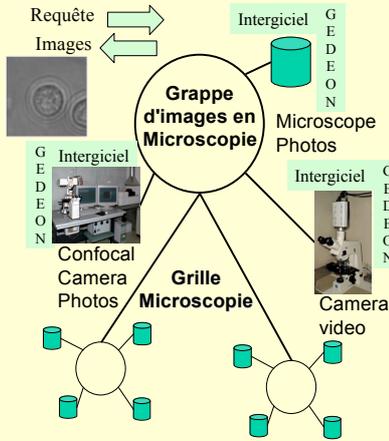
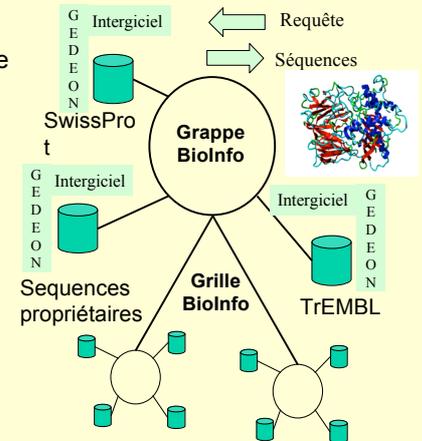


GEDEON est un système **hybride**, combinant les techniques **systèmes de fichiers** hautes performances et **bases de données distribuées**, pour optimiser le déploiement et l'accès à des masses de données sur une grille.



- Gestion distribuée de données à l'échelle d'une grille
- Solution simple pour grilles légères
- Interface de type système de fichier pour la compatibilité avec les applications et les banques de données usuelles du domaine
- Accès aux données enrichis exploitant les méta-données
- Interface pour optimiser les mises à jour régulières des banques de données
- Orientée Performances (File System, caches, pré-chargement)
- Déploiement et configuration flexible et adaptable



Au cœur de GEDEON : Système G

Organisation des méta-données

- **Donnée** : Information brute, utile aux utilisateurs.
Les données (images, séquences) sont stockées sur la grille et référencées par GEDEON. Des méta-données leurs sont associées.
 - **Méta-donnée** : Couple (attribut, valeur)
 - **Enregistrement** : {(méta-donnée+, Id donnée)}
- Un enregistrement est la description d'une donnée dans Gedeon. L'enregistrement est L'UNITE de travail. La donnée n'est pas physiquement présente dans l'enregistrement mais repérée de manière unique par un identifiant.
- Les enregistrements sont regroupés dans des fichiers.
- **Base** : Ensemble d'enregistrements réparti.

- **Noyau minimal** pour la gestion de données sur grilles (rechercher, ajouter, transparence de la distribution).
- Un ensemble de données distribuées sur la grille et indexées par des méta-données.

Interrogation

Une interrogation de type base de données (e.g. SQL) suppose que l'on sait énoncer exactement ce que l'on cherche. Peu adapté pour des recherches approximatives. Gedeon propose :

- Interrogation exacte et incrémentale basée sur les méta-données
- Possibilité de découvrir les méta-données et données

Histologistes

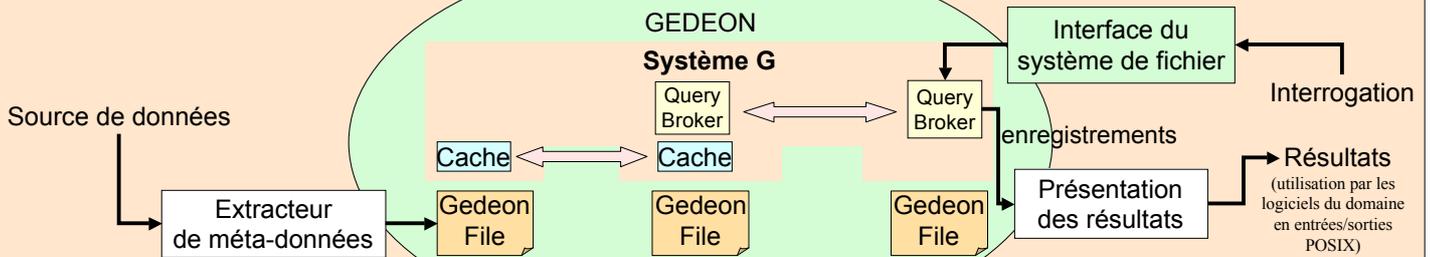
```
cd Microscopy
cd objective=63x,f1.4 / filter=LP650
```

Traitements Batch
sur les images

Biologistes

```
cd SwissProt
cd DT_Update=10-25-2005
cd GN_Name=Jon99cii
cd GN_Synonyms = SER2,SER5
```

Traitement BLAST
des séquences



Cache

- Optimisation des temps de réponse
Gain de temps de communication
Gain de temps de calcul
- Caches de requêtes et caches d'enregistrements
Un Cache de requêtes est un cache sémantique qui identifie les enregistrements qui répondent à une requête (décrites par des méta-données).
- Cache réparties et coopératifs
Répartition de la charge
Utilisation fine des différents types et niveaux de caches.
- Gestion adaptative
Assure une utilisation optimale des caches en fonction du domaine, de l'environnement et de la distribution.

- Hautement performant grâce à l'utilisation de **caches** sur plusieurs niveaux, d'une technologie de transfert efficace des données et d'une gestion optimale de la **distribution**.

Distribution

- Diverses architectures selon le déploiement des serveurs
- Possibilité de séparation des serveurs de méta-données et de données

