

OLAP

Urbanisation des SI-NFE107 Fiche de lecture

Karim SEKRI

20/01/2009

OLAP

1

PLAN

- Introduction
- OLAP
- Les différentes technologies OLAP
- Plate formes et Outils

20/01/2009

OLAP

2

Informatique décisionnelle

BI, Business Intelligence

Système interprétant des données complexes permettant aux dirigeants d'entreprise de prendre des décisions en connaissance de cause.

Aide à la
décision

Connaître les **clients**, analyser
les ventes et les marchés

Mesurer la **performance**

BD clients, historique des
achats ; segmentation ; CRM

Tableaux de bord, reporting...

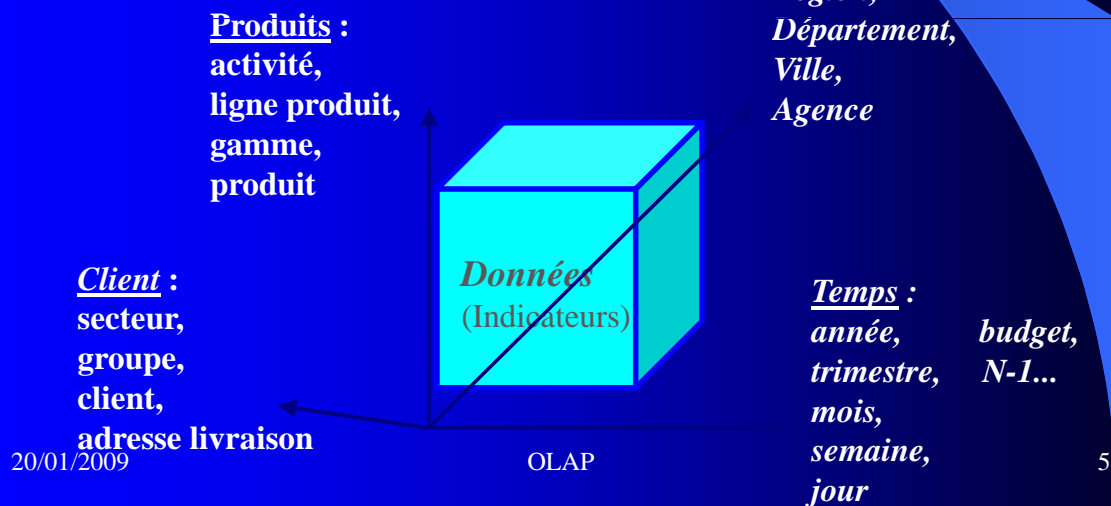
Entrepôts de données

Datawarehouse

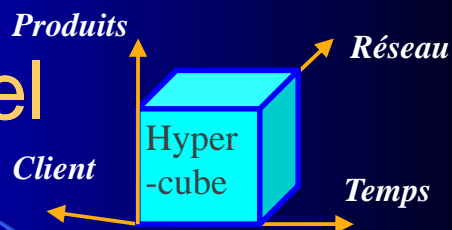
- Le concept d'entrepôt de données a été formalisé pour la première fois en 1990 par Bill Inmon :
- **"collection de données orientées pour un sujet, intégrées, non volatiles et historisées, organisées pour le support du processus d'aide à la décision"**.
- Est créé pour les besoins **décisionnels**
- Est spécialisé par type d'analyse
 - Marketing : BD client, historique des achats
 - Gestion : Suivi coûts / performance des activités
- Est alimenté par les systèmes opérationnels
- A une modélisation **dimensionnelle**...
- ... qui facilite l'analyse selon des **axes** prédéfinis

Entrepôts de données

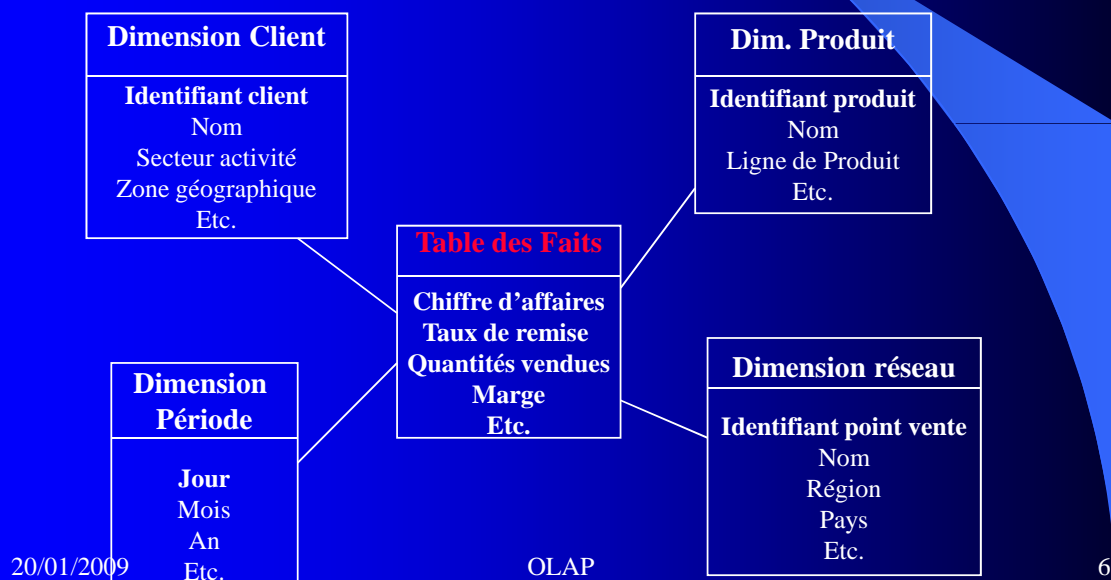
- Modèle dimensionnel : hypercube
 - **Axes** d'analyse (**dimensions**)
 - **Indicateurs** (table des faits)



Modèle dimensionnel



- Schéma BD en « étoile » ou en « flocon »



Entrepôt métier

Datamart

ETL

Outil d'extraction

Datamining

Exploration des données

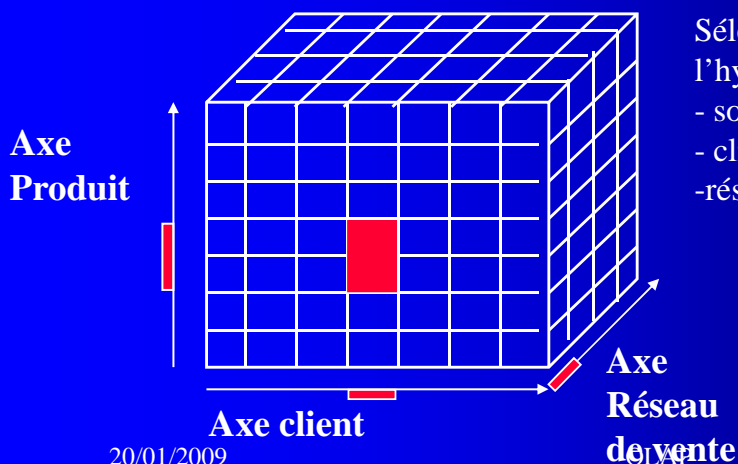
20/01/2009

OLAP

7

Les outils de restitution

- OLAP (On Line Analytical Processing)
- Outils d'interrogation des hypercubes



Sélection d'une **portion** de l'hypercube :

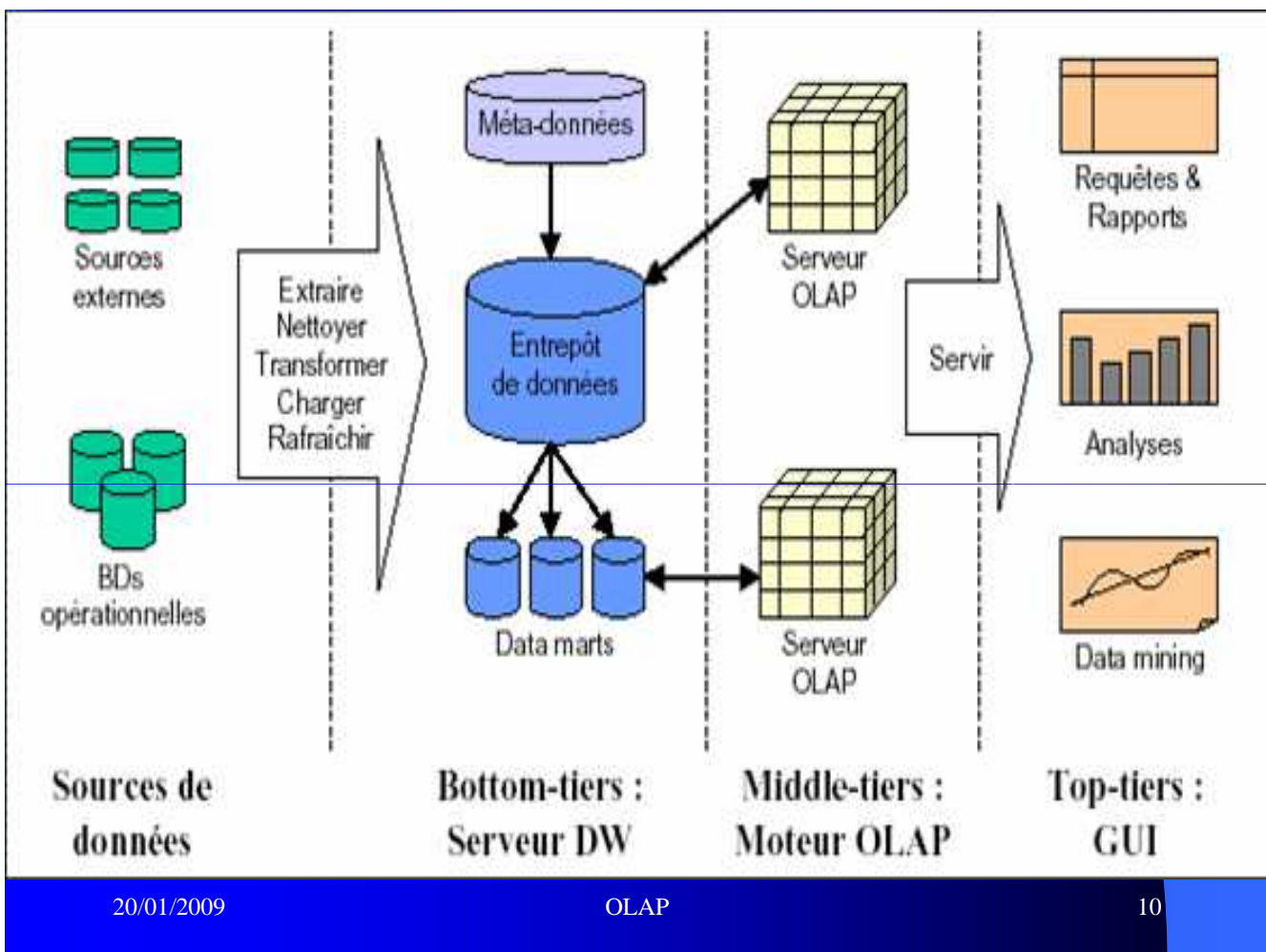
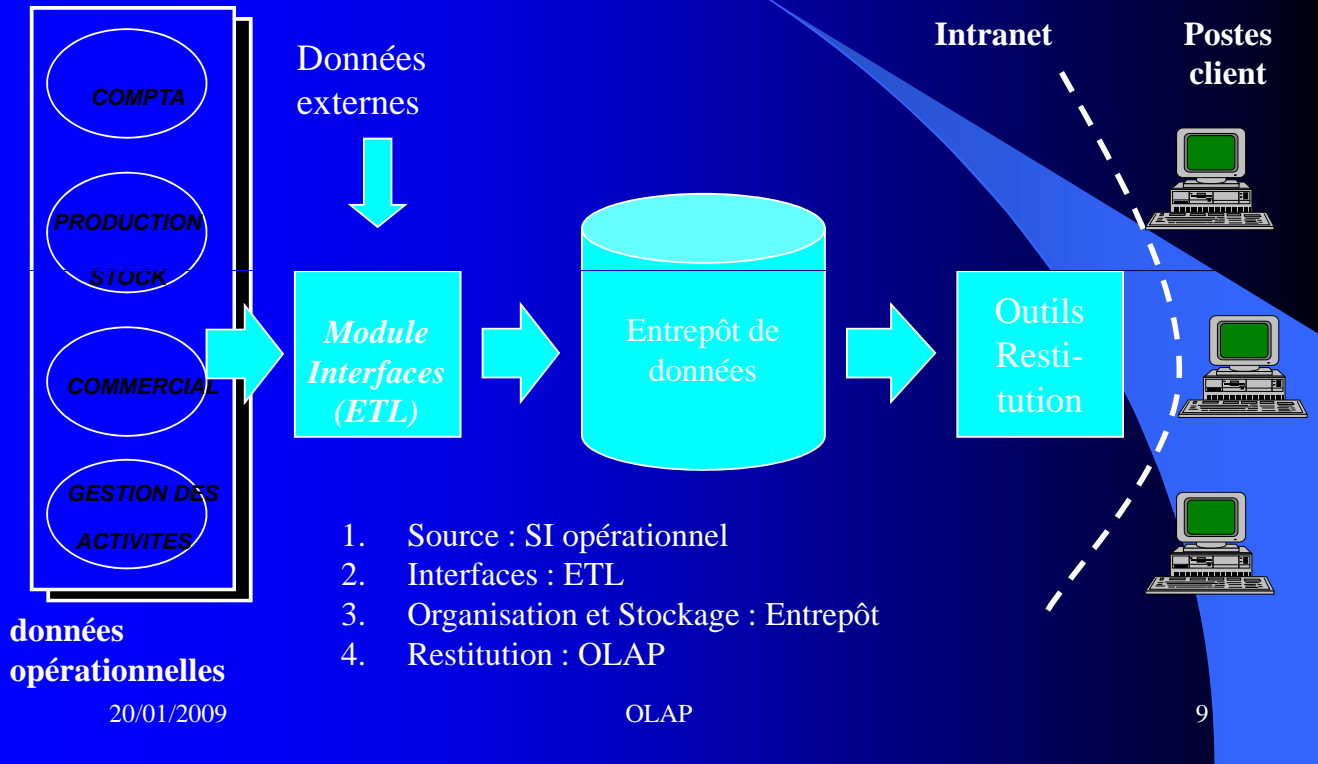
- sous-ensemble de produits,
- client spécifique,
- réseau particulier de distribution

20/01/2009

OLAP

8

Architecture du SI décisionnel



OLAP

E.F. Codd définit 12 règles de base permettant de qualifier le concept global nommé OLAP :

- **Transparence**
- **Accessibilité**
- **Manipulation des données**
- **Souplesse d'affichage**
- **Multidimensionalité**
- **Dimensionnalité générique**
- **Client/serveur**
- **Multi-utilisateur**
- **Accès stable**
- **Gestion des matrices creuses**
- **Croisement des dimensions**
- **Nombre illimité de dimension et de niveaux d'agrégation**

OLAP

- **Approche multidimensionnelle :**
 - Basée sur des thèmes d'analyse (dimensions)
 - Plus intuitive
- **Plusieurs niveaux d'agrégation :**
 - Les données peuvent être groupées à différents niveaux de **granularité** (les regroupements sont pré-calculés, par exemple, le total des ventes pour le mois dernier calculé à partir de la somme de toutes les ventes du mois).
 - **Granularité** : niveau de détail des données emmagasinées dans une base de données.

Vocabulaire OLAP

Dimension :

- Une dimension peut être définie comme un thème, ou un axe (attributs), selon lequel les données seront analysées (en fonction de ...)
 - Ex. Temps, Découpage administratif, Produits
- Une dimension contient des membres organisés en **hiérarchie**, chacun des membres appartenant à un niveau hiérarchique (ou niveau de granularité) particulier
 - Ex. Pour la dimension Temps, les années, les mois et les jours peuvent être des exemples de niveaux hiérarchiques. 1998 est un exemple de membre du niveau Année

Vocabulaire OLAP

Mesure :

- Une mesure est un élément de donnée sur lequel portent les analyses, en fonction des différentes dimensions
 - Ex. coût des travaux, nombre d'accidents, ventes, dépenses

Vocabulaire OLAP

Fait :

- Un fait représente la valeur d'une mesure, mesurée ou calculée, selon un membre de chacune des dimensions (ex. ce qui est recueilli par les systèmes transactionnels).
 - Ex. « le coût des travaux en 1995 pour la région 02 est 250 000 \$ » est un fait qui exprime la valeur de la mesure « coût des travaux » pour le membre « 1995 » du niveau « année » de la dimension « temps » et le membre « 02 » du niveau « région » de la dimension « découpage administratif ».

Vocabulaire OLAP

- **La table des faits:** comme son nom l'indique, contient les faits

Opération: NAVIGATION ou FORAGE

Les outils OLAP utilisent des opérateurs particuliers afin de « naviguer » dans les cubes multidimensionnels :

- **Pivoter (pivot, swap)** : Permet d'interchanger deux dimensions
- **Forer (drill-down)** : Permet de descendre dans la hiérarchie de la dimension. Ex. visualiser le nombre d'accidents par mois au lieu de par année.
- **Remonter (drill-up, roll-up)** : Permet de remonter dans la hiérarchie de la dimension. Ex. visualiser le nombre d'accidents par année au lieu de par mois.
- **Forer latéralement (drill-across)** :
 - Permet de passer d'une mesure à l'autre. Ex. visualiser le coût des travaux au lieu du nombre d'accidents
 - Permet de passer d'un membre de dimension à un autre. Ex. visualiser les données de Montréal au lieu de celles de Québec

Opération: agrégation

- Pour obtenir moins de détails
- Élimination d'une dimension ou regroupement des éléments d'une dimension
- Exemple:
 - Ville < Etat < Province < Pays
 - Au lieu de regrouper les données par ville, elles sont regroupées par pays

Les différentes technologies OLAP

- **MOLAP** (Multidimensionnel)
- **ROLAP** (Relationnel)
- **HOALP** (Hybride)
- **SOLAP** (Spatiale)

20/01/2009

OLAP

19

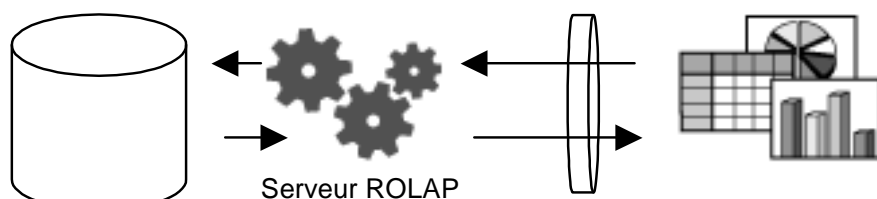


Base de données multidimensionnelle (hypercube)

Serveur MOLAP

Client OLAP

MOLAP



Base de données relationnelle (étoile ou flocon)

Serveur ROLAP

Vue multidimensionnelle

Client OLAP

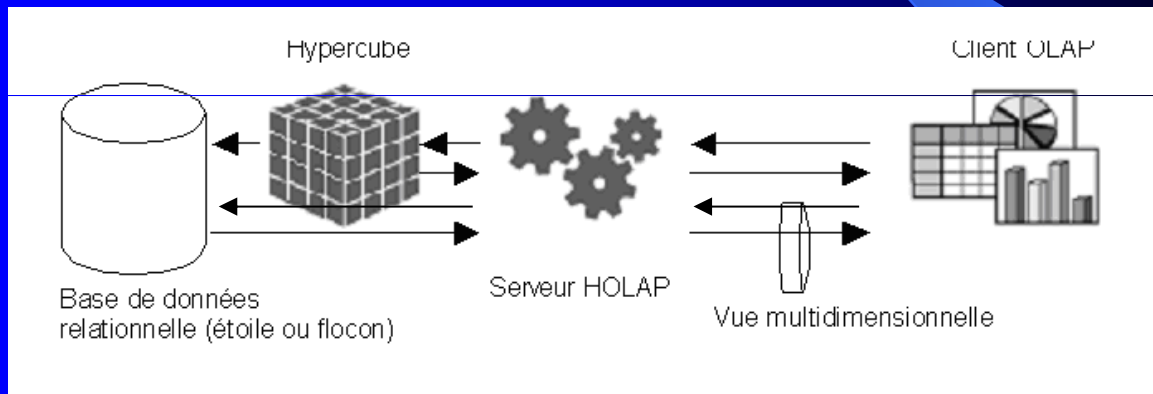
ROLAP

20/01/2009

OLAP

20

HOLAP



MOLAP vs ROLAP vs HOLAP

Critère de comparaison	ROLAP	MOLAP	HOLAP
Stockage des données de base (détaillées)	BD relationnelle	BD multidimensionnelle	BD relationnelle
Stockage des agrégations	BD relationnelle	BD multidimensionnelle	BD multidimensionnelle
Performance des requêtes (habituellement)	Le moins performant	Le plus performant	Performance moyenne

Plate formes et Outils

Pentaho

Kettle. → ETL

Mondrian, JPivot, Jrubik. → OLAP

Weka.

BIRT, JfreeReport, JasperReports, Pentaho Reporting.

Enhydra Shark.

Hibernate.

IDE Eclipse.

Java.

Jboss.

PHP et JSP.

HSQldb, MySQL.

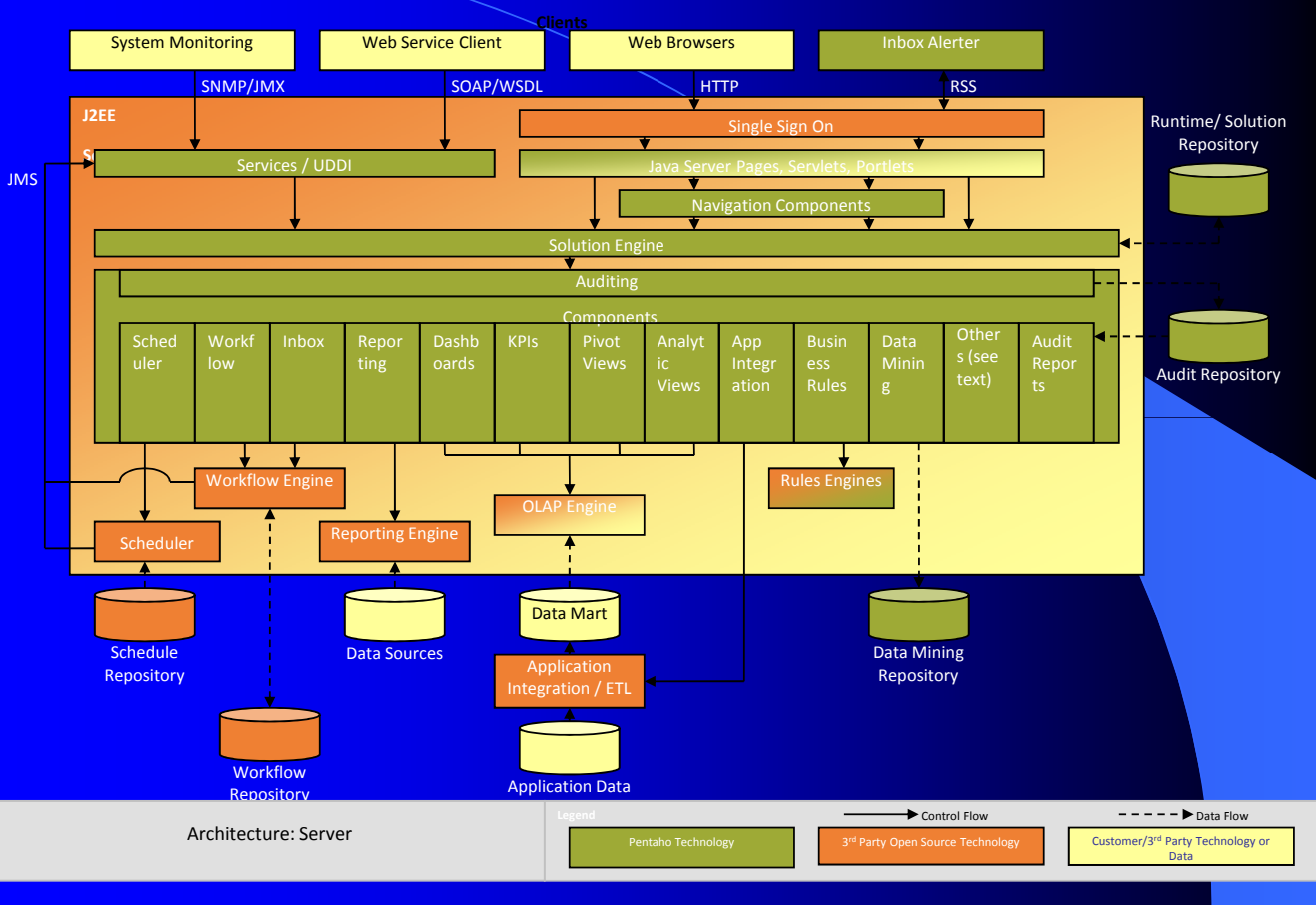
Quartz.

20/01/2009

OLAP

23

Architecture du serveur PENTAHO



Bibliographie

- <http://pagesperso-orange.fr/bernard.lupin/>
- http://www.journaldunet.com/solutions/0301/030108_olap.shtml
- <http://www.piloter.org/business-intelligence/olap.htm>
- <http://fr.wikipedia.org/wiki/OLAP>
- Cours informatique décisionnelle (NFE115) Cnam de basse-Normandie
Gilles LEBRUN et Christophe CHARRIER