

La solution Microsoft BI

1. L'offre Microsoft BI

Microsoft n'est pas un nouveau venu dans le monde de la BI. Depuis le tout début des années 2000, il en est même un acteur majeur et ce, grâce à sa base OLAP Analysis Services et au tableau croisé dynamique Excel.

Bien entendu, son offre s'est largement étendue et structurée depuis, avec le développement de ce nouveau marché. Il est à noter tout de même que Microsoft a largement contribué à démocratiser le recours à l'informatique décisionnelle dans les sociétés de taille moyenne, ainsi que dans les PME.

Sur le fond, l'offre Microsoft BI est structurée autour des trois promesses du décisionnel :

- Améliorer l'accès et la qualité des données : on y retrouve tous les outils destinés à concevoir un entrepôt de données bien modélisé, performant et contenant des données fiables. On y retrouve aussi tous les outils permettant d'identifier et de gérer les données de référence et de s'assurer de leur qualité.
- Gagner en finesse d'analyse et de compréhension de données : on y retrouve tous les outils qui permettent aux utilisateurs finaux d'analyser et de naviguer dans leurs données en toute autonomie, sans avoir à recourir au service informatique.
- Gérer les performances de l'organisation et de ses politiques : on y retrouve tous les outils destinés à partager, à communiquer et à organiser les performances de l'organisation tels que les outils de tableaux de bord et les outils d'intranet.

Sur la forme, l'offre Microsoft BI est structurée au sein de trois licences. Chacune de ces licences contient de nombreux outils à usage décisionnel correspondant aux tâches énoncées plus haut :

- SQL Server 2014 : Integration Services, Master Data Services, Analysis Services, Reporting Services, Data Quality Services.
- Office 2013 : Excel 2013, PowerPivot
- SharePoint Server 2013 : Excel Services, PerformancePoint Services.

Si les outils de la gamme SQL Server 2014 sont plutôt des produits techniques destinés, à l'usage, au service informatique, les outils des gammes Office 2013 et SharePoint 2013 sont plus spécifiquement destinés aux utilisateurs finaux.

Dans la suite de cette partie, nous allons étudier le contenu, l'utilisation et le positionnement de chacune de ces solutions.

Néanmoins, dans cet ouvrage nous n'étudierons pas en détail les outils de la gamme SharePoint Server. Les explications et illustrations de ce livre s'appuient sur la version 2013 de SharePoint Server et de la suite Office mais s'applique tout aussi bien à SharePoint Server 2010 et Office 2010. Les chapitres qui suivent se concentreront sur le bon usage des produits de la gamme SQL Server 2014, ainsi que d'Excel 2013. Pourtant nécessaire et très complémentaire, l'étude de l'utilisation de SharePoint est un sujet en soi, qui est étudié et longuement développé dans l'ouvrage Business Intelligence avec SharePoint Server - Créez un portail décisionnel et pilotez les performances.

	Utilisent	Conçoivent	
Décideurs stratégiques	×		Tableaux de bord stratégiques (KPI)
Décideurs tactiques	×		Tableaux de bord d'activité
Analystes		×	SharePoint 2013
	×		Office 2013
Décideurs opérationnels	×		SQL Server 2014
Informaticiens		×	
			Tableaux de bord
			Outils d'analyse et outils de statistiques
			Rapports opérationnels et rapports de suivi d'activité

L'utilisation des licences Microsoft suivant les profils utilisateurs

2. SQL Server 2014

Si à l'origine, la licence SQL Server correspond uniquement à une base de données relationnelle (OLTP), assez rapidement la licence s'étoffe pour couvrir l'ensemble des outils dédiés au stockage et au traitement de données. Dans le langage courant, SQL Server évoque la base de données relationnelle. Il existe néanmoins d'autres outils ou services couverts par cette même licence, dont la plupart trouve un usage dans le cadre de la mise en œuvre d'un système décisionnel.

Nous ne nous attarderons donc pas sur le service de base de données relationnelle, qui a maintenant fait ses preuves face à la concurrence. La base de données est fiable, performante et hautement disponible. Le service de base de données est parfaitement indiqué pour accueillir un entrepôt de données. On retrouve d'ailleurs des entrepôts de données sous SQL Server Database Services dans de grandes banques nationales, dans de grands consortiums hospitaliers, dans la grande distribution, dans des sociétés de télécom et de téléphonie... SQL Server se retrouve dans les sociétés et organisations de toutes les tailles et dans tous les domaines d'activité.

Nous avons vu précédemment que la construction d'un système décisionnel ne se limite pas à l'utilisation d'une base de données relationnelle. Avec la licence SQL Server, Microsoft nous livre toute la panoplie d'outils dont un service informatique a besoin pour bâtir un système décisionnel dans les règles de l'art.

Pour bâtir notre système d'aide à la décision, nous aurons besoin de :

- SQL Server Integration Services : l'ETL.
- SQL Server Master Data Services : le gestionnaire des données de référence.
- SQL Server Analysis Services : la base de données multidimensionnelle (OLAP) et le métamodèle.
- SQL Server Data Quality Services : la base de connaissances permettant de qualifier et nettoyer certaines données constitutives de vos référentiels de données ou des bases marketing.
- SQL Server Reporting Services : l'outil de reporting opérationnel et de reporting de masse.

Dans les parties qui suivent, nous allons présenter plus en détail chacun de ces outils.

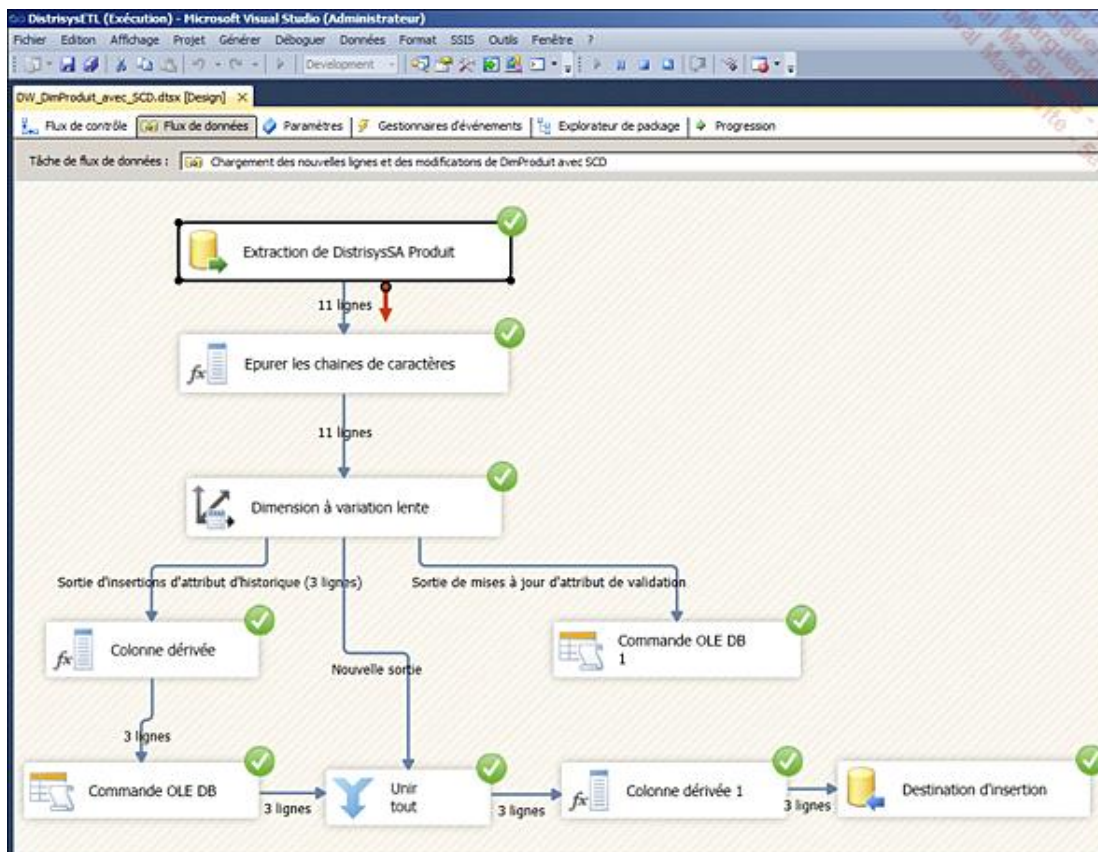
a. SQL Server Integration Services

Nous avons expliqué précédemment que les concepts du décisionnel prennent racine dans le fait de bénéficier d'entrepôt de données contenant et croisant des données provenant de systèmes sources très hétérogènes.

Si la grande valeur ajoutée du décisionnel est d'accéder confortablement aux données contenues dans l'entrepôt de données, il n'en reste pas moins que la majeure partie d'un projet décisionnel se situe dans l'alimentation de l'entrepôt de données et la mise à jour des données de référence.

En effet, la constitution des données de référence et l'alimentation d'un entrepôt de données représentent généralement près de 80 % de la charge du projet. De prime à bord, beaucoup de services informatiques qui découvrent le décisionnel ont largement tendance à sous-estimer le temps nécessaire à récupérer l'information, mais aussi à la croiser.

Pour nous aider dans cette tâche, Microsoft met à notre disposition, dans la licence SQL Server 2014, un outil d'ETL (*Extract Transform Load*) : SQL Server Integration Services (SSIS).



Déroulement d'un flux ETL sous SSIS

Même s'il n'est pas considéré comme le meilleur ETL du marché, il n'en reste pas moins que SSIS est un véritable outil d'ETL, fiable, performant, et qu'il remplit parfaitement le rôle pour lequel il est destiné. SSIS est d'ailleurs maintenant largement déployé et employé dans de très nombreuses entreprises françaises et étrangères.

➤ Nous ne vous conseillons pas d'acheter la licence SQL Server 2014 pour bénéficier spécifiquement de SSIS, surtout, si vous travaillez sur un décisionnel en environnement Linux avec des technologies open source... En revanche, si vous avez pour projet de bâtir un système décisionnel en environnement Microsoft, le rapport qualité-prix de l'outil est véritablement imbattable.



Le point fort de SSIS est que c'est d'une part un bon ETL et qu'il est d'autre part disponible sans surcoût avec la licence SQL Server. Une licence que la plupart des entreprises disposent déjà afin de bénéficier de la base de données relationnelle.

Plus que l'outil, c'est la manière dont il est employé qui est très importante. À ce propos, le chapitre Alimenter l'entrepôt de données avec SSIS, reviendra assez largement sur le sujet afin de vous donner les bonnes pratiques d'usage. Nous vous invitons donc à consulter ce chapitre, à la fois pour découvrir les bonnes pratiques de l'ETL et pour apprendre à vous familiariser avec SSIS.

Généralement, on attend d'un outil d'ETL :

- Qu'il accélère le travail de développement des flux de données : SSIS est un ETL assez simple d'utilisation. Sa prise en main est rapide. SSIS, comme les grands ETL, permet de découper un flux d'alimentation en une multitude de petites tâches de transformation de données distinctes et ordonnancées. L'amélioration de la productivité vient du fait qu'il est plus facile de traiter une multitude de problèmes très simples, plutôt que de traiter un grand problème très compliqué. SSIS permet aussi de suivre très précisément le déroulement du flux de données. Entre chaque tâche de transformation de données, il est possible de visualiser les valeurs, ainsi que les transformations qui leur ont été appliquées.
- Qu'il offre une vision claire et maintenable des flux réalisés : SSIS, comme tout bon ETL, permet d'obtenir une visualisation graphique, logique et simple des flux réalisés. Les règles de transformation pouvant être parfois très compliquées, il est important qu'elles puissent être représentées simplement afin que n'importe quel informaticien, disposant d'une formation ETL, puisse lire et comprendre le déroulement d'un flux de données. Chaque tâche de transformation est représentée par une boîte que le développeur a la liberté de nommer ou plutôt de décrire. Développer un flux ETL avec SSIS ne nécessite pas de notions de développement objet ni même d'un grand savoir-faire SQL. Un flux SSIS correctement développé et respectant la logique ETL est un flux vraiment très lisible et facilement maintenable.
- Qu'il puisse se connecter et travailler avec de nombreuses sources hétérogènes : le propre d'un ETL est d'être ouvert et de pouvoir disposer d'une certaine universalité de connexion. SSIS ne déroge pas à cette règle, il permet entre autres de se connecter nativement à :
 - de nombreuses bases relationnelles par OLEDB : SQL Server, Oracle, DB2...
 - des fichiers : fichier plat CSV, fichier XML, fichier Excel.
 - des sources en ligne : services web, service FTP.
 - etc.

La liste est bien entendu loin d'être exhaustive. Les connecteurs sont cependant de qualité très inégale, mais comme tous les produits Microsoft, le produit est ouvert et la communauté de développeurs est assez efficace. Vous trouverez sur le marché de nombreux autres composants, connecteurs ou tâches, et vous disposerez de la possibilité de créer le vôtre en .NET si vraiment cela s'avère nécessaire. Ce qui est en définitive assez rare toutefois.



Au cours de nos projets, nous avons eu par exemple à développer une tâche récupérant les fichiers en pièces jointes des mails d'une boîte aux lettres électronique. Le besoin est peu banal mais la réalisation a été simple.

- Qu'il soit performant : un ETL dispose souvent de fenêtres de traitement très courtes pour se connecter à un système source et pour charger l'entrepôt de données. Il faut alors que l'outil traite de très gros volumes, très rapidement. Le secret de la performance des outils d'ETL réside généralement dans leur capacité à travailler et faire les transformations sur les données en mémoire vive. C'est le cas de SSIS, qui dispose d'ailleurs de nombreuses possibilités de gestion et d'optimisation du cache.
- Qu'il dispose de nombreuses fonctionnalités de transformation de données : en ce qui concerne la lettre T de ETL, SSIS est plutôt bien fourni et propose de nombreuses tâches de transformation : calcul, contrôle, mise en cohérence des données, conversion, pivotement, union, jointure, nettoyage, regroupement, échantillonnage... SSIS intègre de plus en plus de tâches de transformation avancées issues de la statistique, pour vous permettre de mieux nettoyer

vos données ou contrôler leurs contenus.

- Qu'il puisse se déployer facilement : SSIS gère parfaitement les problématiques de multi-environnements et de déploiement sur un environnement de production. Les flux se lancent et se planifient, soit par le biais de l'agent SQL fourni et intégré à SQL Server, soit par le biais d'une ligne de commande exécutable, si vous disposez de votre propre ordonnanceur d'entreprise.

Enfin SSIS permet de mettre en œuvre pleinement les concepts du métier : audit de données, historisation des dimensions (SCD), traitement des erreurs et des rejets... En ce qui concerne la gestion d'erreur, SSIS gère l'erreur au niveau de chacune de ses tâches et potentiellement au niveau de chacune des lignes chargées. SSIS vous permet ainsi de mettre en place des systèmes d'audit de flux et de données performants et entièrement personnalisés à votre activité et à votre problématique. Cependant, ce genre de système n'est pas quelque chose de standard inclus magiquement dans un outil, il vous faudra le mettre en place. Nous vous incitons à consulter sur ce sujet le chapitre Alimenter l'entrepôt de données avec SSIS - L'audit des flux ETL.

b. SQL Server Master Data Services

SQL Server Master Data Services (MDS) est disponible depuis la version SQL Server 2008 R2. Ce service est une solution de MDM (*Master Data Management*) ou Management des données de référence.

La gestion des données de référence n'est pas exclusive au système décisionnel. Il s'agit plutôt d'une pratique d'urbanisation des systèmes d'information qui contribue sensiblement à la qualité de l'information dans l'entreprise.

Les données de référence sont les données transversales de l'entreprise. Ce sont les éléments clés qui décrivent et définissent un domaine de l'entreprise : clients, produits, fournisseurs, sites, organisations, services, employés...

Gérer ces données de référence devient primordial pour la plupart des entreprises qui ont une organisation cloisonnée des données.

Une société multinationale qui dispose d'un ERP, d'une GPAO ou d'un CRM pour chacune de ses filiales, est forcément confrontée à un moment ou un autre à des problèmes de redondance ou d'incohérence de données. Le risque est le même chaque fois qu'une organisation fait le choix de déployer une solution spécialisée pour gérer une activité spécifique de son activité : déployer un CRM pour les commerciaux, une GPAO pour la production, une GMAO pour les services maintenance, un logiciel de paie spécifique à une branche d'activité... Lors de chaque déploiement, les données de référence sont alors dupliquées. Si l'on n'y fait pas attention, il devient rapidement extrêmement complexe de rapprocher les données de ces différents systèmes.

Le MDM regroupe les pratiques, méthodes et outils qui permettent de rassembler, gérer et partager les données pour lesquelles il est primordial de fournir à l'ensemble de l'organisation une vision contrôlée et unique. Le rôle du MDM est de fournir aux différents systèmes des données de référence exactes, complètes et actualisées.

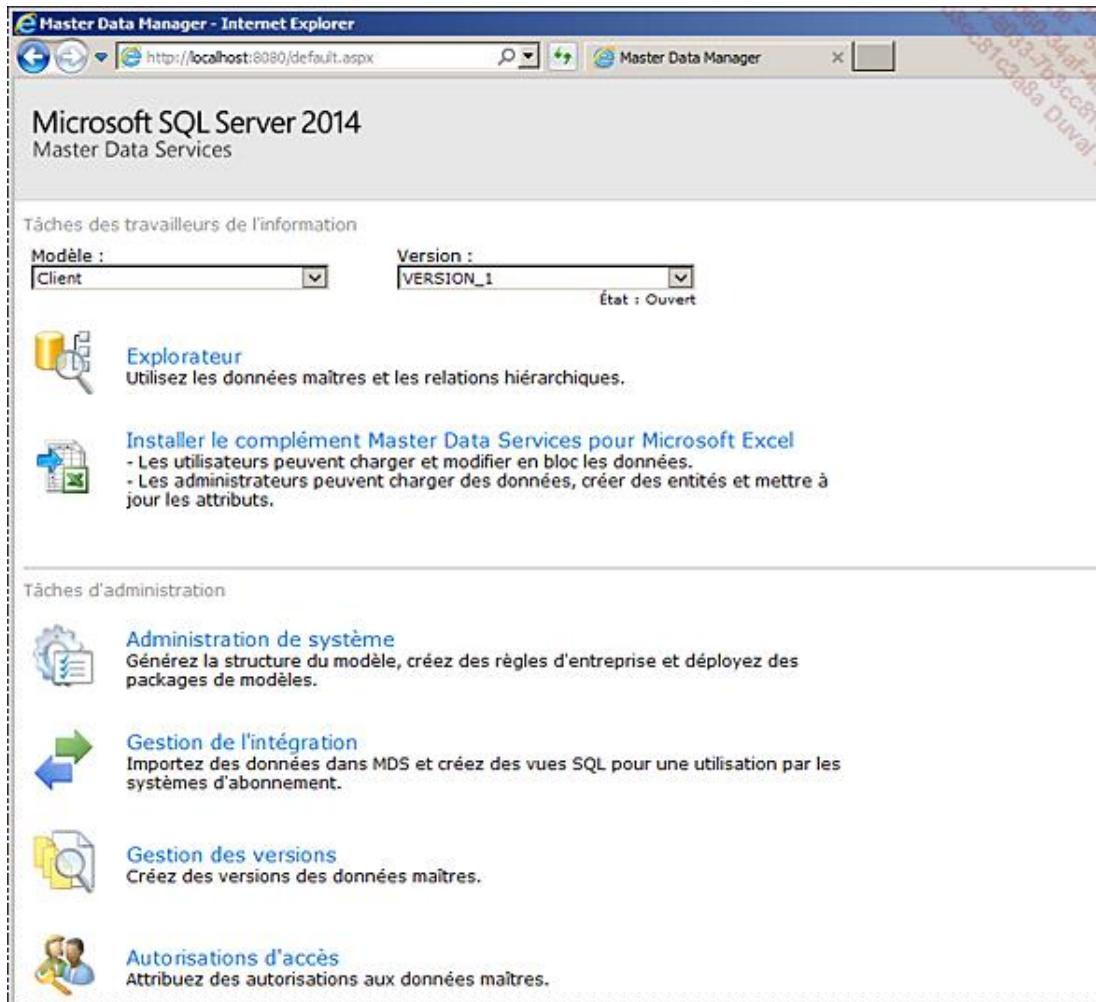
Le MDM crée une ressource centralisée, indépendante des applications et des processus métier, qui gère le cycle de vie des données de référence.

Avec la mise en place d'une telle pratique impliquant les services fonctionnels et le service informatique, la cohérence des données dans les divers systèmes de transactions et d'analyses est garantie. Ainsi, les problèmes de qualité des données peuvent être résolus de manière proactive, plutôt qu'après coup dans l'entrepôt de données.

Microsoft fournit donc à travers MDS, une solution informatique appuyant et permettant la mise en œuvre d'une démarche de MDM dans votre organisation. Le MDM n'est pas spécifiquement l'objet de l'ouvrage, mais elle est abordée dans le chapitre Gérer les données de référence avec MDS. En effet, il est essentiel que vous compreniez l'importance et l'intérêt d'une telle solution lors du déploiement de votre système décisionnel. Le rapprochement des données entre les différents systèmes est l'un des principaux enjeux de l'ETL, mais aussi l'une des raisons

majeures des délais de développement ETL.

Lorsque vous serez confronté à un projet d'entrepôt de données, une des phases techniques préalables de votre projet consistera à identifier les données de référence. Et si celle-ci ne sont pas définies, il vous faudra alors les créer, les alimenter et les mettre à jour et ce, avant même d'avoir créé la base de données de l'entrepôt de données. L'outil SQL Server Master Data Services vous sera alors d'une bonne aide pour faire émerger ce nouveau système dans votre organisation. Un peu comme SSIS, MDS n'est pas la meilleure solution du marché, mais elle s'avèrera très suffisante pour la grande majorité des organisations et dispose d'un rapport qualité-prix inégalable. C'est la solution idéale pour faire émerger un projet de MDM en douceur, en parallèle du projet décisionnel.



Interface d'administration de MDS

Avec MDS, vos utilisateurs vont aussi apprécier de retrouver leurs données de référence dans Excel. La gestion de ces données se fera directement depuis un classeur Excel, facilitant ainsi la conduite du changement et l'adoption du MDM dans votre organisation.

	D	E	F	G	H	I
	Name	Code	Code alpha 2	Nom ISO	Continent	
3	Albanie	ALB	AL	ALBANIE	EU (EUROPE)	
4	Andorre	AND	AD	ANDORRE	EU (EUROPE)	
5	Argentine	ARG	AR	ARGENTINE	MS (Amérique du Sud)	
6	Arménie	ARM	AM	ARMÉNIE	EU (EUROPE)	
7	Autriche	AUT	AT	AUTRICHE	EU (EUROPE)	
8	Azerbaïdjan	AZE	AZ	AZERBAÏDJAN	EU (EUROPE)	
9	Burundi	BDI	BI	BURUNDI	AF (AFRIQUE)	
10	Belgique	BEL	BE	BELGIQUE	EU (EUROPE)	
11	Bénin	BEN	BJ	BÉNIN	AF (AFRIQUE)	
12	Burkina Faso	BFA	BF	BURKINA FASO	AF (AFRIQUE)	
13	Bangladesh	BGD	BD	BANGLADESH	AS (Asie)	
14	Bulgarie	BGR	BG	BULGARIE	EU (EUROPE)	
15	Bahreïn	BHR	BH	BAHREÏN	AF (AFRIQUE)	
16	Bahamas	BHS	BS	BAHAMAS	EU (EUROPE)	
17	Bolivie	BOL	BO	BOLIVIE, ÉTAT PLURINATION	MS (Amérique du Sud)	
18	Brésil	BRA	BR	BRÉSIL	MS (Amérique du Sud)	
19	République centrafricaine	CAF	CF	CENTRAFRICAINE, RÉPUBLIQUE D'AFRIQUE		

La gestion des données de référence dans Excel avec MDS

c. SQL Server Data Quality Services

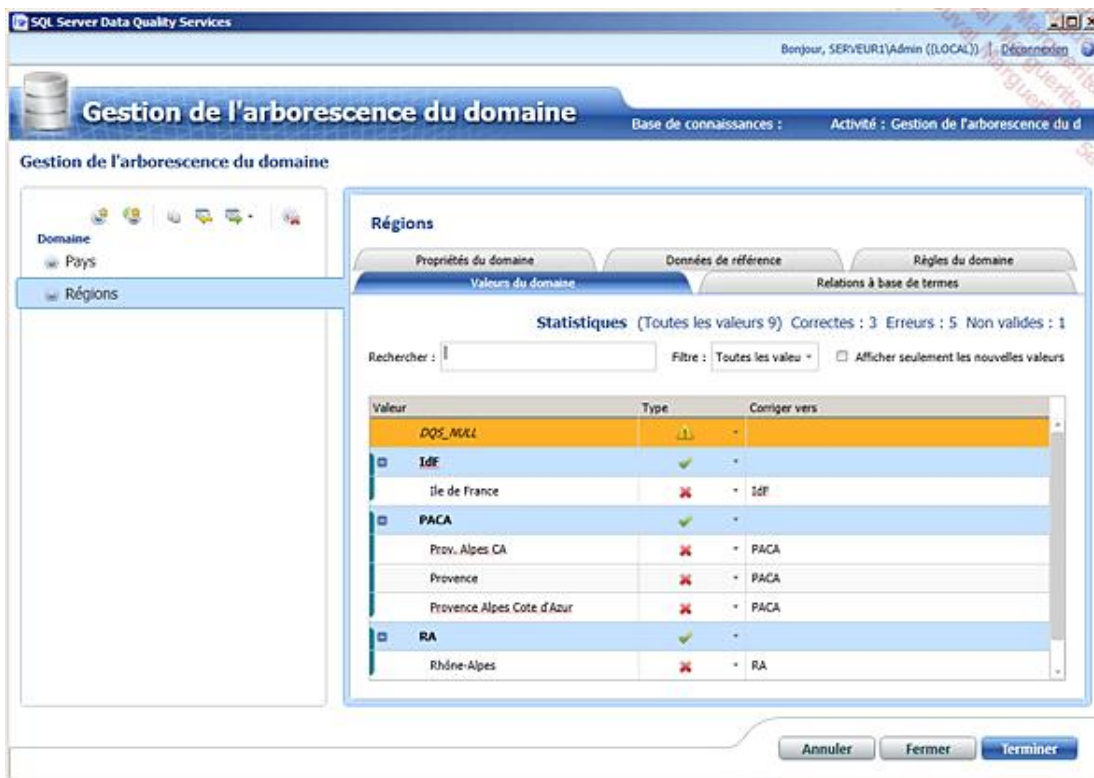
SQL Server Data Quality Services (DQS) a été une nouveauté de SQL Server 2014. Cette solution va vous permettre de contrôler et améliorer la qualité des données qui circulent au sein de l'organisation à l'aide de bases de connaissances liées au contenu et à la structure de vos données.

Typiquement, on peut utiliser DQS pour dé-doublonner une base client, contrôler le format d'un numéro de téléphone ou d'un code postal, qui doivent être composés de 10 et 5 chiffres. On peut également implémenter des contrôles plus fonctionnels, comme par exemple un employé ne peut pas avoir plus de 90 ans et moins de 16 ans.

Les bases de connaissances vont également vous permettre de capitaliser sur la correction des erreurs. Ainsi, une correction manuelle d'une erreur pourra être transformée en règle qui corrigera ensuite automatiquement toute donnée comportant la même anomalie.

Cette solution trouve tout son sens dans la constitution de fichier client ou lors de la récupération de données provenant de fichiers non structurés saisis par divers employés, fournisseurs ou partenaires. Dans ce cas, DQS offre la possibilité d'auditer une base ou un fichier pour vous signaler tous les champs en désaccord avec la base de connaissances.

Cette base de connaissances employée conjointement avec MDS et SSIS va permettre d'identifier la non-qualité des données et faciliter le travail de nettoyage des fichiers client, prospect, employé... Mais l'utilisation dans le cadre d'un projet strictement décisionnel est limitée ou du moins annexe, nous ne l'aborderons donc pas dans cet ouvrage.



Base de connaissances DQS

d. SQL Server Analysis Services

SQL Server Analysis Services regroupe en fait plusieurs services destinés à faciliter l'exploration d'une quantité importante de données.

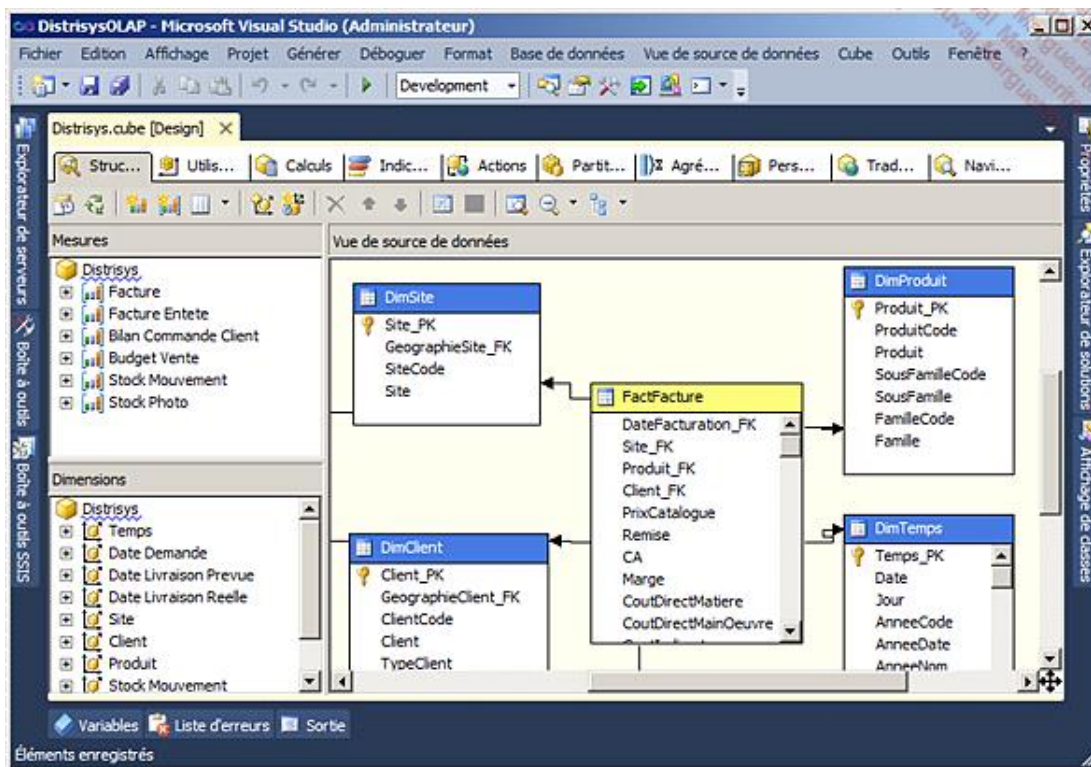
Les trois facettes de la solution sont :

- Le mode multidimensionnel
- Le mode exploration de données ou data mining
- Le mode tabulaire

Le mode multidimensionnel

Il s'agit du mode historique de SQL Server Analysis Services (SSAS). C'est ce mode-ci que l'on désigne peut-être abusivement sous la dénomination SSAS.

Dans ce mode, SQL Server Analysis Services (SSAS) est la base multidimensionnelle (OLAP) de la licence SQL Server. On présente souvent SSAS comme étant la solution de cubes de Microsoft. SSAS est une solution parfaitement fiable et très robuste, leader de son marché.



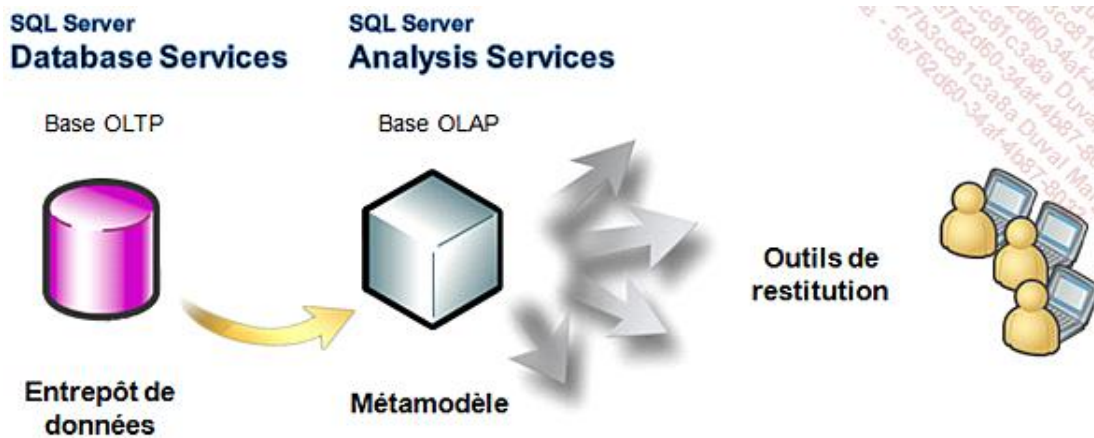
Interface de conception Analysis Services

Avec l'augmentation des volumes de données, les bases de données OLAP se sont imposées comme des solutions incontournables pour représenter les données contenues dans l'entrepôt de données. À l'opposé des bases OLTP, plus les requêtes utilisateurs portent sur des données globales et agrégées, plus la réponse est rapide. À l'inverse, plus la requête porte sur les données de détails, moins la requête est performante. Il ne faut donc pas considérer Analysis Services comme une base de données permettant des extractions de données. Les modèles Analysis Services doivent être conçus pour fournir la finalité de l'analyse attendue par l'utilisateur.

Les données détaillées de l'entrepôt de données sont contenues dans la base relationnelle, mais c'est Analysis Services qui les agrège et les présente aux utilisateurs finaux. Étant la partie émergée de l'entrepôt de données, il offre ainsi la possibilité de gérer véritablement de très grands volumes de données avec des temps de réponse de l'ordre de la seconde. Les résultats sont souvent jugés époustouffants pour des utilisateurs habitués à travailler avec des bases de données OLTP. SSAS ne craint pas les gros volumes de données, il est conçu pour cela. Attention toutefois à conserver une modélisation appropriée, car la performance de son moteur OLAP est directement liée à la modélisation de la base de données sous-jacente.

Nous ne saurions que trop insister tout au long de l'ouvrage sur l'importance de la modélisation de l'entrepôt de données. Nous vous suggérons à ce sujet les chapitres Réaliser son premier système décisionnel, puis La modélisation dimensionnelle, pour vous apprendre à concevoir des modèles d'analyse cohérents et performants.

Les données de l'entrepôt de données étant manipulées directement et exclusivement par le biais d'Analysis Services, celui-ci offre aux utilisateurs une interface simplifiée d'accès aux données et intuitive. Dans les faits, les utilisateurs n'accèdent pas directement à Analysis Services. Ce sont les outils de restitution qui proposent et consomment les données du cube. Il est donc indispensable, lorsque vous mettez en place SSAS, de proposer parallèlement aux utilisateurs, des outils de restitution adaptés (par exemple Excel).



Analysis Services est l'interface entre l'entrepôt de données et les outils de restitution

Le mode exploration de données

Cette autre facette de SQL Server Analysis Services est dédiée à l'élaboration de modèles statistiques. Cette solution est installée et disponible nativement en même temps que la solution SSAS en mode multidimensionnel. Dans ce mode, SQL Server offre de faire du data mining, c'est-à-dire la possibilité de créer des modèles statistiques.

Le data mining est une facette des solutions décisionnelles mais rarement employée en France de manière intégrée au système décisionnel de l'entreprise. On retrouve par contre assez largement ces solutions dans les services marketing ou prospection des entreprises ou des associations. Ces solutions sont alors employées quotidiennement pour aider l'entreprise à se développer.

Sans rentrer dans le détail, le mode explorateur de données de SSAS propose plusieurs algorithmes :

- Les algorithmes de segmentation : ces algorithmes se proposent à partir d'une population donnée, de distinguer certains groupes jugés récurrents afin de regrouper cette population en fonction de leur comportement ou de leurs propriétés intrinsèques...

Par exemple, c'est ce type d'algorithme qui permet de segmenter les clients d'une société de vente par correspondance. Les classes ou segments sont obtenus en fonction du comportement d'achat constaté par le passé et par certains attributs : l'âge, le sexe, le revenu annuel, la domiciliation...

- Les algorithmes de classification : ces algorithmes proposent pour un élément donné l'appartenance à une classe.

Par exemple, lorsqu'un nouveau prospect est entré dans la base, le modèle est alors capable de déterminer de quel segment l'individu fait partie. C'est ce type de modèle qui permet à une banque ou à une société d'assurances par exemple de juger statistiquement le risque encouru par le prospect et donc de lui proposer des prix ou des offres adaptés.

- Les algorithmes de régression : ces algorithmes permettent à partir des données passées d'anticiper les résultats avenir. Ce sont des algorithmes de prévisions temporelles.

Par exemple, un service marketing qui réalise depuis un certain nombre d'années des opérations commerciales similaires peut anticiper le résultat d'une action commerciale à venir.

- Les algorithmes d'association : ces algorithmes se proposent de trouver des corrélations entre des comportements.

Ces algorithmes sont très utilisés dans la détection des paniers d'achat. Par exemple, le modèle peut constater que

statistiquement une personne achetant un jouet électrique achètera aussi des piles électriques, ou alors une personne achetant une pizza surgelée achètera un soda.

Souvent les résultats des modèles sont anticipables : nous n'avons pas besoin d'un ordinateur pour deviner que l'achat de rouleaux à tapisserie s'accompagnera de la colle à tapisser. Mais parfois, ils sont plus surprenants et c'est là tout l'intérêt de ce type d'algorithme.

- Les algorithmes d'analyse de séquence : ces algorithmes proposent d'identifier les séquences ou épisodes fréquents dans les données.

Par exemple, utilisé sur un site web, le modèle peut identifier le ou les séquencements de pages les plus couramment utilisés. Il peut aussi prévoir à partir d'une page, la ou les pages suivantes que consultera un utilisateur.

La solution est assez complète et propose une intégration parfaite à Excel.

Toutefois, il faut bien comprendre que ces algorithmes ne sont pas magiques. Il faut toute l'expérience du service marketing d'une organisation pour élaborer un modèle. Un modèle valide étant un modèle offrant de bien meilleurs résultats que le hasard, durablement et ce, quel que soit l'échantillon qu'on lui propose. C'est un métier qui est assez éloigné du décisionnel, mais une fois que ce modèle est élaboré et validé, le système décisionnel pourrait bien entendu parfaitement l'intégrer pour générer des objectifs ou des cibles à certains indicateurs par exemple. Nous verrons en conclusion que c'est une ouverture vers laquelle pourraient tendre ou prétendre les systèmes décisionnels les plus matures.

Le mode tabulaire

Le mode tabulaire de SQL Server Analysis Services a été une nouveauté de la version 2012, et c'est la plus grande évolution de SQL Server Analysis Services depuis la version 2005 et la sortie du mode multidimensionnel actuel.

Tout comme le mode multidimensionnel, le mode tabulaire offre un modèle sémantique qui se veut ouvert à tous les types de restitution : reporting, analyse, tableau de bord...

Le mode tabulaire propose une autre façon de concevoir des modèles et de les stocker. Contrairement au mode multidimensionnel qui nécessite une modélisation dimensionnelle (cf. chapitre La modélisation dimensionnelle) et un stockage OLAP, le mode tabulaire utilise une modélisation relationnelle et un moteur d'analyse en mémoire vive pour le stockage et le calcul.

De ce fait, le mode tabulaire va être plus proche des modèles relationnels classiques, tout en permettant une vitesse de restitution de données extrêmement rapide de la donnée détaillée comme de la donnée agrégée. Par contre, il ne dispose pas de la maturité et de toutes les fonctionnalités du mode multidimensionnel.

En ce qui nous concerne, nous considérons qu'il n'est toujours pas concevable avec la version SQL Server 2014 de faire reposer un entrepôt de données d'entreprise sur le mode tabulaire. Trop de fonctionnalités sont encore manquantes : relation plusieurs à plusieurs, traduction, sécurité à la cellule... Les investissements de Microsoft étant toutefois tournés vers ce nouveau mode, il semblerait que les versions à venir vont combler progressivement l'écart de fonctionnalités.

Actuellement, la solution répond très bien soit à une approche Business Intelligence en silo pour répondre à un besoin sectoriel, soit à un besoin de reporting sur une base applicative existante.

En ce qui concerne l'interface de conception, elle est quasiment identique à celle utilisée par le PowerPivot d'Excel, mais incluse dans l'outil SQL Server Data Tools (SSDT). De ce fait, même si dans cet ouvrage nous ne traiterons pas du mode tabulaire, le chapitre Restituer les données décisionnelles - Excel vous donnera un aperçu plus que satisfaisant de la solution et de ses capacités.

Pour utiliser le mode tabulaire, il vous faudra toutefois créer une instance dédiée et nous vous conseillons de dédier une machine physique et puissante à cette instance.

e. SQL Server Reporting Services

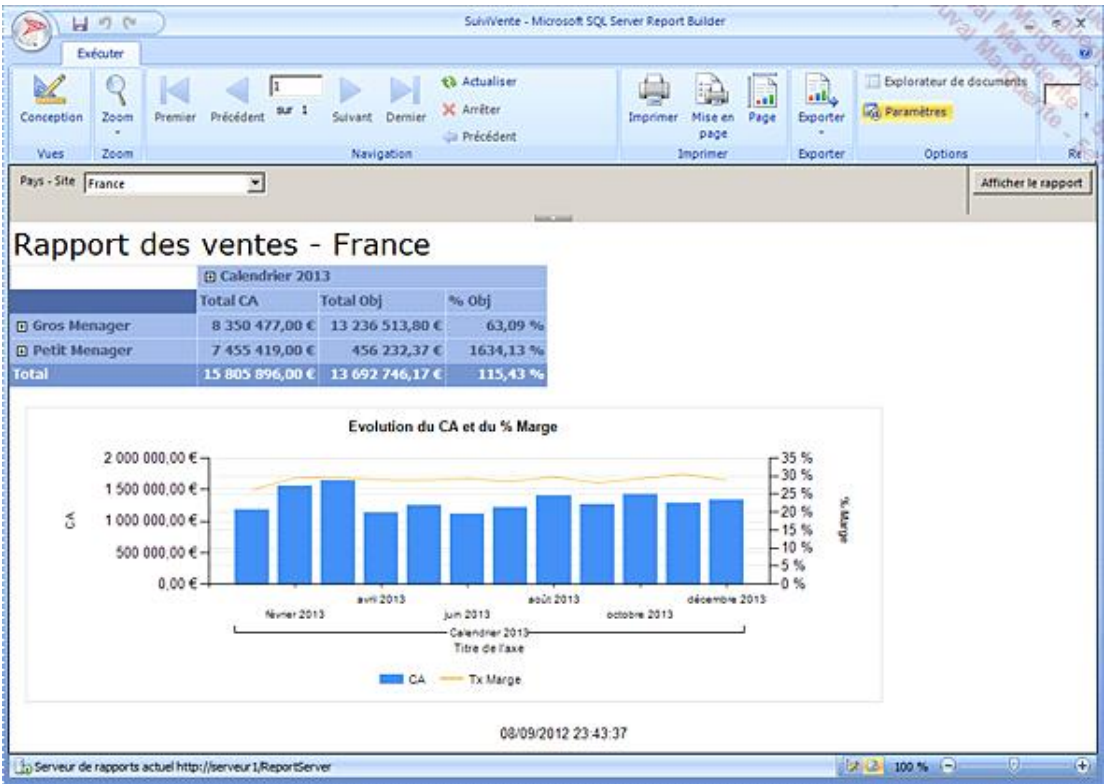
Enfin, la suite SQL Server dispose d'un serveur de rapports permettant d'afficher et de diffuser des informations.

Reporting Services est avant tout un produit destiné à un public d'informaticiens. Sa parfaite intégration avec l'environnement .NET, sa capacité de mise en page, ses possibilités de diffusion en font un excellent outil de reporting opérationnel. Reporting Services est l'outil parfait pour mettre en page une facture, un bon de livraison, un suivi de commande, un inventaire, un catalogue produit, la liste des clients à relancer... Tous les états dont une application de gestion a besoin sont parfaitement réalisables avec Reporting Services.

En ce qui concerne le besoin décisionnel, Reporting Services a été survendu par Microsoft depuis la sortie de SQL Server 2005. Reporting Services n'est pas un outil d'analyse destiné à des utilisateurs finaux. En revanche, il a totalement sa place dans la diffusion de rapports de masse, c'est-à-dire dans les rapports décisionnels destinés à de nombreux décideurs opérationnels.

Par exemple, la diffusion par mail au format PDF du rapport mensuel de suivi des ventes à tous les commerciaux de l'entreprise.

Nous reviendrons largement sur ce sujet dans le chapitre Restituer les données décisionnelles - Reporting Services. Nous aborderons alors plus en détail la différence entre rapport opérationnel et rapport décisionnel, puis nous apprendrons à faire bon usage de Reporting Services.



Exemple de rapport réalisé avec Reporting Services

3. Office 2013

Microsoft Excel est sûrement et de loin, le premier outil décisionnel dans le monde et ce, depuis de nombreuses

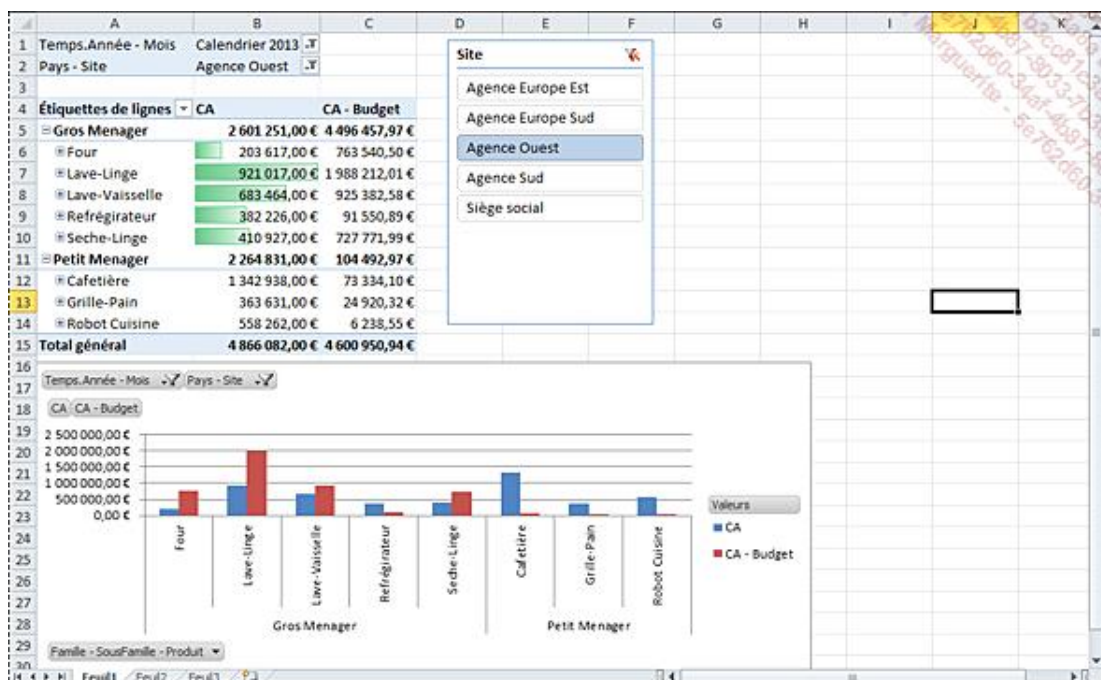
années. Il répond aux besoins d'analyse de tous les services, de toutes les organisations et sert à toutes les tâches : stockage de données, traitement de l'information et restitution.

Si Microsoft Excel seul répond assez bien à des problématiques sectorielles (propres à un seul service), il atteint toutefois ses limites lorsque :

- Il s'agit de croiser les données de référence de plusieurs applications : les fichiers Excel deviennent alors de véritables usines à gaz très difficiles à maintenir.
- Il s'agit de réduire les délais de production des tableaux de bord : il est difficile d'automatiser le traitement de données dans Excel. On peut toujours y arriver par le biais de macro, mais on augmente alors sensiblement la difficulté de maintenance des rapports.
- Les données à traiter deviennent trop importantes : Excel 2003 gère quelque 65000 lignes, Excel 2013 jusqu'à un peu plus de 1 million. Mais les systèmes produisent toujours plus de données et les demandes des décideurs ont aussi tendance à se complexifier.
- Il s'agit de sécurité : un fichier Excel diffusé par mail contient l'intégralité des données détaillées qu'il affiche et ses données sont potentiellement modifiables.
- Il s'agit d'automatiser la diffusion des rapports : les fichiers sont de plus en plus volumineux. La limite admise par le serveur de messagerie est parfois atteinte.

Pour toutes ces raisons et bien d'autres encore, Excel a besoin de s'adosser à un système décisionnel. Les utilisateurs continuent toutefois de plébisciter l'utilisation d'Excel dans leur quotidien. Microsoft a compris l'avantage substantiel qu'il peut tirer de cette position privilégiée.

Au sein de la solution BI Microsoft, Excel est l'outil d'analyse des utilisateurs finaux. Toutefois, Excel n'est utilisé que pour accéder, manipuler et naviguer dans les données d'Analysis Services. Les données ne sont plus contenues directement dans le fichier, mais sur un serveur. Et ces données ne sont plus traitées directement par les utilisateurs fonctionnels mais par le service informatique par le biais de l'ETL. Excel conserve toutefois toutes ses capacités de représentation graphique, de mise en page et de personnalisation à l'aide de formules.



Le tableau croisé dynamique d'Excel

Depuis la version 2010, Excel dispose d'un nouvel outil : le PowerPivot. Un outil puissant mais qu'il faut bien

comprendre et repositionner dans la solution Microsoft BI.

Nous étudierons en détail, dans le chapitre Restituer les données décisionnelles - Excel, l'utilisation d'Excel adossée à la solution décisionnelle Microsoft, mais aussi l'utilisation du PowerPivot. Vous apprendrez ainsi à mieux comprendre et à cerner l'utilisation de ces outils.

4. SharePoint 2013

SharePoint 2013 est une plate-forme de services de portail. SharePoint est notamment utilisé pour réaliser des portails Intranet/Extranet et des plates-formes d'espaces collaboratifs et documentaires.

SharePoint est une solution très vaste regroupant de très nombreux services, tels que :

- Le moteur de recherche de l'entreprise.
- La gestion documentaire.
- La gestion des processus métier par le biais des flux de travail (workflows).
- La gestion de contenus (*Content Management Services* ou CMS).
- L'affichage des données applicatives.

SharePoint, comme son nom l'indique, se veut être le point de convergence de tous les contenus de l'entreprise. Les informations décisionnelles font bien évidemment partie de ces contenus.

Comme nous l'avons vu plus haut dans ce chapitre, le décisionnel a pour objectif de mettre à disposition des décideurs, tous les éléments nécessaires à la prise de décision. Les contenus nécessaires à la prise de décision sont très vastes et ne se limitent pas uniquement aux informations quantitatives mises à disposition par l'entrepôt de données. L'intranet, c'est-à-dire SharePoint dans la solution Microsoft, est le relais idéal du système décisionnel pour toucher les décideurs et concevoir des espaces de décision complets, contenant :

- Rapports et analyses chiffrées.
- Commentaires et analyses des analystes métier.
- Remarques des confrères.
- Bibliothèques de documents Word et Excel.
- Lien direct pour contacter un collaborateur et engager une action.
- Lien direct vers le moteur de recherche pour une ouverture sur des informations internes ou externes à l'organisation, etc.

SharePoint est véritablement l'outil idéal pour s'adresser aux décideurs de l'entreprise.

Pour aller plus loin dans les données de l'entrepôt de données et réaliser de véritables tableaux de bord, SharePoint dispose de deux solutions :

- Excel Services
- PerformancePoint Services



Ces services nécessitent la version et la licence Enterprise de SharePoint.

Ces solutions et leur mise en place nécessitent une attention et une perception qui dépassent l'objectif de cet

ouvrage. Bien qu'essentielles, elles ne seront donc pas abordées.

Vous pourrez toutefois mettre en œuvre les solutions décisionnelles de SharePoint avec l'ouvrage Business Intelligence avec SharePoint Server - Créez un portail décisionnel et pilotez la performance. L'installation, la configuration et l'utilisation de chacune de ces solutions y étant décrites pas à pas.

a. Excel Services

Le service Excel est un serveur offrant la possibilité de transformer tout ou partie du contenu d'une feuille Excel en une page web au format HTML. La grande force des services Excel est de rendre dynamique ce contenu. Cela signifie que l'utilisateur qui publie sur SharePoint un tableau croisé dynamique basé sur Analysis Services verra le contenu de celui-ci, sur son portail SharePoint, actualisé pour afficher dynamiquement les toutes dernières données. Le tableau croisé dynamique publié conserve aussi ses capacités de navigation et de filtre comme le tableau croisé dynamique du fichier Excel originel.

Avec les services Excel, les contenus et analyses d'Excel conçus par les utilisateurs finaux deviennent alors très facilement diffusables et partageables à l'ensemble de l'organisation.

Le fonctionnement est assez simple : l'utilisateur conçoit ses analyses sous Excel. Puis, il publie le fichier sur une liste SharePoint (un répertoire virtuel en quelque sorte). Le contenu de ce fichier publié est alors calculé par le serveur Excel pour être restitué à l'aide d'une webpart SharePoint.

Dans le chapitre Restituer les données décisionnelles - Excel nous évoquerons un peu le sujet sans toutefois entrer dans les détails spécifiquement propres à SharePoint.

Étiquettes de lignes	CA	% Marge
Allemagne	4 726 528,00 €	29,08%
Gros Menager	4 458 410,00 €	28,89%
Petit Menager	268 118,00 €	32,23%
Espagne	4 743 335,00 €	28,95%
Gros Menager	4 552 849,00 €	28,83%
Petit Menager	190 486,00 €	31,85%
France	13 437 214,33 €	29,10%
Gros Menager	12 956 427,61 €	28,98%
Petit Menager	480 786,72 €	32,31%
Total général	22 907 077,33 €	29,06%

Tableau croisé dynamique Excel Services dans SharePoint 2010

b. PerformancePoint Services

Le dernier service dont dispose SharePoint 2013 est PerformancePoint. Ce dernier service, destiné à des

utilisateurs métier (type contrôleurs de gestion), est un outil destiné à élaborer et à gérer la performance de l'entreprise. Ce service s'inscrit totalement dans une approche de gestion globale de la performance de l'entreprise. On retrouve cette démarche dans la littérature sous les acronymes BPM (*Business Performance Management*), EPM (*Enterprise Performance Management*) ou CPM (*Corporate Performance Management*).

Le service PerformancePoint permet à la solution BI de Microsoft de répondre à la dernière attente de l'informatique décisionnelle : gérer les performances de l'organisation et de ses politiques.

Le service PerformancePoint est une des toutes meilleures solutions sur ce sujet et il est aussi le seul outil de la suite Microsoft à s'ouvrir sur cette démarche.

Le management de la performance a pour objectif l'alignement et la mise en cohérence des objectifs des décideurs tactiques et opérationnels sur les objectifs stratégiques de l'organisation. Cette démarche favorise l'émergence de tableaux de bord pensés et construits autour d'indicateurs clés (ou KPI pour *Key Performance Indicator*).

Cette démarche est de plus en plus présente dans les organisations tant privées que publiques (la LOLF menée par le gouvernement français en est un bon exemple). De plus en plus, vous entendrez sûrement parler de tableaux de bord, d'indicateurs, d'objectifs... Tous ces termes relèvent d'une démarche de management de la performance.

Le service PerformancePoint est alors un complément pour accompagner cette démarche, pour répondre aux exigences des décideurs de votre organisation et pour maîtriser la communication officielle sur la performance de l'organisation.

Le service PerformancePoint offre aux utilisateurs métier non-informaticiens la possibilité de :

- Créer et gérer des espaces de décision complets.
- Créer et gérer des tableaux de bord dynamiques constitués d'indicateurs clés.
- Créer et gérer des indicateurs clés, modifier les seuils d'atteinte ainsi que les visuels de météo (feu vert en cas d'atteinte d'un objectif par exemple, flèche rouge vers le bas lors de tendance à la baisse, etc.).
- Créer des analyses de données plus dynamiques que ne le permettent les services d'Excel. C'est en ce sens un bon complément au service Excel.
- Référencer et gérer les ressources officielles complémentaires aux tableaux de bord : tableau croisé dynamique Excel Services, graphique d'analyse en mode web, commentaires, ressources documentaires...

La finalité de ce service est de délivrer aux décideurs des espaces de décision en mode web, complets, riches et dynamiques, comme le montrent les copies d'écran suivantes :

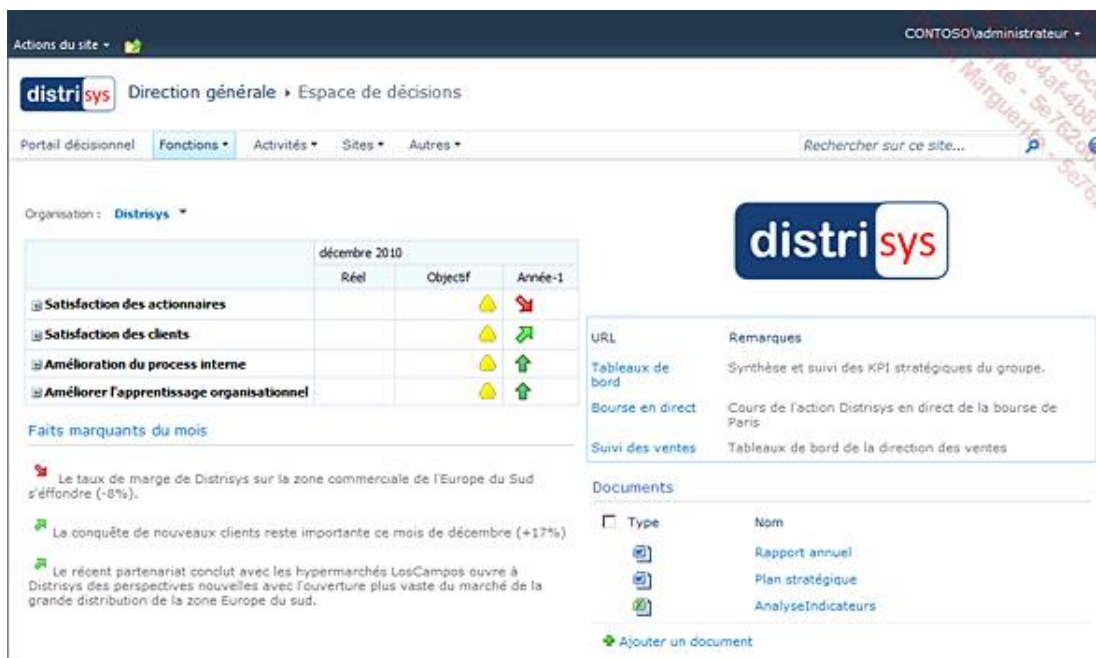
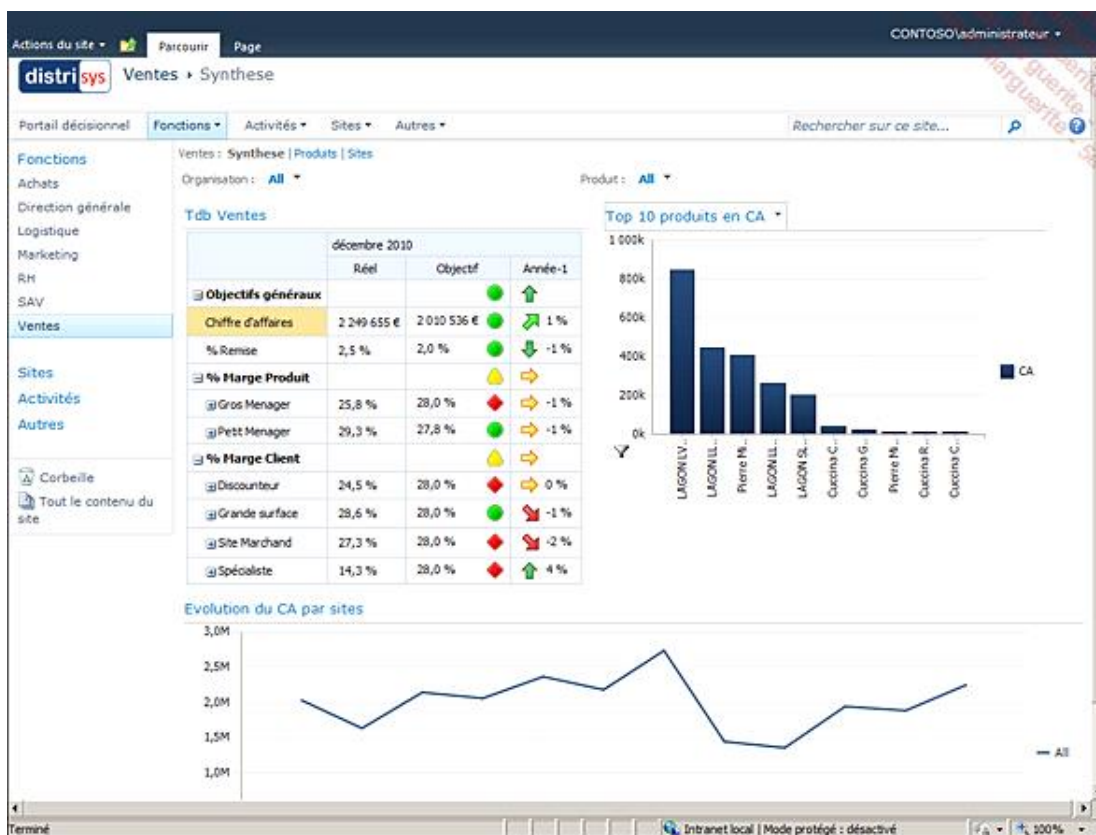
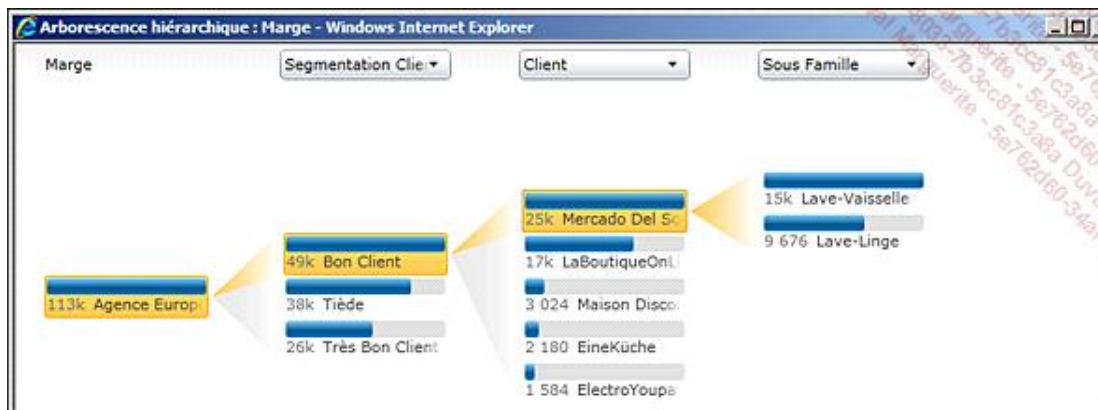


Tableau de bord intégré à l'espace portail



Espace de décision avec tableau de bord et analyse



Représentation graphique en arborescence hiérarchique

Toutes ces copies écran proviennent du portail réalisé avec le livre Business Intelligence avec SharePoint Server - Créez un portail décisionnel et pilotez la performance.