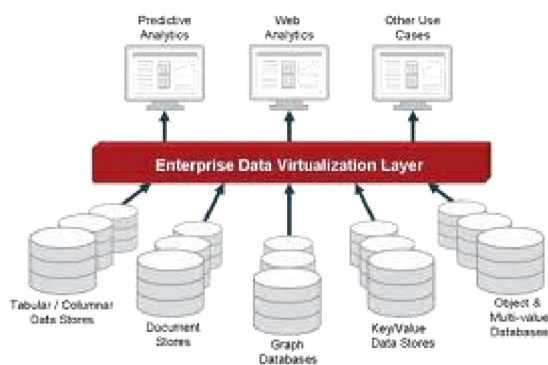




Blog Talks Publications ▾ Tools ▾ Radar / Plan ▾ Team About

"Data Virtualization" : perspectives et points d'attention

Posted on 2013-08-21 by [Isabelle Boydens](#)



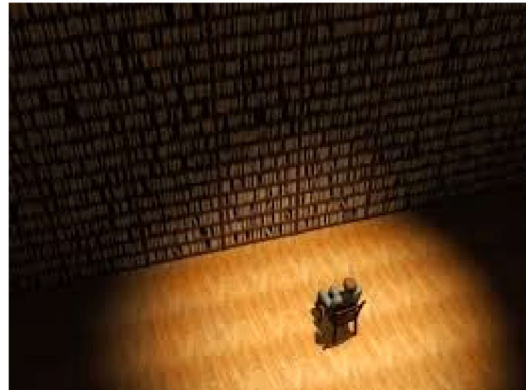
L'approche "Data Virtualization" émerge depuis quelques années dans la littérature et sur le marché IT (voir par exemple : R. F. Van der Lans, *"Data Virtualization for BI Systems. Revolutionizing Data Integration for Datawarehouse"*, Morgan Kaufmann, Elsevier, 2012).

A l'heure du "Big Data", cette approche propose de séparer couches physiques, logiques et conceptuelles, pour faciliter l'accès aux données via la modélisation d'une vue unique virtuelle. Sur le principe, l'approche n'est pas neuve mais la technologie est nouvelle car elle intègre une large gamme de modèles logiques actuels de données : RDBMS, NoSQL databases (modèles réseau, hiérarchique, ...), information non structurée, Web & Linked Open Data, etc.

Le "Data Virtualization" vise ainsi à la mise en place d'une plateforme d'accès et d'échange homogène à un grand ensemble de bases de données hétérogènes liées entre elles pour le "business". Des outils apparaissent sur le marché à cette fin (parmi les leaders en 2013, citons : [nodo](#), [Informatica](#) et [Composite Software](#) ainsi qu'un outil Open Source [Red Hat](#)) visant à gérer notamment :

- la performance (gestion de mémoires cache);
- la sécurité;
- la conception graphique de vues avec des fonctionnalités de "Sandbox" (prototypes rapidement conçus), d'audit ("lineage"), d'opérateurs spécifiques ("mapping", "matching", "nesting", ...);
- la génération d'une vue homogène moyennant éventuellement des pertes d'information (par exemple, les structures des NoSQL databases peuvent être converties en structures relationnelles pour obtenir une vue homogène, si la vue virtuelle est de type relationnelle);
- des transactions bidirectionnelles "théoriquement", toutefois, ces outils sont surtout utilisés à l'heure actuelle en "read only" et peu en "write back";
- certains aspects de la qualité des données (via le recours à des techniques de "data profiling", par exemple);
- des domaines d'application spécifiques;
- le lien avec des bases de données externes internationales fréquemment consultées, comme "Dun and Bradstreet", dans certains cas.

Sur **le plan conceptuel**, la mise en place de la vue virtuelle unique, reposant sur des requêtes (queries), n'est pas triviale et demande une **expertise humaine pointue** :



- des **techniques d'intégration de schémas** ("top down", "bottom up", "inside out", ...) et éventuellement, d'intégration de valeurs sur la base de règles de priorité doivent être mobilisées (voir par exemple : Ben Hassine-Guetari S. et al., « *La gestion de données multi-sources : de la théorie à la mise en œuvre dans le cadre d'un référentiel client unique* ». In Berti Equille L. éd., La qualité et la gouvernance des données au service de la performance des entreprises. Paris : Hermès, 2012, p. 179 à 214).
- la **gouvernance des données** (sur le plan sémantique, de la gestion des versions, ...) doit être gérée rigoureusement en amont de façon à disposer de données documentées et homogènes (ces questions sémantiques n'étant pas gérées "automatiquement" par un logiciel et

demandant de la main d'oeuvre humaine qualifiée et l'implication des utilisateurs). Cette gouvernance est capitale dès le début du projet mais aussi tout au long de la maintenance et du cycle de vie des applications. L'intérêt d'un outil de "data virtualization", sur ce plan, réside éventuellement dans le fait qu'il attire potentiellement l'attention sur la nécessité d'une gouvernance et d'une organisation.

- Une telle plateforme peut également faire appel à des **sources authentiques validées** dont la conception est stratégique (à propos de la conception de telles sources, voir l'ouvrage de référence : J. Bizingre, J. Paumier et P. Rivière, "*Les référentiels du système d'information. Données de référence et architecture d'entreprise*", Paris, Dunod, 2013).

Enfin, les points d'attention suivants doivent être mentionnés **sur le plan fonctionnel** :

- Par définition, les outils de "Data Virtualization" sont surtout intéressants si l'on dispose de **DBMS hétérogènes** différents sur les plans logiques et physiques. Si l'on dispose d'un seul type de DBMS, l'approche reste intéressante sur le principe mais l'acquisition d'un outil peut sembler superflue.
- La technologie est récente, en 2013, il existe **peu de case studies** à grande échelle de par le monde;
- Les outils sont **coûteux et parfois propriétaires** (voir plus haut).
- Même s'ils s'appliquent aussi aux bases de données "real time" transactionnelles, pour des raisons de performance évidentes, les domaines de prédilection actuels des outils de "Data Virtualization" sont le **BI**, les **ETL** ou les **DWH** (où l'on trouve fréquemment des DBMS logiques et physiques hétérogènes).

Les techniques de "Data Virtualization" présentent dans tous les cas un grand intérêt, dans le contexte du "data analytics" et du "big data" et mériteront un suivi attentif, en prenant en considération des défis conceptuels et fonctionnels qu'elles soulèvent.

big data

data quality

database management

information management

MORE POSTS

Je data beschermen tegen beheerders: 'on-premise' Confidential Computing

2026-03

Protéger ses données des administrateurs : l'informatique confidentielle « on-premise »

2026-03

De performance van LLM's: Een vergelijkende analyse tussen Frans en Nederlands

2026-03

Made by Smals Research – Privacyvriendelijk Kruisen van Persoonsgegevens

2026-02

Search

Newsletter & webinars:

Dutch French

Keywords:

[analytics](#) [artificial intelligence](#) [big data](#) [blockchain](#) [bpm](#) [chatbot](#)
[cloud computing](#) [cost cutting](#) [cryptography](#) [data center](#)
[data quality](#) [development](#) [eda](#) [egov](#) [event](#) [gis](#)

information management machine learning managing it costs
methodology mobile natural language processing open source privacy
productivity security social software design
software engineering standards

Smals Research

© Smals Research – License/Disclaimer: FR / NL