

# Présentation de la suite Informatica Power Center

Par DARGES Jean-Sébastien 

Date de publication : 5 novembre 2009

Dernière mise à jour : 4 décembre 2009

Ce tutoriel va vous présenter de façon rapide la suite Informatica PowerCenter le leader du marché des ETL.

**Commentez**

I - Introduction.....	3
II - L'architecture Informatica.....	3
III - Repository Manager.....	3
IV - Repository Server Application Console.....	6
V - Designer.....	6
VI - Workflow Manager.....	8
VII - Workflow Monitor.....	10
VIII - Conclusion.....	11
IX - Remerciements.....	12

## I - Introduction

Informatica PowerCenter est un ETL (Extract Transform Load) qui permet de centraliser des données provenant de sources variées en y appliquant plus ou moins de transformations. Il est effectivement possible d'extraire des données de fichiers plats (délimités, XML, positionnels), ou de bases de données comme Oracle, IBM DB2, Informix, MSSQL, d'y appliquer des transformations, suivant des règles métiers, et de charger le tout dans un entrepôt de données (datawarehouse) ou même dans des fichiers plats.

Ce tutoriel va donc vous présenter les différentes applications faisant partie de la suite Informatica PowerCenter. Une connaissance même minimale des concepts ETL suffiront pour une bonne compréhension de cette présentation.

## II - L'architecture Informatica

L'architecture Informatica PowerCenter se présente de la manière suivante :

- **PowerCenter Repository** : l'installation d'Informatica PowerCenter crée des tables dans une base de données permettant la gestion des référentiels, des utilisateurs, des serveurs... qui seront ainsi accessibles à distance pour les applications et outils PowerCenter.
- **PowerCenter Repository Server** : permet la gestion des connexions aux référentiels depuis les applications clientes et permet ainsi de garantir la cohérence des objets.
- **PowerCenter Client** : ce terme définit tout simplement les applications clientes disponibles à l'installation d'Informatica PowerCenter à savoir :



*Repository Manager : gestion des référentiels, des répertoires propres aux référentiels, des utilisateurs, des droits...*



*Repository Server Application console : gestion des connexions, des verrous sur les objets, des logs, des sauvegardes...*



*Designer : application permettant de développer des mappings afin d'extraire, transformer et charger des données dans une base de données par exemple*



*Workflow Manager : ordonnancement et configuration des mappings*



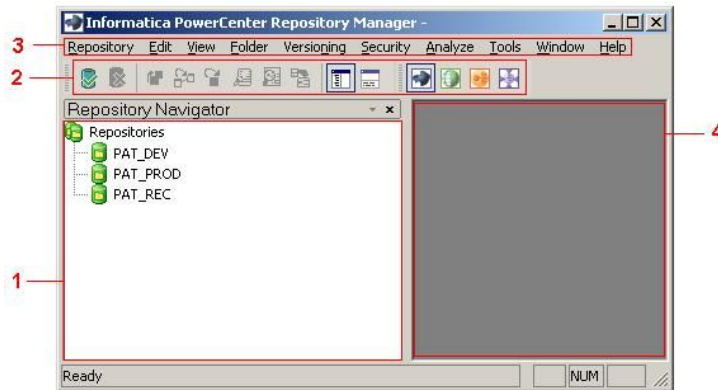
*Workflow Monitor : suivi de l'exécution des workflows*



*Ce tutoriel va présenter les 5 applications clientes précédemment citées.*

## III - Repository Manager

Cette application permet d'accéder aux référentiels afin de gérer les dossiers, les utilisateurs, les droits d'accès, les versions...



- 1** : liste des référentiels de travail disponibles. On peut voir ici que j'ai créé un projet PAT en 3 versions :
- PAT\_DEV : référentiel de développement qui permettra de développer de nouveaux mappings
  - PAT\_REC : référentiel qui sera disponible pour la phase de recette de mon projet
  - PAT\_PROD : référentiel qui sera utilisé en production

Ainsi les développeurs auront accès au référentiel PAT\_DEV pour développer de nouvelles évolutions, de nouveaux mappings, le référentiel PAT\_REC permettra aux développeurs, une fois les développements finis, de mettre à la disposition des personnes opérant sur la recette de faire des tests sur les développements réalisés. Une fois ces tests terminés il sera ainsi possible de modifier les mappings ne correspondant pas à la demande en DEV puis de les réimporter en recette pour être enfin chargés en production (PAT\_PROD), une fois la recette OK.

On aurait pu également ajouter un autre référentiel, par exemple PAT\_PERF, pour tester les développements avec un nombre important de données pour ne pas gêner les autres environnements.

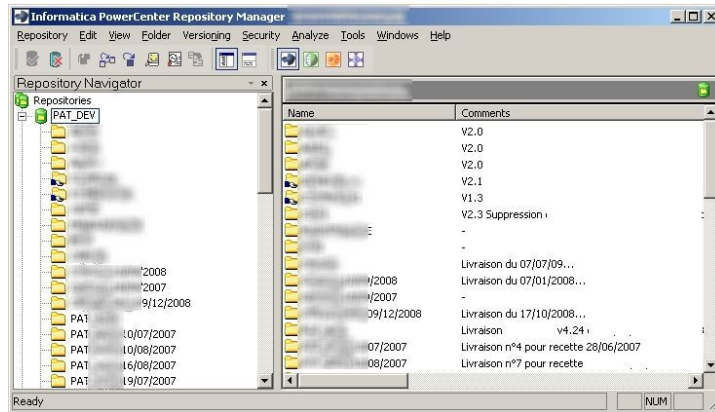
**2** : raccourcis de la barre de menus

**3** : barre de menus permettant entre autres de :

- Se connecter à un référentiel
- Modifier, supprimer éditer un référentiel
- Accéder aux différentes vues de la fenêtre
- Ajouter des dossiers au référentiel
- Gérer les versions des dossiers
- Gérer les privilèges utilisateurs Analyser le contenu des référentiels
- Importer / exporter un référentiel et/ou ses dépendances
- Accéder aux autres applications de la suite PowerCenter

**4** : la 4ème vue permet d'accéder au contenu du référentiel afin d'organiser un projet

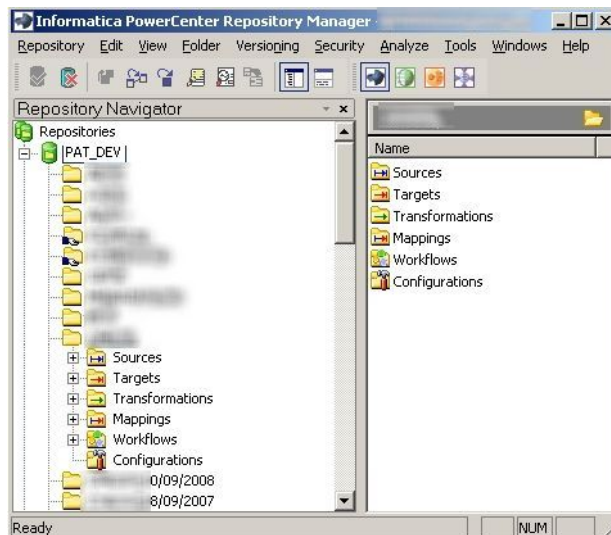
- organisation des dossiers



En se connectant à un référentiel on peut accéder aux dossiers lui appartenant. Ces dossiers constituent des sous-espaces de travail ou des versions de ces sous-espaces. Ce qui est pratique pour revenir sur une version antérieure par exemple.

On peut, bien sûr, affecter des commentaires à ces dossiers pour plus de compréhension lors des livraisons ou copies de dossiers.

- contenu des dossiers

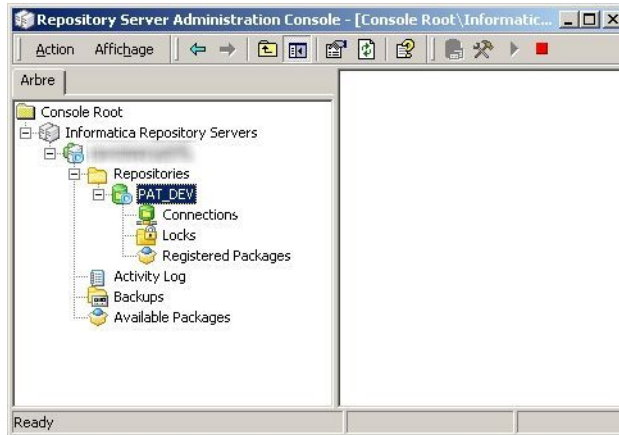


Une fois un dossier développé, on peut accéder à son contenu à savoir :

- les sources de données
- les cibles
- les transformations
- les mappings
- les workflows
- les différentes configurations associées

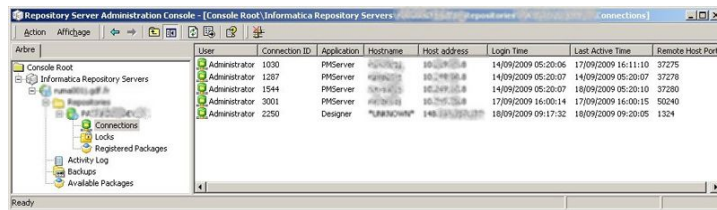
Il est ainsi possible d'avoir une vue détaillée du contenu de chaque dossier et d'accéder plus spécifiquement à chaque objet.

## IV - Repository Server Application Console

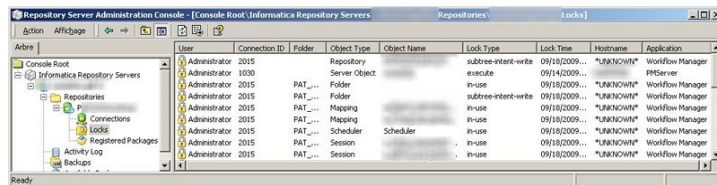


Cette application permet d'afficher chaque serveur où est installée la suite PowerCenter ainsi que les informations relatives à un référentiel telles que :

- les connexions courantes à telle ou telle application de la suite PowerCenter



- les verrous des objets du référentiel (mappings, sources, cibles)



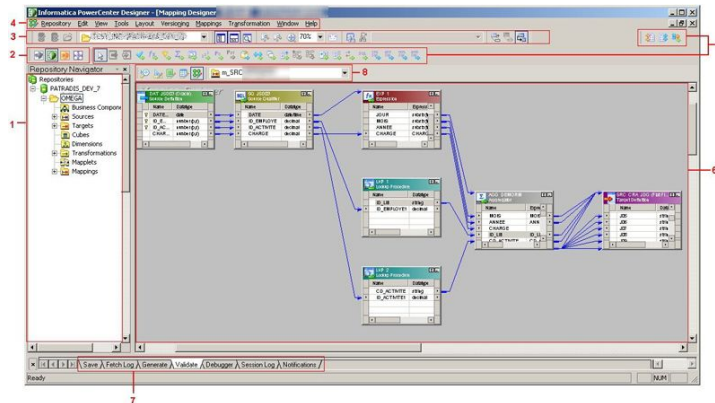
- les logs
- les backups planifiés
- ...

Pour simplifier, cette application permet de savoir qui fait quoi sur tel objet à un moment donné pour pouvoir libérer des ressources si nécessaire.

## V - Designer



Le designer sert à définir les sources de données, les cibles, développer les mappings, les transformations, les mapplets et bien d'autres choses



**1** : comme toutes les applications de la suite PowerCenter, vous avez à gauche l'arborescence des différents référentiels et dossiers que vous avez créés. Chaque dossier contenant les définitions des données sources, cibles, mappings

**2** : raccourcis pour accéder aux applications de la suite PowerCenter

**3** : permet d'accéder aux dossiers du référentiel ouverts ainsi que de montrer / cacher certaines vues.

**4** : barre de menus du Designer

**5** : raccourcis pour accéder aux transformations possibles à savoir :

- Update Strategy Transformation : définit le comportement de sortie vers une base de données (insertion, mise à jour...)
- Expression Transformation : permet entre autres, de faire des calculs, d'affecter de nouvelles valeurs à des champs
- Filter Transformation : permet de filtrer un flux d'entrée
- Aggregator Transformation : permet d'agréger un flux
- Lookup Transformation : permet de faire un lookup sur une source / cible de données
- Sequence Generator : permet de générer une séquence de nombres internes à Informatica. Cela permet de ne plus passer par une séquence Oracle, par exemple, qui ralentirait une insertion en masse.
- Stored Procedure Transformation : fait appel à une procédure externe
- Custom Transformation : transformation personnalisée
- Joiner Transformation : permet de joindre plusieurs sources de données
- Normalizer Transformation : permet de normaliser un flux
- Mapplet : fait appel à un mapplet (sorte de " mapping fonction " pouvant être utilisé dans plusieurs mappings distincts)
- Router Transformation : permet d'écarter un flux d'entrée en groupe de données afin de les diriger vers d'autres transformation
- Sorter Transformation : ordonne les colonnes d'un flux de données par ordre croissant ou décroissant
- Transaction Control Transformation : contrôle une transaction
- Source Qualifier Transformation : spécifie les champs à sélectionner d'après le schéma d'une ou plusieurs tables d'une base de données
- XML Source Qualifier Transformation : spécifie les champs à utiliser parmi un flux de données XML

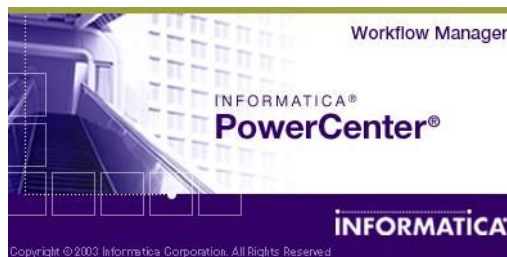
**6** : l'espace de travail où vous pourrez définir vos sources / cibles et développer vos mappings.

**7** : accès aux différents logs et informations relatives au développement

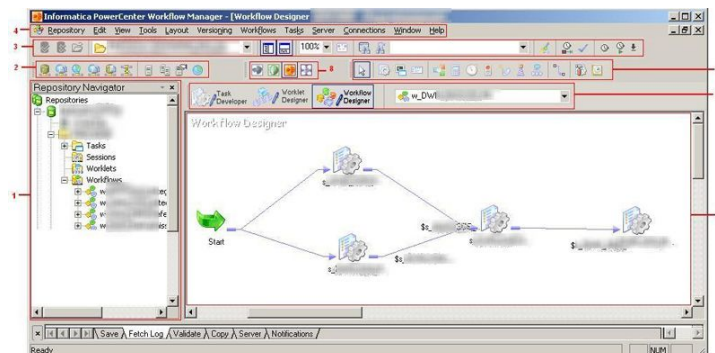
**8** : ces 5 icônes sont très importantes. Ils vous permettront d'accéder aux espaces de travail suivants :

- Source Designer : vous pourrez y définir vos sources de données (fichier délimité, positionnel, XML, Cobol, WSDL et base de données Oracle, Informix, IBM DB2...)
- Warehouse Designer : ici vous pourrez définir vos cibles
- Transformation Developer : permet de définir des transformations personnalisées
- Mapplet designer : permet de développer des sortes de mappings pouvant être utilisés dans plusieurs autres mappings afin de ne pas avoir à redévelopper une suite de composants identiques à plusieurs mappings. En gros une sorte de fonction.
- Mapping Designer : c'est ici que vous développerez vos mappings en y ajoutant les sources définies dans la vue Source Designer, les cibles définies dans Warehouse Designer et d'y apporter des transformations.

## VI - Workflow Manager



Le Workflow Manager permet d'organiser l'exécution des mappings suivant un ordre précis. Vous pourrez ainsi exécuter vos mappings en parallèle.



**1** : arborescence du Workflow manager qui vous permet d'accéder aux tâches, aux sessions, aux worklets et aux workflows

**2** : ces icônes permettent d'ajouter / éditer des connexions à des serveurs locaux ou distants

**3** : affiche les vues mise à disposition et permet également de planifier un workflow et de lancer son exécution en accédant directement au Workflow Monitor.

**4** : barre de menus permettant d'accéder à tous les raccourcis de l'application (import / export, planificateur, sessions, tâches...)

**5** : composants pouvant être déposés sur l'espace de travail du Workflow Manager. On peut ainsi paramétrer une session, une commande externe, un envoi de mail, un chronomètre, un capteur d'erreur...

**6** - Task Developer : permet de créer des tâches à accomplir dans le workflow

- Worklet Designer : permet de développer des worklets. Un worklet est un regroupement de tâches / sessions pouvant être utilisé dans plusieurs workflows.

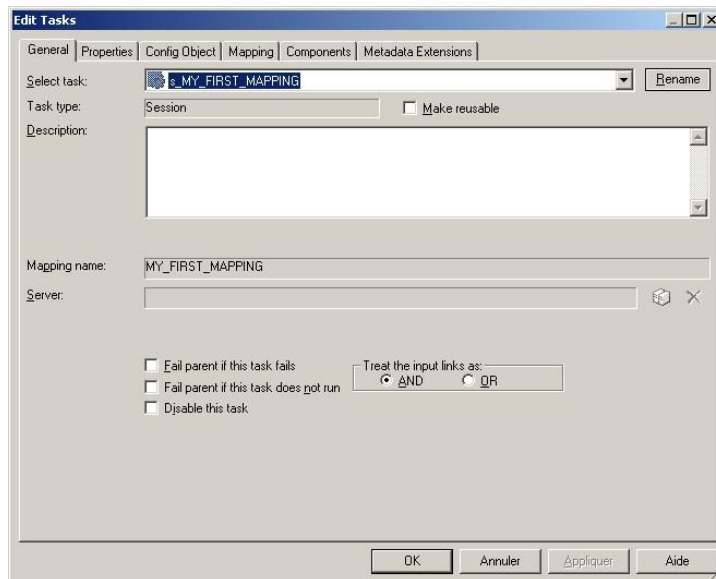


- Workflow Designer : permet de développer les workflows à proprement parler (l'orchestration des tâches, sessions et worklets)

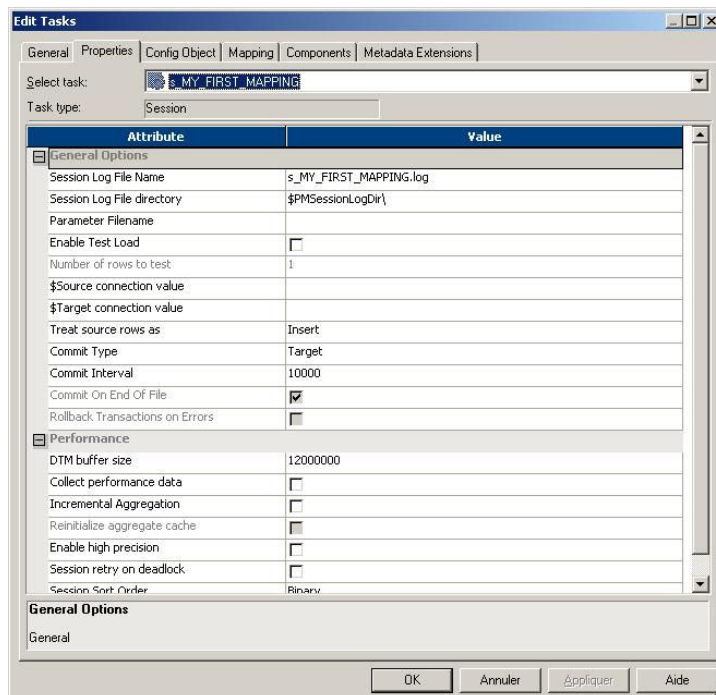
**7** : espace de travail du Workflow Manager. Comme on peut le voir sur l'image, un workflow commence toujours par le composant "Start" ajouté par défaut à la création d'un nouveau workflow. On peut ensuite orchestrer les sessions, worklets et tâches comme on le veut, en parallèle ou non.

**8** : accès aux autres applications de la suite PowerCenter.

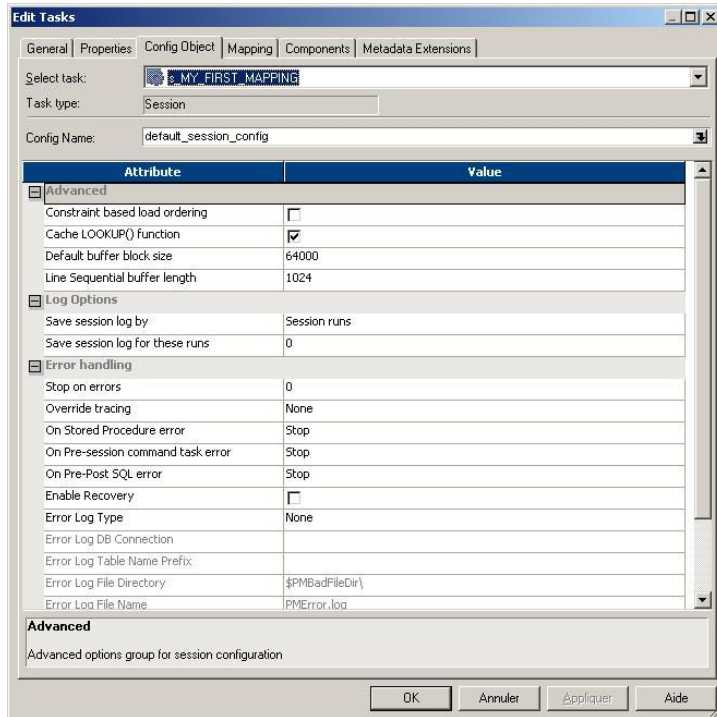
Si vous double-cliquez sur une session, vous pourrez définir les paramètres tels que le nom des fichiers d'entrée / sortie, les connexions aux bases de données, les fichiers de logs, le fichier d'erreur... Voici les vues disponibles :



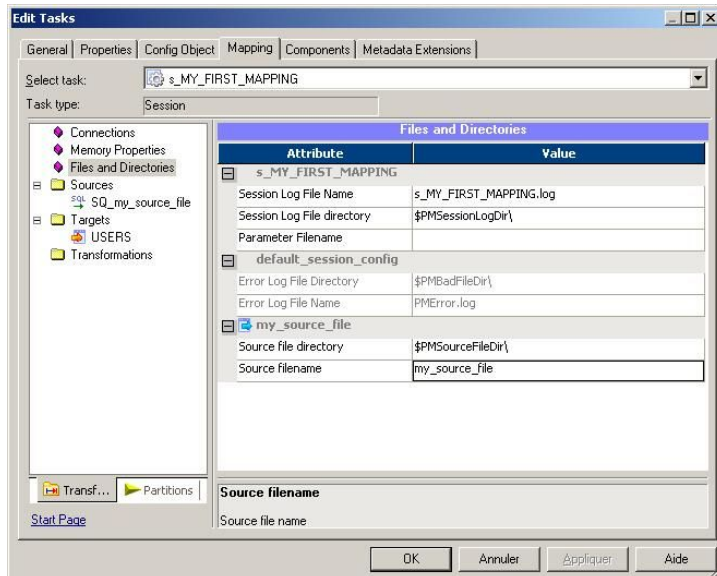
*General*



*Properties*

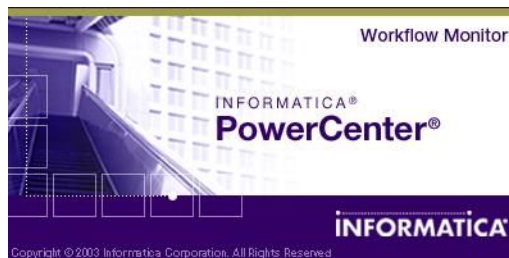


Config Object

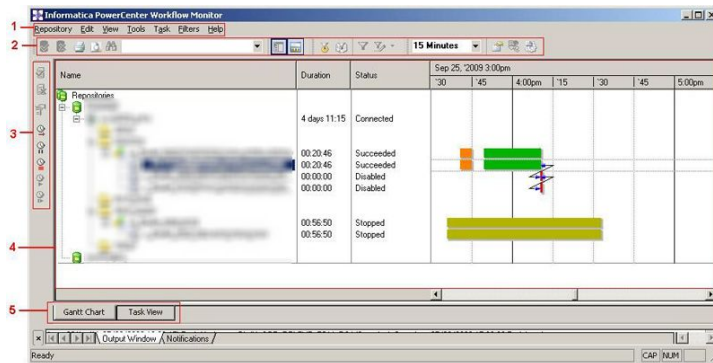


Mapping

## VII - Workflow Monitor

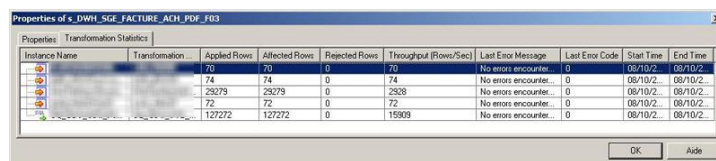


Le Workflow Monitor permet de visualiser l'exécution des workflows. Vous pourrez ainsi lancer un traitement, l'arrêter, accéder à ses propriétés pour voir combien de données ont été traitées. C'est aussi ici que vous pourrez accéder aux logs de traitement.



**1** : barre de menus où vous retrouverez tous les raccourcis

**2** : barre des tâches permettant entre autres de changer de vue, de filtrer l'affichage des workflows, d'afficher les propriétés des workflows... Si vous cliquez sur le bouton de propriété d'une session voici ce que vous pourrez visualiser : tout d'abord dans l'onglet "Propriétés" vous pourrez voir les propriétés de la session à savoir son nom, son type, le serveur d'exécution, le début du traitement, la fin du traitement, son statut, le message d'erreur (s'il y a lieu)... Dans l'onglet "Transformation statistics" c'est ici que vous verrez combien de données ont été traitées, combien on été rejetées, combien sont allées dans telle ou telle cible...



**3** : barre d'exécution des workflows permettant de démarrer, d'arrêter un traitement

**4** : affichage en temps réel de l'exécution des workflows et de leur état. Sur l'image ci-dessus vous pouvez voir comment sont lancés les traitements, s'ils sont en cours d'exécution (couleur verte avec les secondes défilantes), s'ils sont en erreur (couleur orange), s'ils ont été désactivés (couleur rouge), s'ils ont été arrêtés (couleur jaune) ou s'ils ont été abordés (couleur rose foncée). Pour accéder aux logs de traitement, il suffit de faire un clic droit sur la session puis de choisir "Get Session log". Les logs du serveur seront alors automatiquement chargés dans un éditeur de texte.

**5** : l'exécution des workflows peut aussi se visualiser dans l'onglet task view qui affiche uniquement le nom du traitement et son état. Privilégiez la vue "Gantt Chart" pour avoir une vue d'ensemble et plus détaillée sur l'exécution des workflows.

## VIII - Conclusion

Ce tutoriel vous a rapidement présenté la suite Informatica PowerCenter et plus particulièrement les applications clientes.

Vous pourrez retrouver, dans un prochain tutoriel en cours de rédaction, comment créer votre premier mapping sous Informatica. Je rentrerai alors plus dans les détails du Designer.

## IX - Remerciements

Je remercie **jacques\_jean** et **Antoun** pour leur relecture et **Bérénice MAUREL** et **Fleur-Anne Blain** pour leurs conseils et remarques.