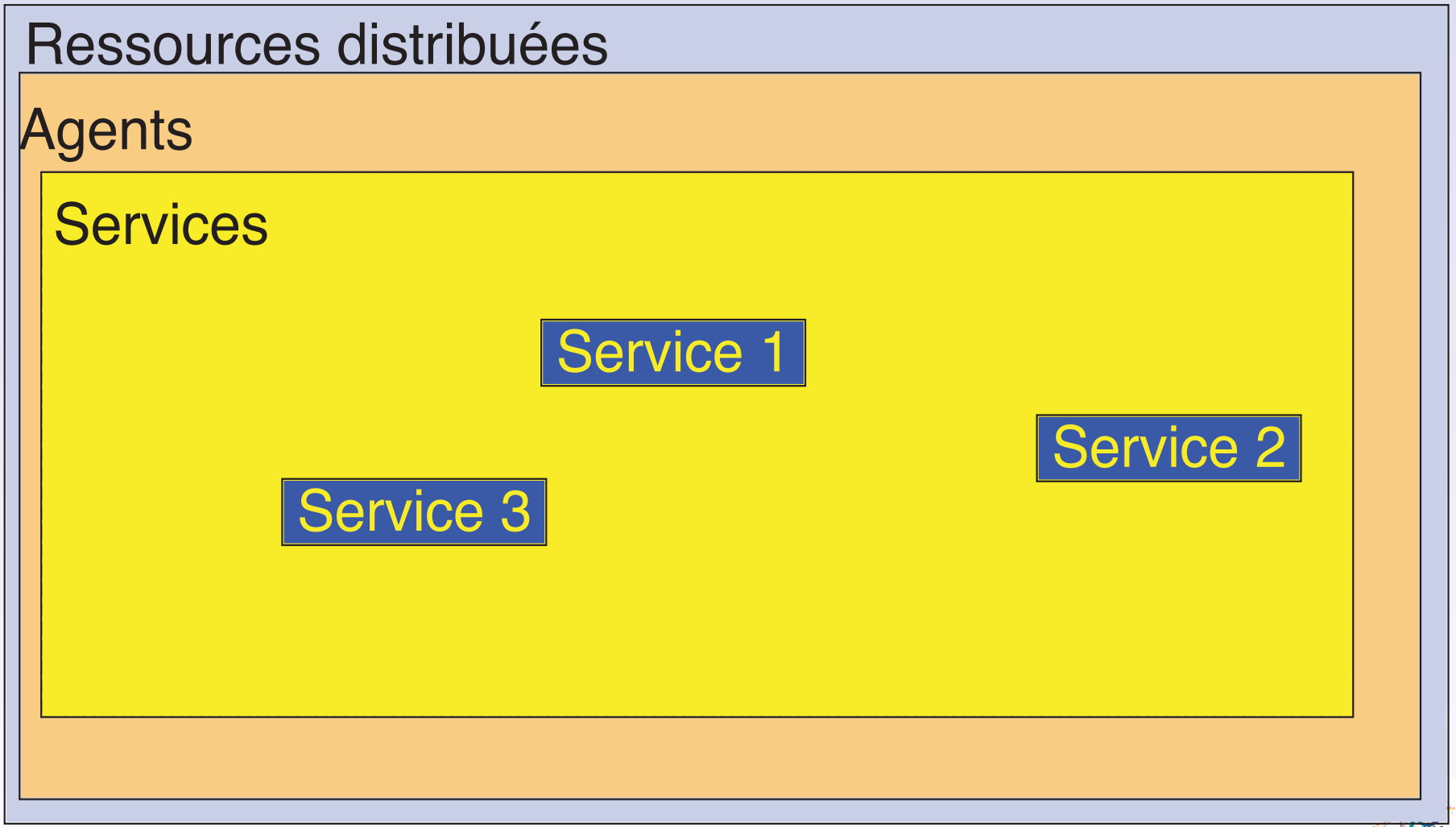
La virtualisation communautaire



Les services



L'architecture orientée service



Les prochaines générations de grilles

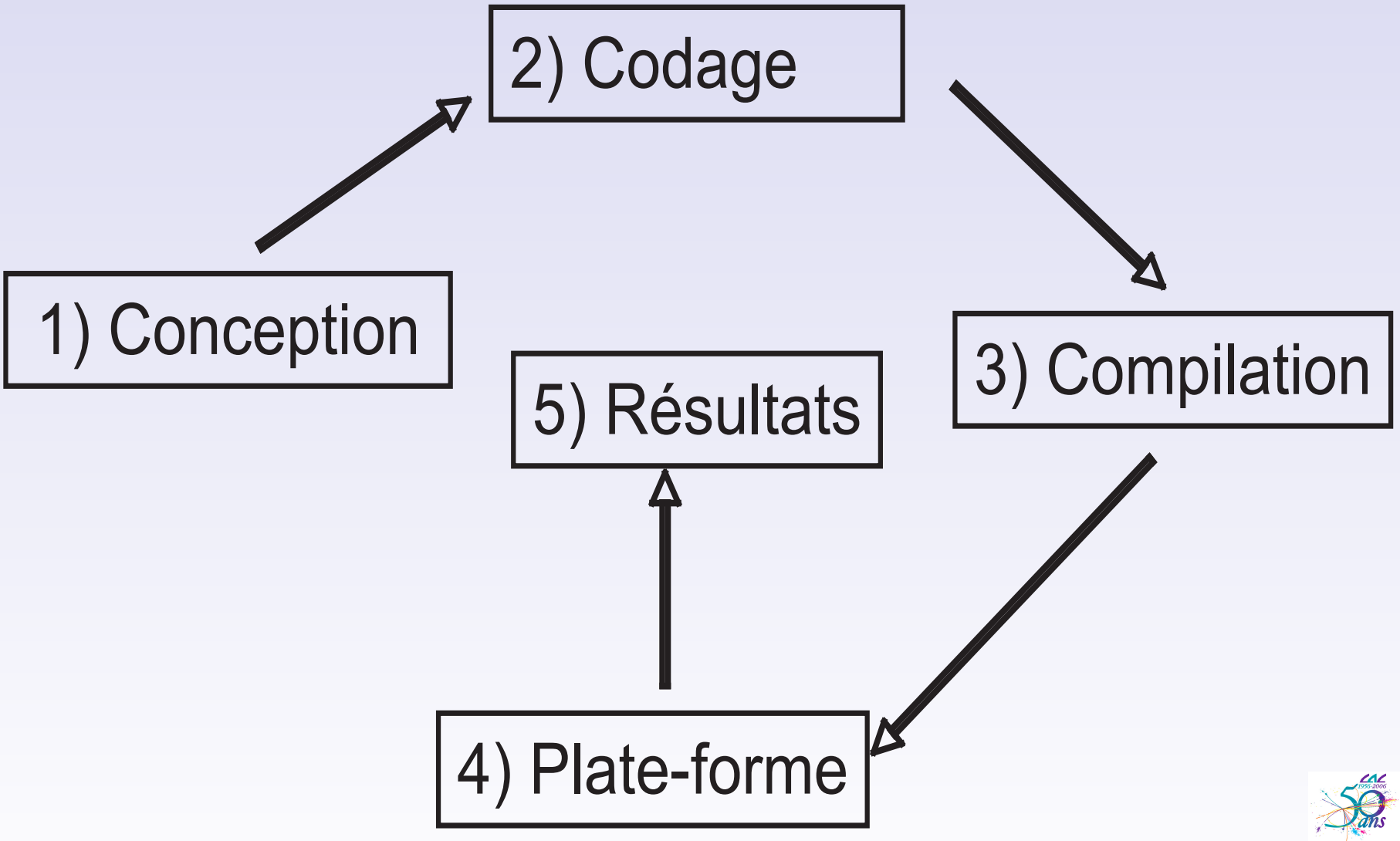
Les prédictions de Mme Irma

- Dans le court terme :
  - Convergence des grilles de productions et légères
  - Plus de services identifiés, standardisés et incorporés dans l'infrastructure générique de base
- Dans le long terme :
  - le web/grille sémantique, les systèmes multi-agents, pervasifs et multi-threadés

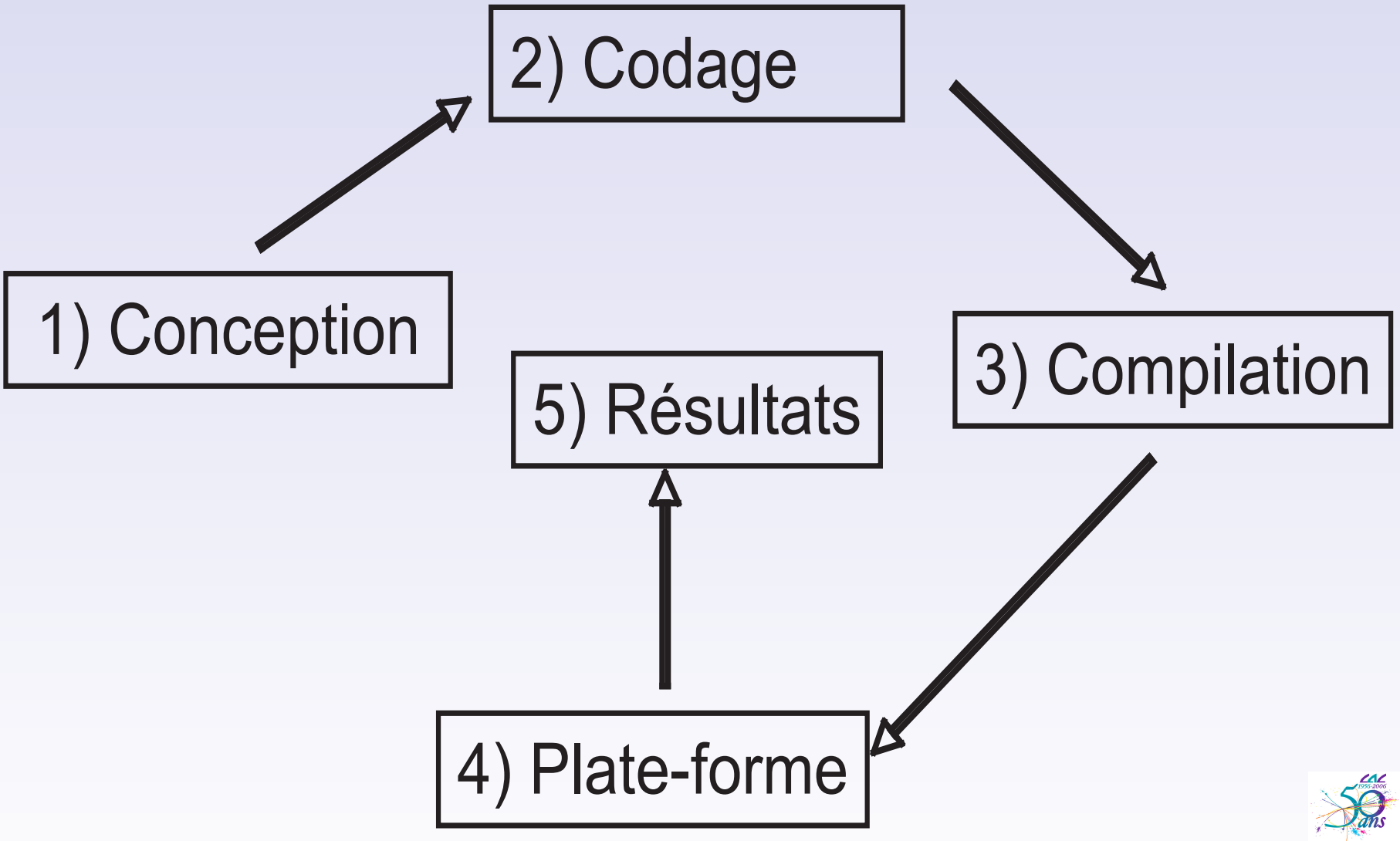
oui très bien mais : quel software pour 2010-2015 en HEP ?

- Comment faire de la production et de l'analyse de données ?
- Comment coupler les codes applicatifs ?
- Comment gérer les données, résultats et informations ?

Cycle de développement actuel



Adéquation dans 10 ans ?



Détails I

La conception

- Les concepts de l'O.O (UML)
- Définition des structures de données et algorithmes
- Découpages des classes
- Beaucoup de choses déjà faites

L'implémentation

- Quel langage ? C++, Csharp, java, python ou ? ? ?
- Possibilité d'utiliser différents langages
- La réutilisabilité du code existant : Geant4

Détails II

La compilation

- gcc8 ?

La plate-forme

- Développement multi-plateforme
- Contraintes plus importantes sur la gestion de configuration (cmt)

Détails III

execution

- De plus grande contraintes du à l'environnement distribué et hétérogène
- Favorise une plus grande modularité du code
  - Quid de Root ? - monolithique par philosophie, boot un premier pas ?
- Distribution du soft
- Surveillance du déroulement de l'application

Détails IV

Les outils

- Développement : Visual, Xcode + shark, profilage de code
- Debogeur : gdb, valgrind
- Version : SVN, Darx
- Mais aussi d'autres outils d'aides à la parallèlisation

Les résultats

- Beaucoup de résultats , comment les analyser ? (problématique du « result mass checking »)
- Outils d'aide à l'analyse des résultats (problématique d'apprentissage et de représentation des connaisances)

Conclusions

- Les futurs environnements apporterons dès le départ des contraintes sur la conception/implémentation du soft
- Beaucoup de problématiques concernent le domaine de la recherche en informatique
- Problématiques seulement compréhensibles par des experts ;)
- Nos convictions :
  - Manque de comités transversaux d'experts de standardisation dans plusieurs domaines (services grilles, persistances des données, visualisation, etc.)
  - Besoin de standard « open source », modulaire et inter-opérable
  - « Une fonctionnalité, plusieurs implémention » vs. ré-invention multiple de la roue
  - Quel rôle pour le LAL ?

Questions ?