

Portail Numérique

VO Paris datacenter

Journée Workflow - 21 Décembre 2006

Franck Le Petit

Laboratoire Univers & Théories

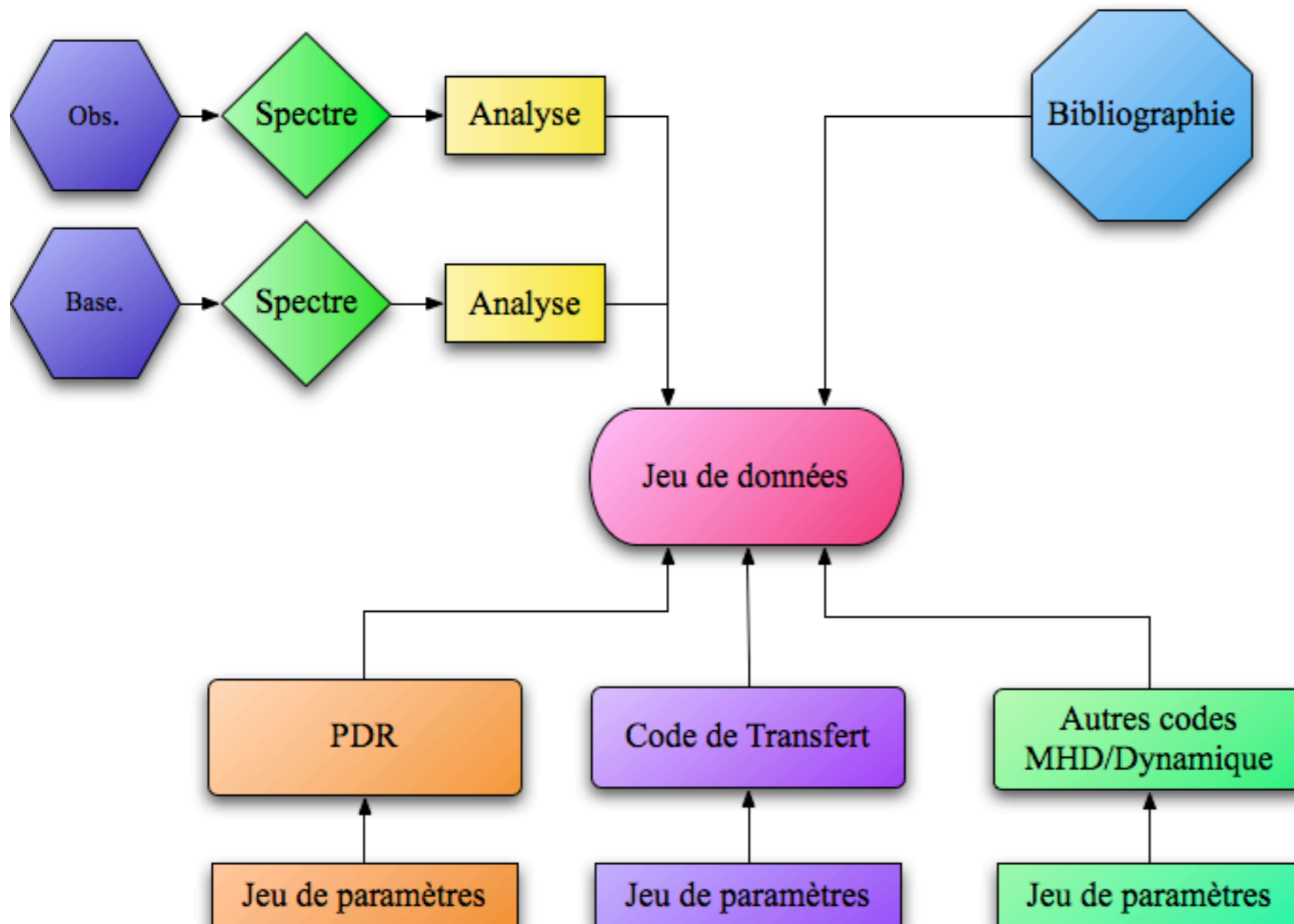
<http://vo.obspm.fr/simulation>



Virtual Observatory Paris Data Center

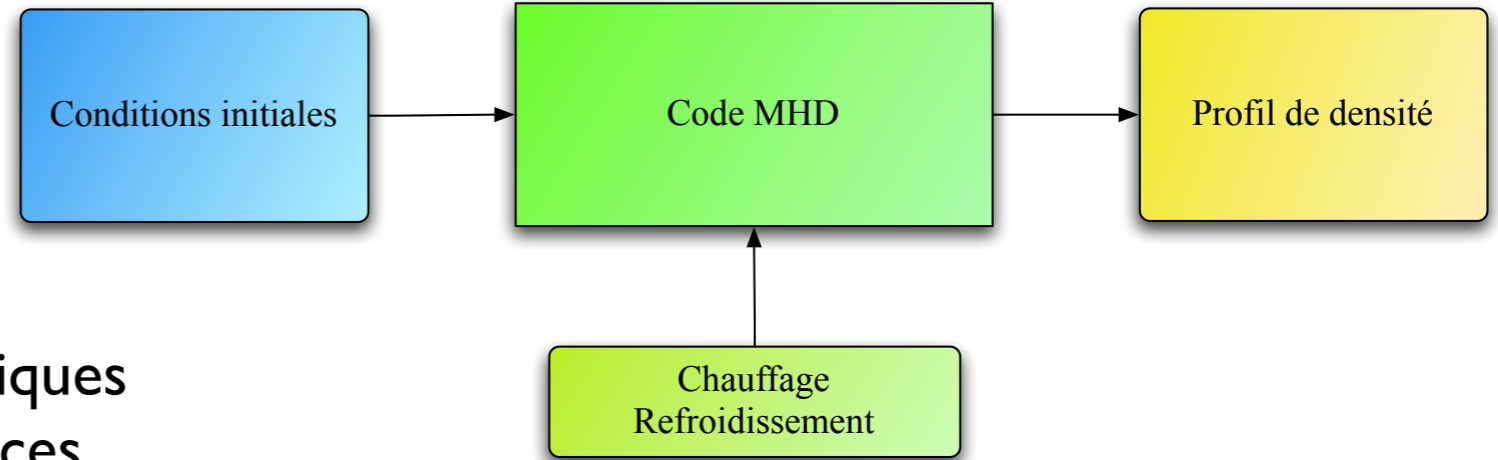
Observations dans des milieux moléculaires

- nuages moléculaires : diffus & sombres
- disques circumplanétaires
- systèmes Lyman alpha dampés



Code de MHD

- résout l'évolution dynamique
 - Profil de densité
- ne résout pas les processus microscopiques
- ne calcule pas les abondances des espèces



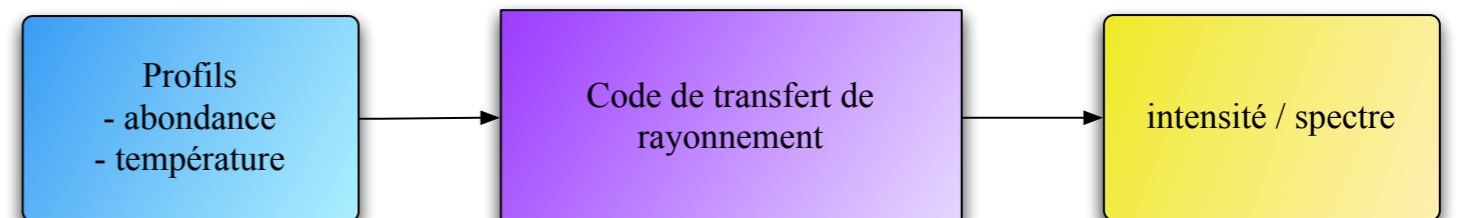
Code PDR

- calcule la structure microscopique
 - Profil d'abondance des espèces
 - Profil de température
- ne résout pas la dynamique : profil de densité ad-hoc
- ne calcule pas l'émission de toutes les espèces



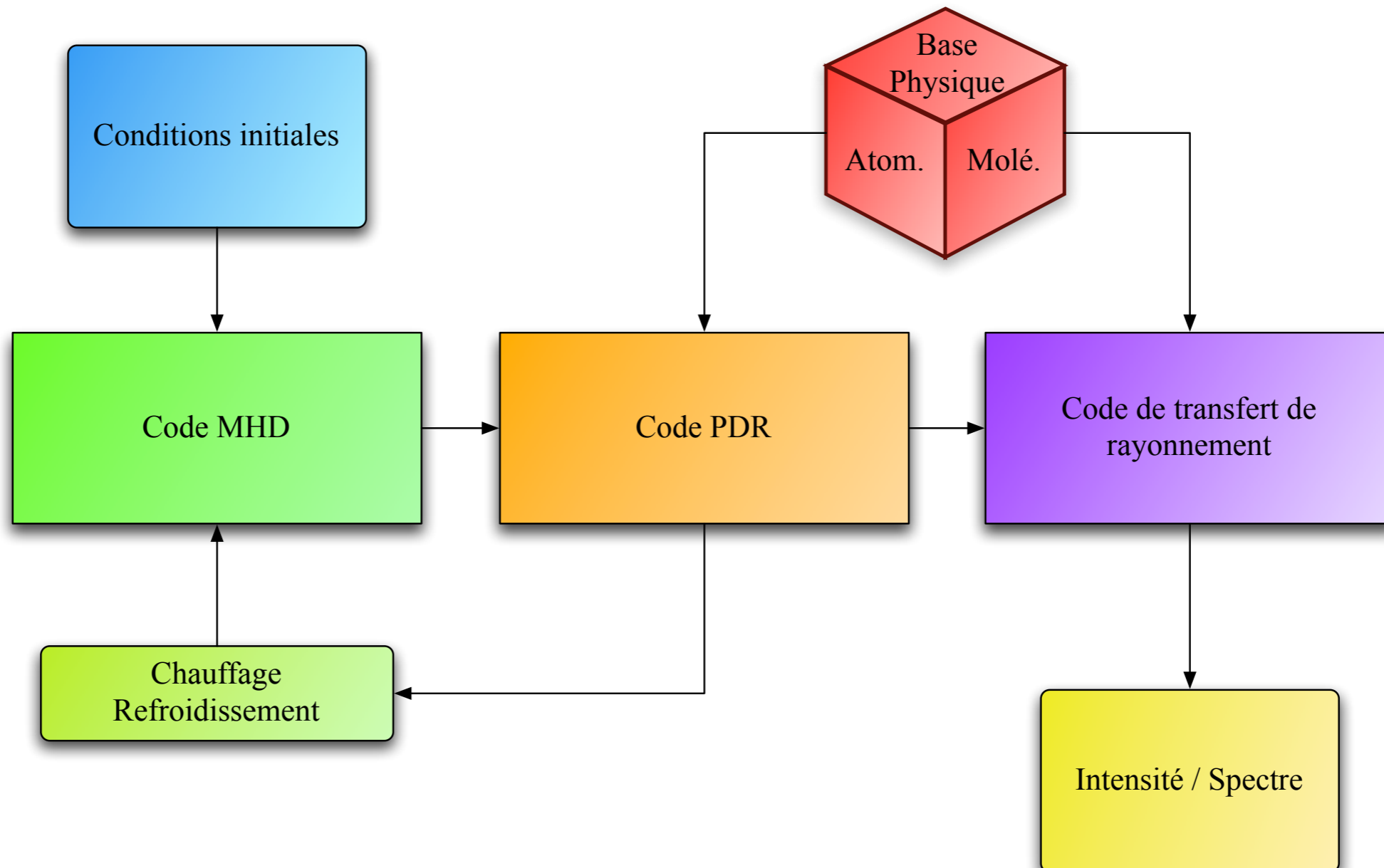
Code de transfert de rayonnement

- résout l'équation de transfert
 - émissivités ou spectres
- ne résout pas la chimie / bilan thermique



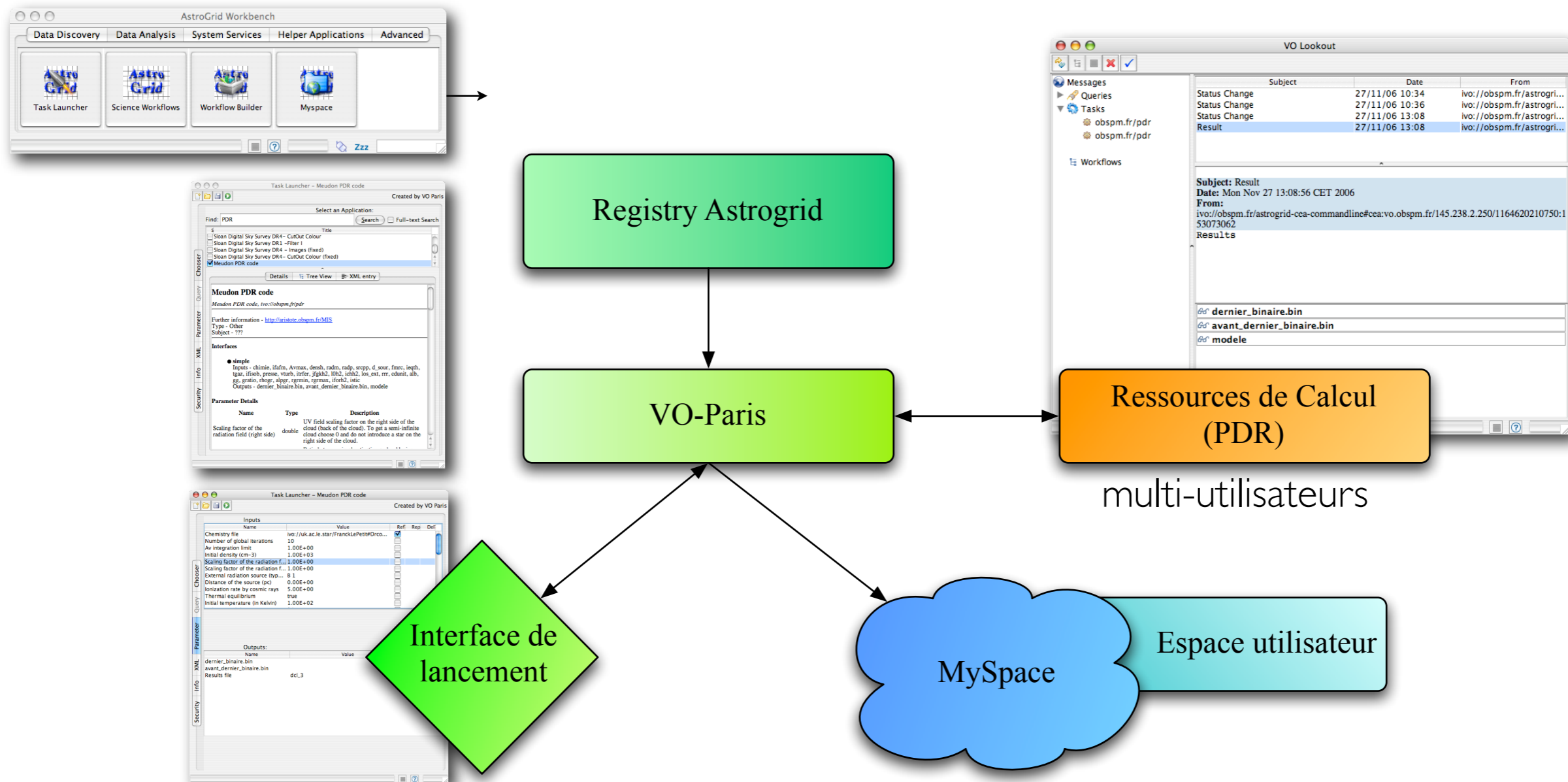
Objectif :

- Utiliser des workflows pour relier les différents éléments
- Faire simplement des opérations aujourd'hui complexes
- Aller plus loin dans l'analyse scientifique



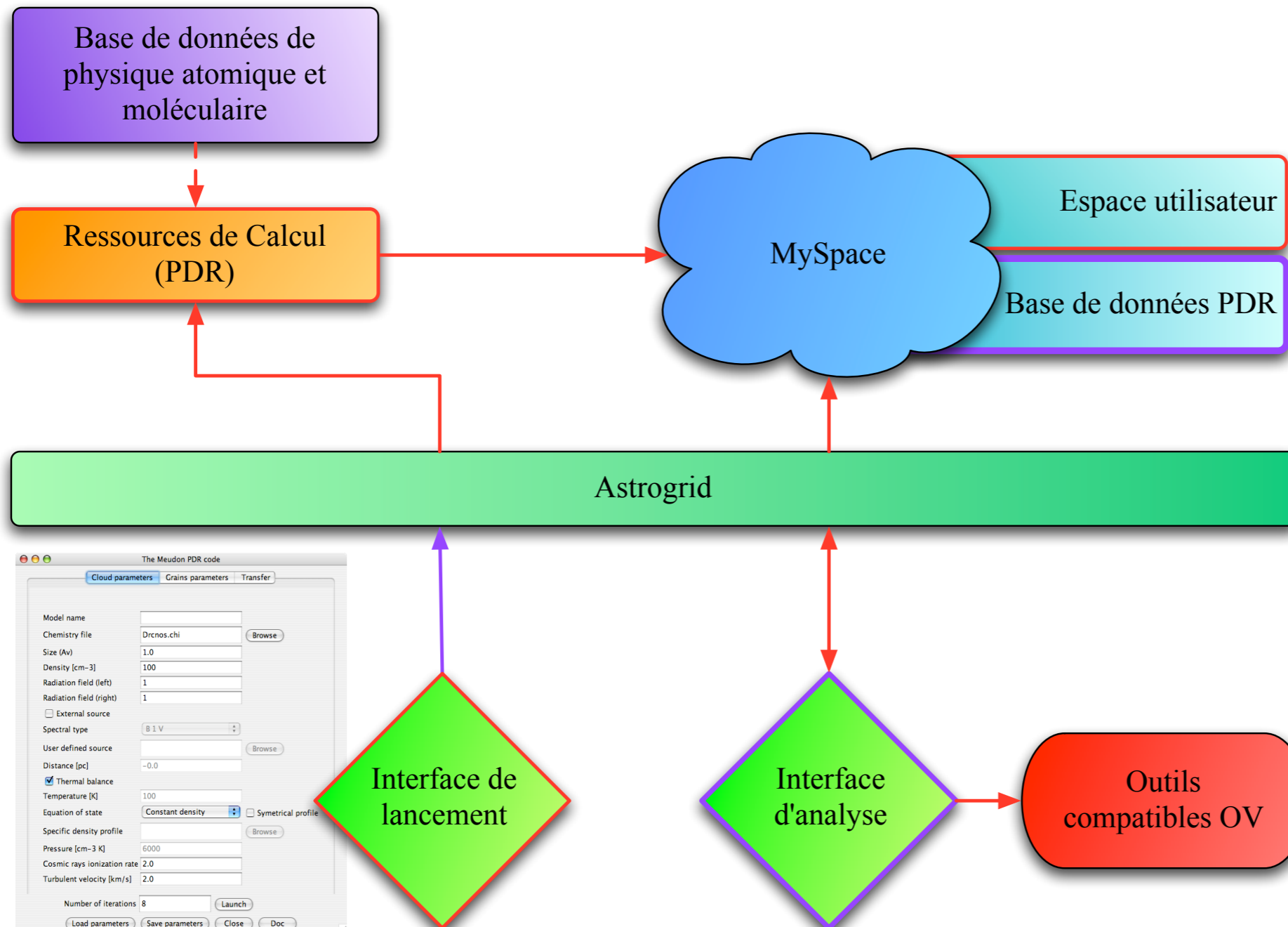
- Première étape :
 - inscrire des codes de simulations dans le VO
 - pouvoir lancer des codes via le VO sur des serveurs distants
 - pouvoir extraire des résultats du binaire créé
- => fonctionnel mais **manuel**

PDR dans Astrogrid



Introduction du code PDR dans les OV

- Lancement via Astrogrid à partir de l'interface de lancement
- Développement de l'interface d'analyse
- Bases de données de résultats



Développement en collaboration avec :

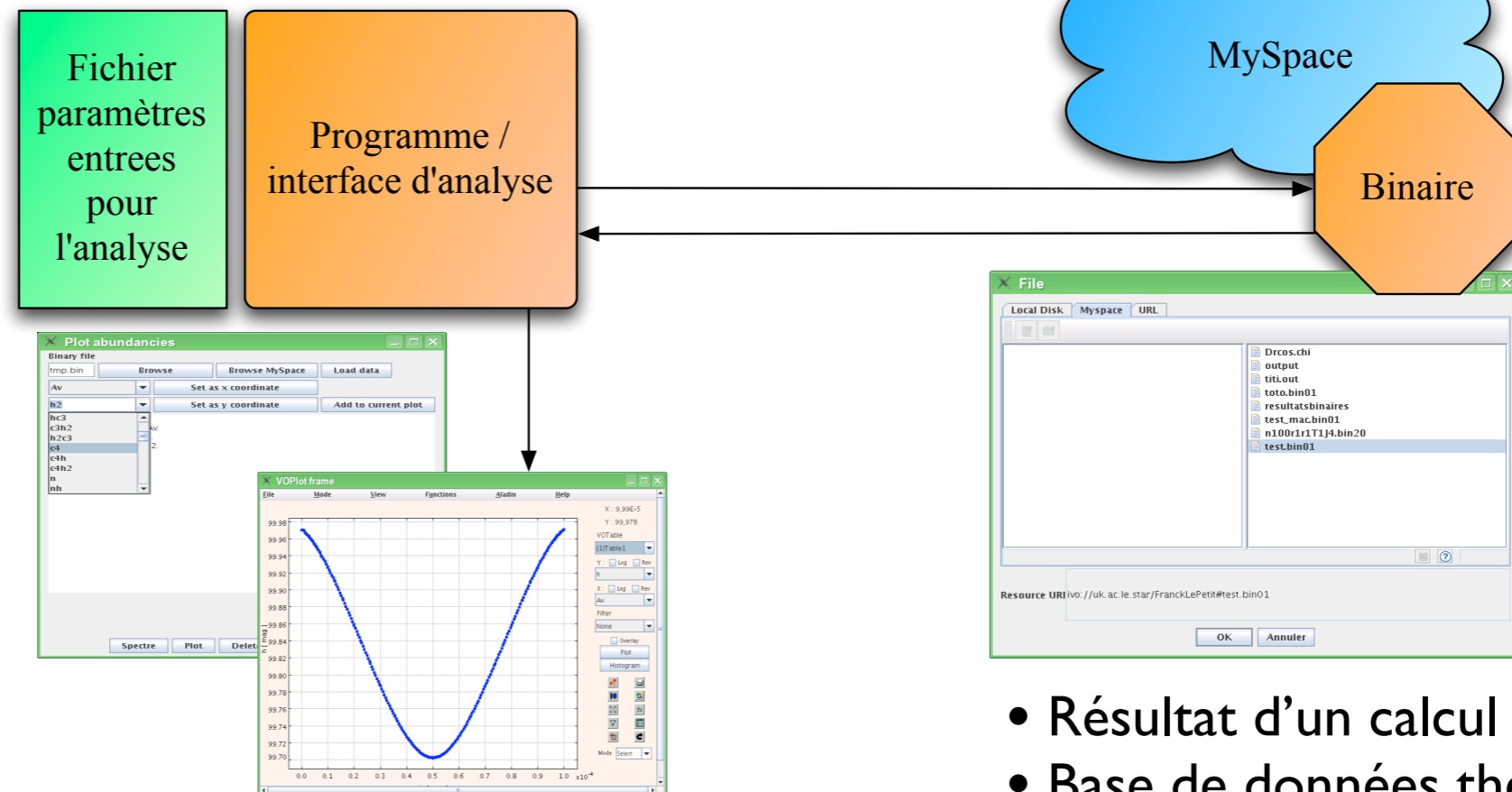
- VO-Paris
- LUTh
- Astrogrid

- **Première étape** : Codes dans le VO
 - inscrire des codes de simulations dans le VO
 - pouvoir lancer des codes via le VO sur des serveurs distants
 - pouvoir extraire des résultats du binaire créé

=> fonctionnel mais **manuel**

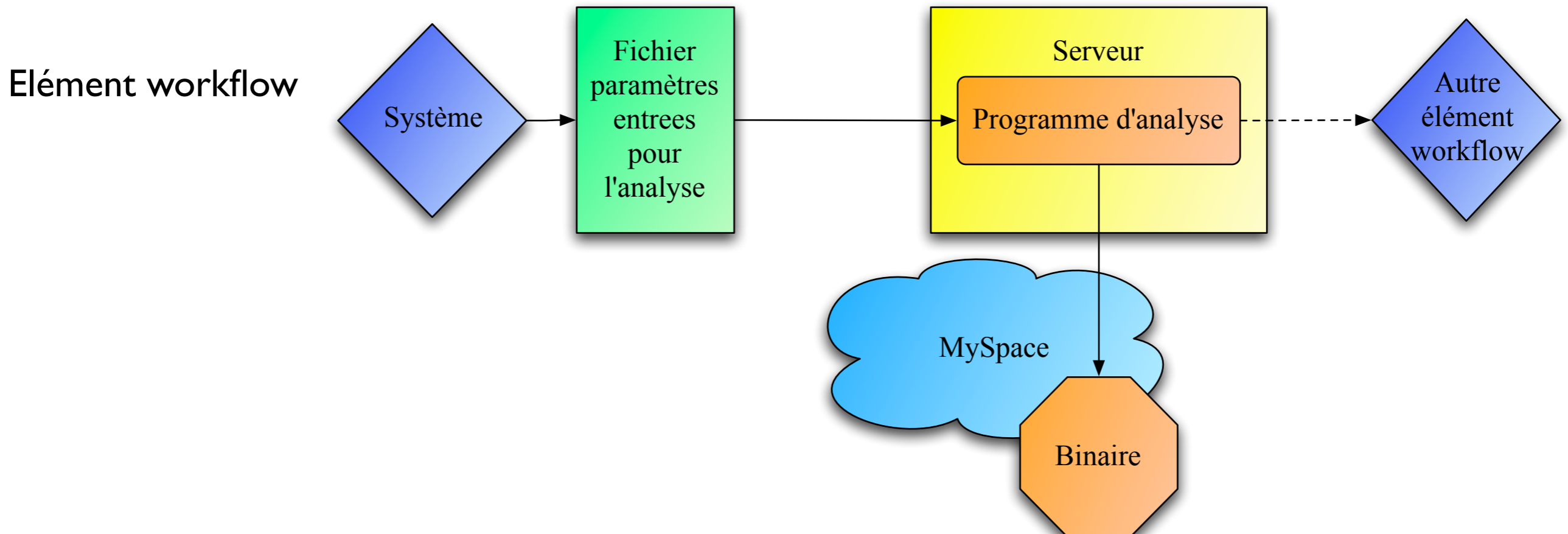
- **Seconde étape** : élément d'un workflow
 - récupération de résultats de façon **automatique**
 - création d'un fichier lu par le programme d'analyse
 - processus "analyse" tourne sur le serveur

Fonctionnement actuel



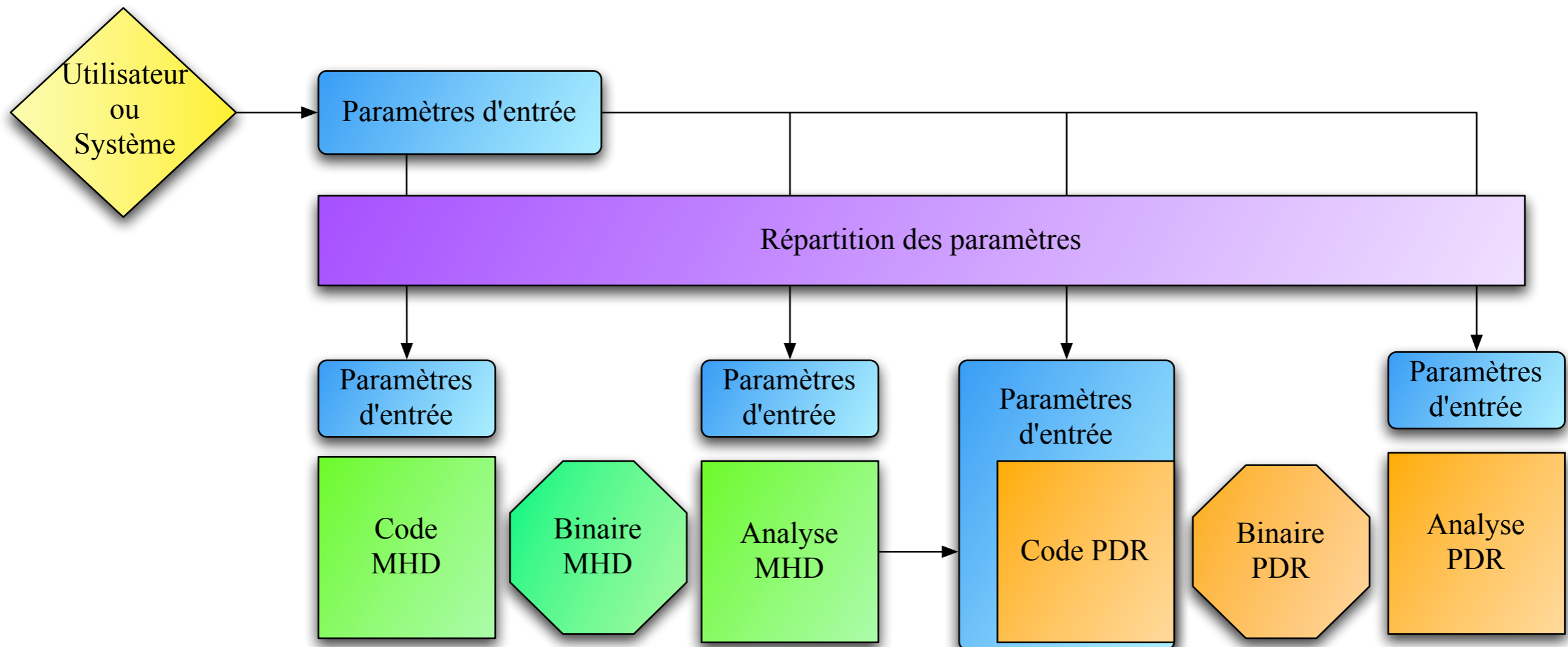
- Résultat d'un calcul
- Base de données théoriques

- **Première étape** : Codes dans le VO
 - inscrire des codes de simulations dans le VO
 - pouvoir lancer des codes via le VO sur des serveurs distants
 - pouvoir extraire des résultats du binaire créé=> fonctionnel mais **manuel**
- **Seconde étape** : élément d'un workflow
 - récupération de résultats de façon **automatique**
 - création d'un fichier lu par le programme d'analyse
 - processus "analyse" tourne sur le serveur



Code de simulation :

- paramètres d'entrée
- code
- fichier de résultats binaire
- code d'analyse
- sortie "human readable"



- **Première étape** : Codes dans le VO
 - inscrire des codes de simulations dans le VO
 - pouvoir lancer des codes via le VO sur des serveurs distants
 - pouvoir extraire des résultats du binaire créé=> fonctionnel mais **manuel**

- **Seconde étape** : élément d'un workflow
 - récupération de résultats de façon **automatique**
 - création d'un fichier lu par le programme d'analyse
 - processus "analyse" tourne sur le serveur

- **Troisième étape** : relier des codes
 - Workflow entre PDR + transfert : test avec Aïda
 - difficulté : langage standard entre les différents membres d'une communauté

Problèmes scientifiques

- Cohérence physique du couplage entre codes
Exemple : Couplage entre un code MHD et un code PDR pour la résolution du bilan thermique
- Définition des quantités physiques peut être différentes entre les codes
Exemple : densité $n = n(H) + 2 n(H_2)$ ou $n = n(H) + n(H_2) + n(He)$?

Problèmes techniques

- UCD / Utypes pour faire communiquer les éléments
Code de transfert au code PDR : "Envoie moi le profil de température et d'abondance de HCO^+ "
- Structuration des données en sorties : datamodel
- Protocoles d'interrogation
 - Travaux en cours pour les codes de simulation sur grille (cosmologie, MHD, ...) : SNAP
 - Pour les codes de transfert le problème est différent mais on doit pouvoir s'inspirer de la "philosophie"
- Problèmes des fichiers binaires
 - Les résultats des codes de simulation sont écrits dans un "gros" fichier binaire
 - En fonction du système sur lequel le binaire a été créé on peut avoir des problèmes pour le relire sur un autre système : *Exemple typique : little endian / big endian*

Solutions :

- Lire le binaire là où il a été créé : pas toujours possible (Transfert des données sur une machine de stockage)
- Solution des images .tiff : code un nombre connu en début de fichier : lourd à mettre en oeuvre

- Couplage entre codes, bases de données théoriques
- Calcul asynchrone
- Changements de types de codes pour étudier effet sur résultats physiques
- Comparaison aux observations

