

Architecture SAS

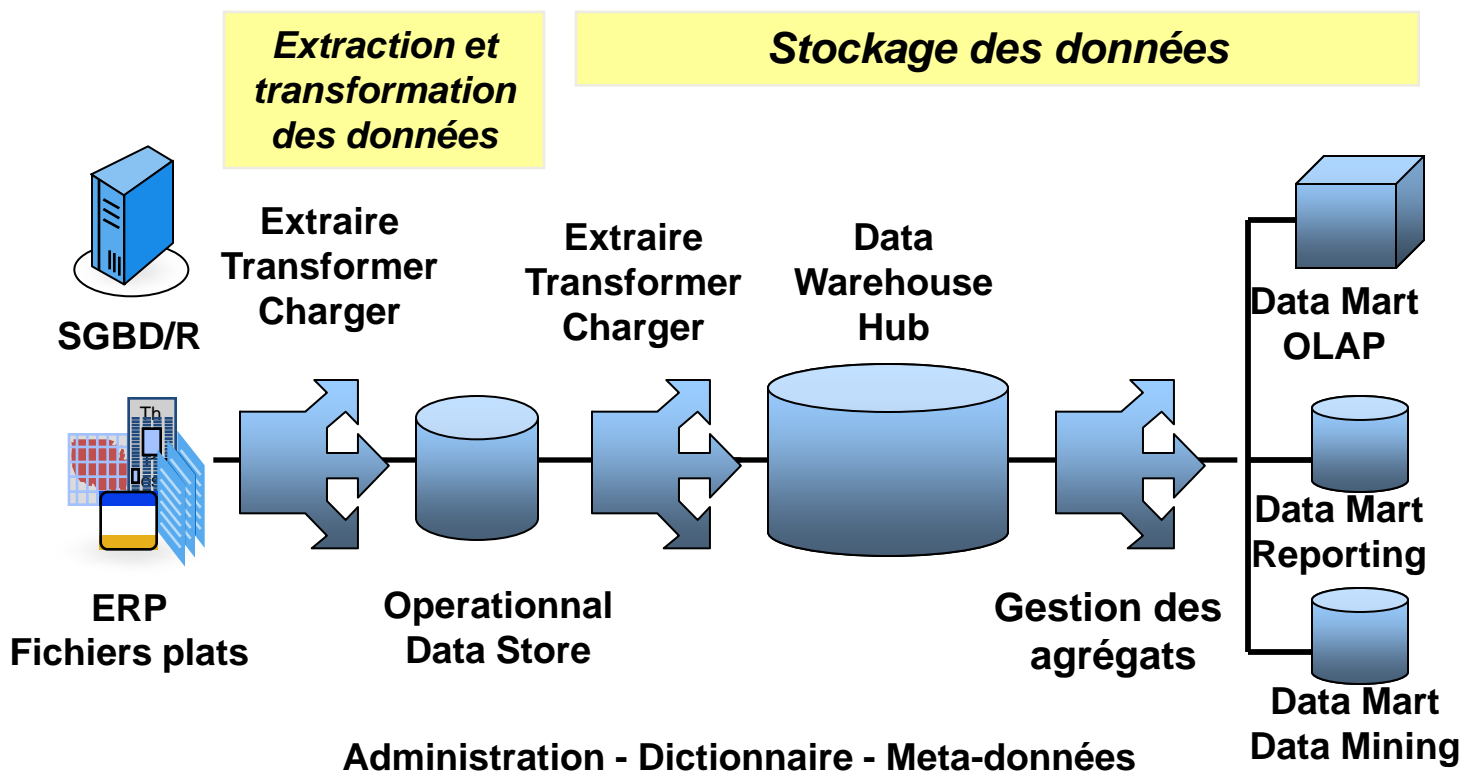
Benjamin TREGUIER

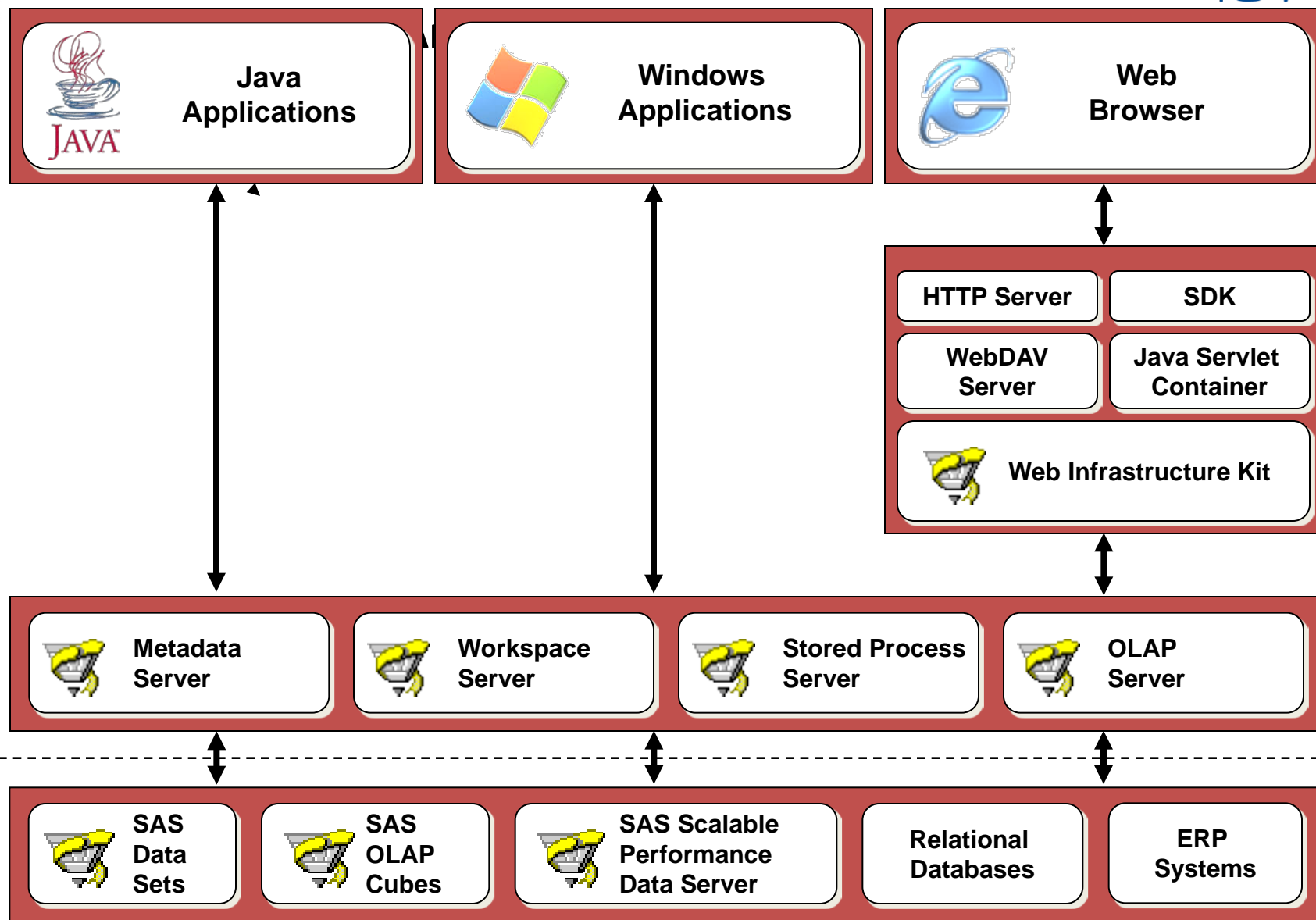
Durée : 24h

Plan de cours

- Introduction à la BI et à SAS
- Sas Management Console et Sas Data Integration
- SAS Enterprise Guide, Proc stockée et Add'in For Microsoft Excel
- Sas Information Map Studio, Sas Web Report Studio

Modélisation : ODS, DWH, Datamarts





Informatique décisionnelle

- Ces données sont pérennes dans le temps, intégrées, uniformisées, organisées de manière structurée...et permettent de piloter.
- Objectif des services informatique et décisionnels : mettre à disposition des utilisateurs métier, les données de l'entreprise utiles, à des fins de prises de décision.

SAS Management Console 9.2

- Permettre aux utilisateurs, de manière libérée (pas de contrainte technique) l'accès aux données décisionnelles
- Différentes applications légères permettant d'accéder aux données selon différentes formes (Cube, Map, etc.)
- Pour du reporting, de l'analyse stat, de la consultation, etc.
- Pourquoi mettre en place des entrepôt de données?

SAS Management Console 9.2

- Application JAVA permettant l'administration centralisée des serveurs SAS et la gestion des MD
- Sécurisation, Contrôle et Documentation de l'entrepôt
- Notions de Méta données
- Quelles sont les MD existantes?

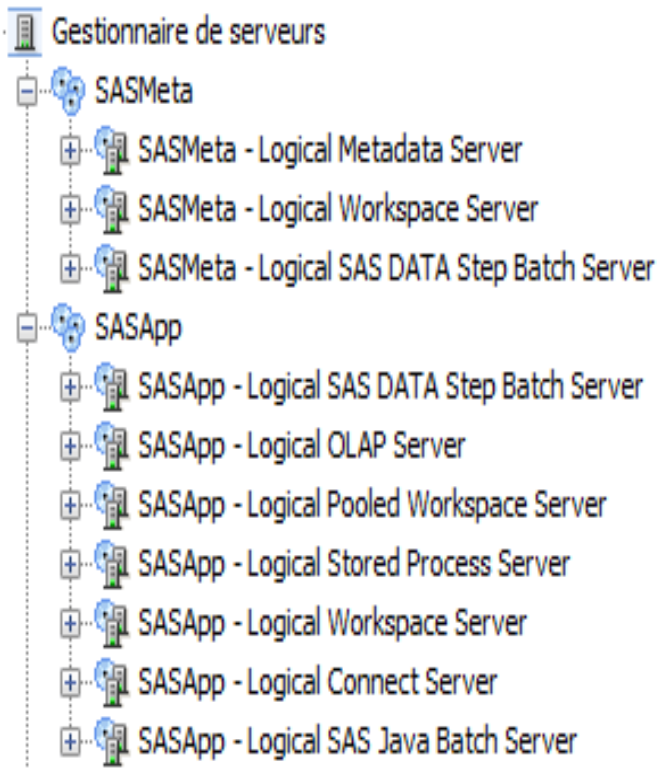
Définition Méta données

- « Données sur les données » = description
- Information partagée et centralisée
- Source, Librairie, Format, Type, etc.
- Utilisateurs de l'information
- Comment sont stockées les Méta Données?

Serveurs SAS et Connexion

- Administrateur se connecte au serveur de MD et gère l'administration
- On associe à cette connexion un **référentiel de MD**
 - **FOUNDATION**
- De ce référentiel on gère les **accès** (utilisateurs), les **serveurs d'application**, les **librairies** rattachées etc.
- D'autres référentiels pourront dépendre de Fondation

Serveurs SAS (exemple)



Référentiels et droits d'accès

- Pour chaque référentiel = un référentiel d'accès
- **ACT** (Acces Control Template) pour déclarer les utilisateurs BI
- Accès aux ressources gérées par le référentiel de MD
- Obtention de permission selon :
 - Groupe
 - Utilisateurs

Rôle de la SMC

- Se balader dans les Méta données
- Définir des serveurs
- Gérer les Méta données
- Affecter des librairies aux serveurs
- Gérer les accès aux ressources
- Définir et regrouper les utilisateurs et leurs droits
- Assurer la maintenance du produit SAS

ETL

- L'etl est l'abréviation de l'anglais Extract
Transfert/Transform/Loading pour Extraction, Transfert et
Transformation des données et chargement (Load en anglais)
dans des base de données (Entrepôt).
- Les ETL sont des outils récents. Le principal concurrent d'un
ETL... est le code développé manuellement.

ETL : Les avantages

- **Interface graphique** : améliore la maintenance, et la lisibilité d'un chargement.
- **Industrialisation** : Un ETL doit pouvoir être paramétré pour déclencher de manière récurrente un même chargement de données.
- **Analyse d'impact** : Capacité de mesurer quelles sont les impacts d'un changement de données (type, format, table d'origine) dans l'alimentation de l'entrepôt.

ETL SAS

- SAS a depuis longtemps été utilisés par des sociétés afin d'alimentés des entrepôts ou datamart.
- Depuis 2000, une première version d'un outil ETL appelé SAS/Warehouse Administrator a fourni aux sociétés les capacités de chargement au sein d'un entrepôt.

ETL SAS

- La version SAS9, avec son architecture centralisée et sécurisée a permis à SAS de mutualiser l'ensemble des métadonnées générées. Les développements JAVA récents ont engendrés une nouvelle suite d'outils dont **SAS Data Integration**.
- SAS Data Integration permet de gérer l'**alimentation** des données d'un entrepôt. L'addin SAS OLAP Cube Studio, option de SAS Data Integration, offre à l'ETL la capacité de construire et rafraichir des structures **multidimensionnelles**.

ETL SAS : Utilisation & démarche

- Se connecter via un **profil** de métadonnées.
- Définir les sources
- Définir les cibles
- Définir le **flux de processus/alimentation** (job) reliant les cibles aux sources.
 - Mode de chargement
 - Règles de gestion
 - Mapping
- Travaux stockés au sein des Méta Données

Cube Multidimensionnel

- Structure multidimensionnelle
- Se balader dans les données (niveau très agrégé → moins agrégé)
- Notion de **hiérarchie** et de **dimension**
 - Les dimensions composent la hiérarchie
 - On se balade dans la **hiérarchie** à travers les niveaux appelés **dimension**
- Sur l'ensemble des dimensions, on pré-calcule à tous les croisements la valeur des indicateurs (NWAY)

Cube Multidimensionnel OLAP

- **On Line Analycis Processing**
- Indépendance des dimensions :
 - Les dimensions se doivent de ne pas avoir d'information commune.
- Gestion des matrices creuses
 - Croisement de dimension vide : Planches de surf en Alsace... Le CA est nul
- Accès multi-utilisateurs
- **SAS Olap Cube Studio : Client SAS de création de cube**

SAS Enterprise Guide 4.3

- **Client SAS d'analyse et de reporting**
- Interface ergonomique
- Accès aux données du serveur d'application
- Permet d'analyser, de requêter les données
- Création de **Procédures stockées** réutilisables
- Accès aux **cubes** multidimensionnels

SAS Enterprise Guide 4.3

- Notion de **projet** et de flux de **processus**
- Un projet SEG contient au minimum 1 flux de processus
 - Création de code : « mode programmation »
 - « Mode assisté »
 - Utilisation de tâche prédéfinie (Stat, graph, tableau de bord, etc.)
 - Création de requête (nouvelles colonnes, filtre, jointure, etc.)
 - Imbrication du mode programmation et assisté
 - Import / Export tableau, graphiques, tables
 - Industrialisation des processus

Procédure Stockée

- Programme SAS enregistré sur le serveur SAS
- Administrée par le serveur de Méta Données
- Serveurs logiques utilisés pour l'exécution :
 - Workspace Server
 - Stored Process Server
- Créée par SEG et exécutable par d'autres applications :
 - SAS Web Report Studio
 - SAS Stored Process Web Application
 - SAS Add'in For microsoft Excel

Procédure Stockée : utilisation

- Utilisation des STP à des fins divers :
 - Sortie ODS de SAS
 - Sources de données SAS
 - Fichiers Externes
 - Publication, etc.
- Peut également être un programme SAS complexe pour des non programmeurs : accès à travers Add'in Excel

Add'in Excel

- « **Permet d'utiliser toute la puissance de SAS à travers Excel** »
- Accès aux données du serveur SAS
- Utilisation de tri, filtre et des tâches de SEG
- Visualisation des données du serveur dans vos feuilles Excel
- Exécution de **Procédures Stockées** et affichage des résultats
- Accès au **cube multidimensionnel** précédemment créé
- **Pour quel type d'utilisateur?**

SAS Information Map Studio

- Création de « MAP » à comparer avec des univers BO
- « Vue métier » transverse des données de l'entreprise
- Principe de création (à partir de tables ou cube) :
 - Création de jointure entre les tables (si tables utilisées)
 - Définition de mesure et de catégorie
 - Création de filtre prédéfini

SAS Information Map Studio

- Principale fonction : Transformer les MD créés en Vue Métier
- Organisation des variables dans des dossiers « Métiers »
- Utilisation des Maps dans Web Report Studio par des utilisateurs « Métier »

SAS Web Report Studio

- Concurrent de BO pour les utilisateurs Métier
- Client Web d'accès et d'exécution
- Création de rapport dans le processus d'aide à la décision
- A partir de :
 - Information Map
 - Procédures Stockées
- Mise en page réalisable en vue de diffusion des rapports créés
- Ne peut être utilisé sans une architecture BI complète