

8 janvier 2014

Mineure SOA – Cours 8

Olivier BESNARD

Consultant sénior

Practice Architecture des Systèmes d'Information



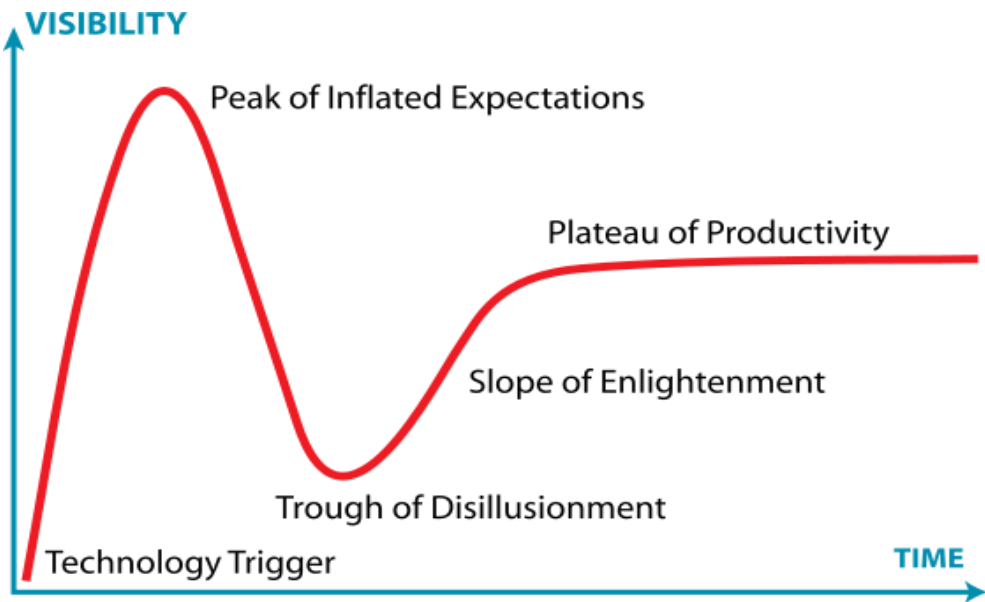
Agenda

- ▶ 1. Hype Cycle
- 2. BigData
- 3. Entreprise 2.0
- 4. BYOD
- 5. Business Intelligence

Introduction

Gartner Hype Cycle

Le Hype Cycle est un outil graphique développé et utilisé par la firme Gartner pour représenter la maturité, l'adoption et l'application sociale de technologies spécifiques



Phase	Description
Technology Trigger	A potential technology breakthrough kicks things off. Early proof-of-concept stories and media interest trigger significant publicity. Often no usable products exist and commercial viability is unproven.
Peak of Inflated Expectations	Early publicity produces a number of success stories-often accompanied by scores of failures. Some companies take action; many do not.
Trough of Disillusionment	Interest wanes as experiments and implementations fail to deliver. Producers of the technology shake out or fail. Investments continue only if the surviving providers improve their products to the satisfaction of early adopters.
Slope of Enlightenment	More instances of how the technology can benefit the enterprise start to crystallize and become more widely understood. Second- and third-generation products appear from technology providers. More enterprises fund pilots; conservative companies remain cautious.
Plateau of Productivity	Mainstream adoption starts to take off. Criteria for assessing provider viability are more clearly defined. The technology's broad market applicability and relevance are clearly paying off.

Introduction

Atelier Post it Hype Cycle

Placer les technologies sur la courbe

Brain-Computer Interface

Cloud Computing

Prescriptive analytics

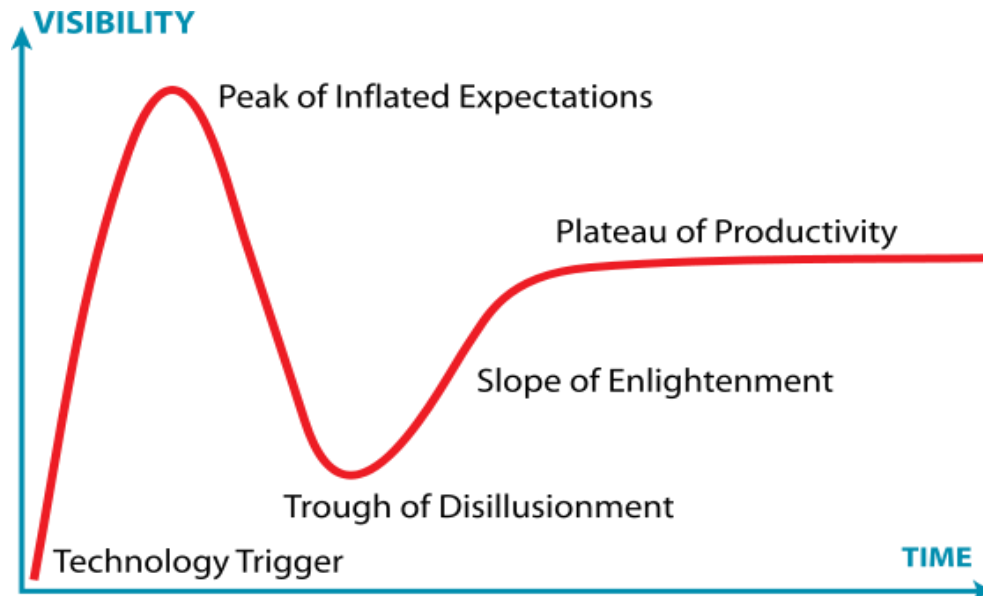
Activity Streams

Internet of Things

Predictive Analytics

Big Data

Complex Event Processing



- 3 post-it
- 1 technologie par post-it
- Indiquer 1 phase

- 5 minutes 
- *[Option] Indiquer quand le plateau sera atteint*

Plateau will be reached in:

- less than 2 years
- 2 to 5 years
- 5 to 10 years
- △ more than 10 years
- ⊗ obsolete before plateau

Agenda

1. Hype Cycle

► 2. **BigData**

2. 1 Enjeux

2. 2 BigData : c'est pas ça

2. 3 Glossaire

3. Entreprise 2.0

4. BYOD

5. Business Intelligence

Qu'est-ce que Big data?

- ▶ Ensemble de données à la volumétrie telle que leur gestion exige des moyens (organisationnels et techniques) nouveaux
- ▶ Par extension, le terme désigne ces **moyens humains ou technologiques extraordinaires**
- ▶ Tout aussi importantes, des caractéristiques complémentaires ajoutent à cette complexité
 - ▶ **En croissance rapide et soutenue**
 - ▶ **De sources et formats disparates**
 - ▶ Non structurées (textes, images, vidéo)
 - ▶ Incomplètes
 - ▶ Distribuées mondialement

Exemples de Big data

- ▶ Données transactionnelles (commandes, paiements etc.)
- ▶ Logs
- ▶ Clickstreams
- ▶ Evénements géolocalisés
- ▶ Contenus des réseaux sociaux
- ▶ Données comportementales
- ▶ Relevés de compteurs intelligents
- ▶ Appels au service client

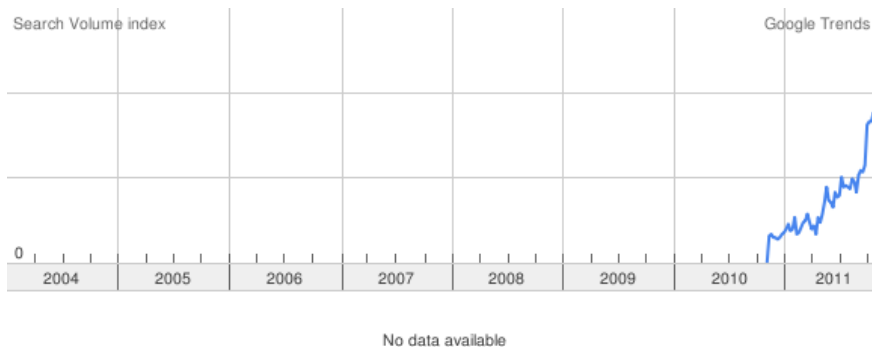
■ Contexte

- ▶ La **production de données augmente drastiquement**, par l'action des individus tout autant que celles des organisations
 - ▶ En 2010, l'humanité a déversé 800 milliards de Gygabytes, soit, comme l'a dit un jour Eric Schmidt, plus que la totalité de ce que l'humanité avait écrit, imprimé, gravé, filmé ou enregistré de sa naissance jusqu'en 2003
 - ▶ Les moteurs des Boeing produisent 10 téraoctets de données toutes les 30 minutes de vol. Ils n'en produisaient pas il y a 20 ans
- ▶ La mesure des volumes de données commence à se faire en **pétaoctets (10¹⁵ octets)**
- ▶ Les perspectives de traitement ont déjà séduit, notamment la recherche scientifique, la finance, l'indexation web etc.
- ▶ Avec l'émergence de technologies plus accessibles, beaucoup d'entreprises (même petites) souhaitent évaluer **comment tirer un avantage concurrentiel** (monétiser) de leurs gisements de données ou de ceux publics (web, open data)

Buzz ou Big Bang?

Buzz

- ▶ L'expression est **en vogue depuis fin 2010**, même si la question des gros volumes n'est pas nouvelle (moteur de recherche, grilles informatiques)



Graphe Google Trends pour "Big data"

- ▶ Le **sujet dépasse** les directions métiers et informatiques qui s'interrogent
 - ▶ Quelles sont ses opportunités business?
 - ▶ Comment s'en emparer opérationnellement?
- ▶ Son usage est **alimenté par des éditeurs et intégrateurs** qui positionnent leur offre de gestion des données dans le paysage Big data et contribuent à étendre son périmètre

Big Bang

- ▶ Certains analystes évoquent un Big Bang
- ▶ Le terme **Big Data désigne aussi pour eux les transformations majeures, nécessaires pour « passer à l'échelle »**
 - ▶ De **stratégie** d'entreprise pour capter, contrôler et donner du sens aux données
 - ▶ **Organisationnelles** pour revoir les pratiques (journalisme, relations publiques)
 - ▶ **Technologiques** avec la refonte nécessaire des applications et infrastructures IT
 - ▶ **Légales et éthiques** pour éviter les abus (façon Big Brother)



"Your recent Amazon purchases, Tweet score and location history makes you 23.5% welcome here."

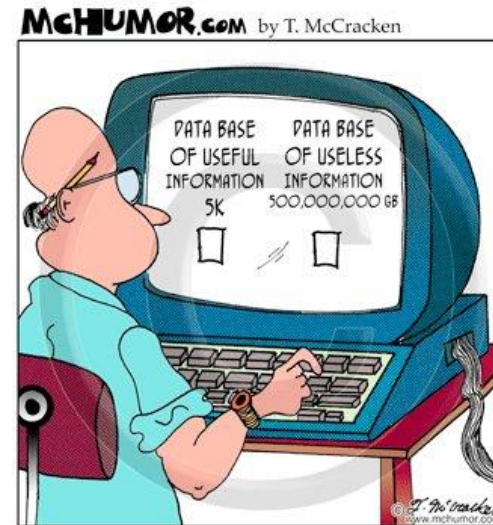
Quels sont les enjeux?

■ Métiers

- ▶ **Extraire des données brutes**, de l'entreprise ou publiques, **les informations utiles** capables de supporter ses activités, de gagner en productivité et d'innover

■ DSI

- ▶ **Offrir des moyens efficaces et évolutifs** (« **scalables** ») pour toutes les activités
 - ▶ **Capture** (sourcing élargi, enrichissement des données de méta-données)
 - ▶ **Stockage**
 - ▶ **Analyse** (algorithmique parallèle, apprentissage automatique, analyse en langage naturel, statistiques etc.) **et ré-analyse** sur de nouveaux axes ou par de nouvelles corrélations
 - ▶ **Restitution**
- ▶ **Refondre** les architectures réseaux, les traitements et logiques des gestions des données
- ▶ **Faire converger les travaux à tous les niveaux**



Exemples d'activités métiers concernées

- ▶ Pilotage opérationnel
- ▶ Aide à la décision à partir de modèles analytiques
- ▶ Analyse de processus, de tendance, d'opinion
- ▶ Amélioration de la performance opérationnelle par expérimentation
- ▶ Etude des affinités et comportements d'achats des clients pour mieux adapter produits et services
- ▶ Gestion de la relation client
- ▶ Développement de services innovants à partir des données d'utilisation des services existants (after-market services)

Agenda

1. Hype Cycle

► 2. **BigData**

2. 1 Enjeux

2. 2 BigData : c'est pas ça

2. 3 Glossaire

3. Entreprise 2.0

4. BYOD

5. Business Intelligence

Domaines technologiques proches

- ▶ **Les solutions** dans ces catégories technologiques **n'adressent pas spécifiquement le Big data** mais apportent une solution partielle ou globale à ses problématiques
- ▶ **Les opportunités business** du Big data conduisent **les offreurs à positionner** leurs produits et services **dans le paysage du Big data**

■ Cloud computing : externaliser tout ou partie des activités

- ▶ Une solution **rapide, économique** et, donc, peu risquée (en terme de ROI) pour s'attaquer aux problématiques techniques du Big data est d'externaliser l'infrastructure (au sens large) et l'expertise associée à son exploitation: idéal pour les petites entreprises et les expérimentations
- ▶ Les offreurs PaaS et SaaS se positionnent sur les problématiques Big Data en développant **l'Analytics as a Service**
- ▶ La **sécurité des données** et **l'intégration des sources internes avec l'infrastructure** externe exige une attention particulière lors des études préalables à l'adoption d'un service du cloud

■ Grid computing : accélérer l'analyse

- ▶ Depuis plus d'une dizaine d'années, le grid computing est la solution pour supporter le calcul intensif. Le principe est de **paralléliser les traitements massifs ou/et longs** en les envoyant **sur une « grille » de machines**
- ▶ Le Big data rend obsolète les logiques simplistes de distribution des données via le **réseau** qui **devient goulot d'étranglement**. (notamment dans les grilles basées sur le bénévolat des internautes). Il faut repenser la logique d'accès des nœuds aux données (Map/Reduce) et/ou intégrer des briques supplémentaires (systèmes de fichiers distribués, data grid, caches distribués)
- ▶ Le grid computing « pur » demeure **pertinent pour des traitements essentiellement consommateurs de CPU** pour lesquels le temps de traitement écrase celui de transfert des données

■ CEP : capter et analyser des événements en temps réel ET réagir

- ▶ Le CEP (Complex Event Processing) détecte à partir d'un grand nombre d'évènements, **des évènements complexes** et y apporte des réponses adaptées dans des contraintes de durée de type « temps réel »
- ▶ Dans le cadre du Big Data, le CEP se distingue par sa capacité à offrir des **alertes, des indicateurs et des réponses automatisées en temps réel** tout en étant disposé à agir sur des données variées et volumineuses
- ▶ Le CEP travaille par essence sur des flux de données (MOM, ESB en sont le coeur), une fois exploitée, l'information événementielle n'a pas forcément vocation à être stockée

Agenda

1. Hype Cycle

► 2. **BigData**

2. 1 Enjeux

2. 2 BigData : c'est pas ça

2. 3 Glossaire

3. Entreprise 2.0

4. BYOD

5. Business Intelligence

■ Acidité

- Ensemble de propriétés qui garantissent que les transactions de bases de données sont exécutées de manière sûre
 - **A**tomacité garantit qu'une transaction incomplète ne peut exister
 - **C**onsistance garantit que l'état de la base de données, à travers les transactions, respecte ses règles de fonctionnement
 - **I**solation garantit qu'il n'y a pas d'interférence entre transactions
 - **D**urabilité garantie que le résultat d'une transaction s'inscrit dans la durée
- « La majorité des SGBD hiérarchiques comme relationnels du marché permettent de réaliser des transactions atomiques, cohérentes, isolées et durables. **Les systèmes NoSQL n'y prétendent pas encore** » (source wikipédia). Précisément, ils **en font fi volontairement pour gagner en souplesse et en performance**

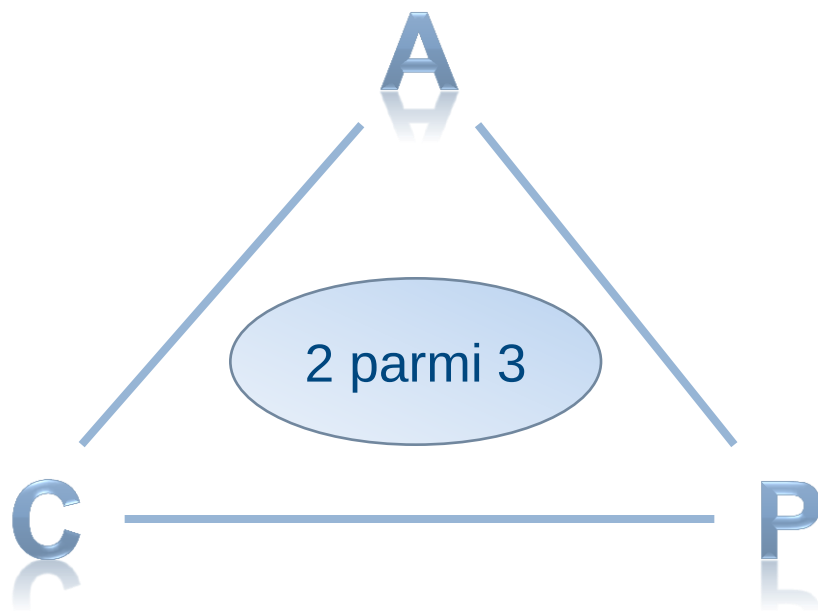
■ Lac de données (Data Lake)

- **Zone de stockage** des architectures Big data où les données sont déversées de manière brute, **sans classement, ni organisation avancée présumant des analyses** qu'elles peuvent subir (pour permettre leur ré-analyse)

Glossaire

Théorème CAP de Brewer

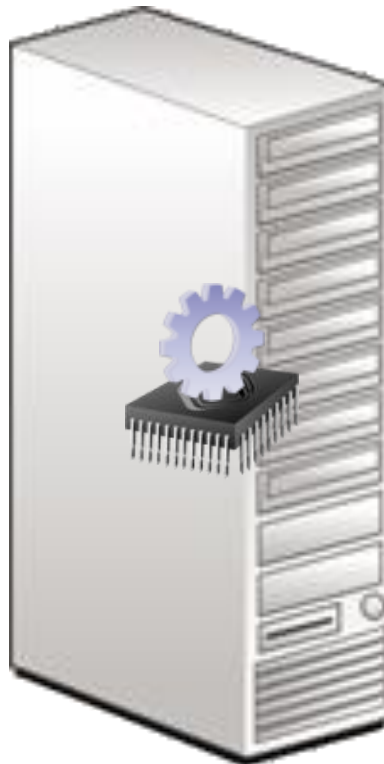
- Au sein d'un système distribué, il est impossible de garantir, en même temps, les trois contraintes suivantes :
 - Cohérence
 - Disponibilité (Availability)
 - Résistance au partitionnement (Partition Tolerance)



■ Exemples

- CA
 - SGBD Traditionnel en Cluster
 - AsterData
 - Greenplum
- CP
 - Memcached
 - Redis ?
 - Hadoop/HBase
- AP
 - Cassandra
 - CouchDB

- Scalabilité interne (Scale in)
 - Capacité d'un logiciel à pouvoir tirer parti d'UNE ressource de plus en plus grosse



Glossaire

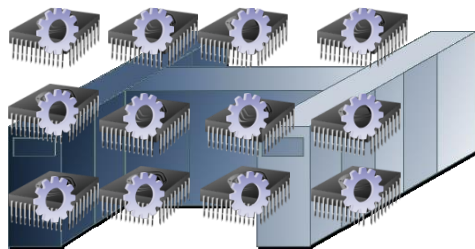
Scalabilité (2/2)

■ Scalabilité verticale (Scale up)

- Ce dit d'une solution capable d'utiliser de plus en plus de ressources au sein d'un ordinateur

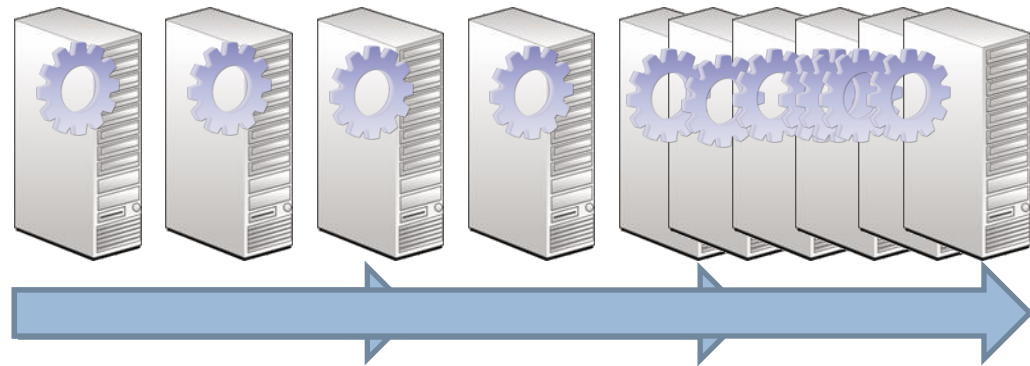


Attention : le « logiciel » doit pouvoir suivre la scalabilité



■ Scalabilité horizontale (Scale out)

- Ce dit d'une solution capable d'utiliser des ressources de plus en plus d'ordinateurs

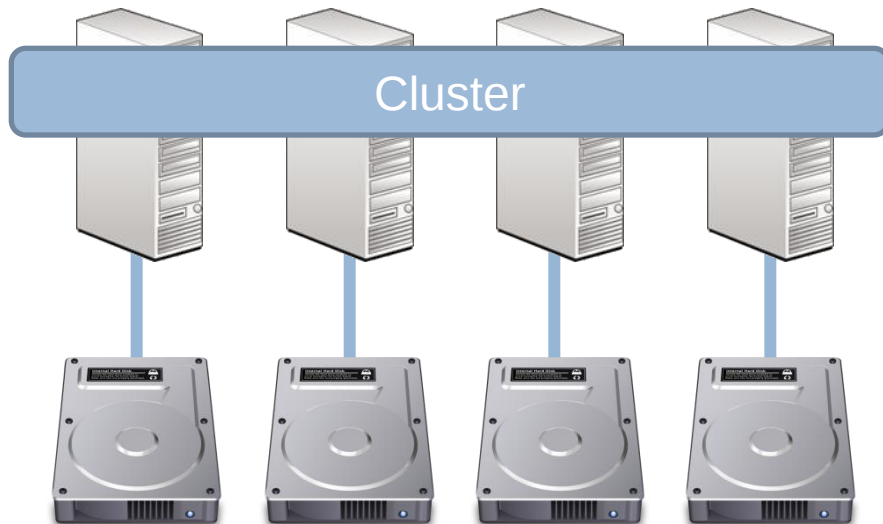


Glossaire : Cluster

Shared-nothing vs Shared-everything

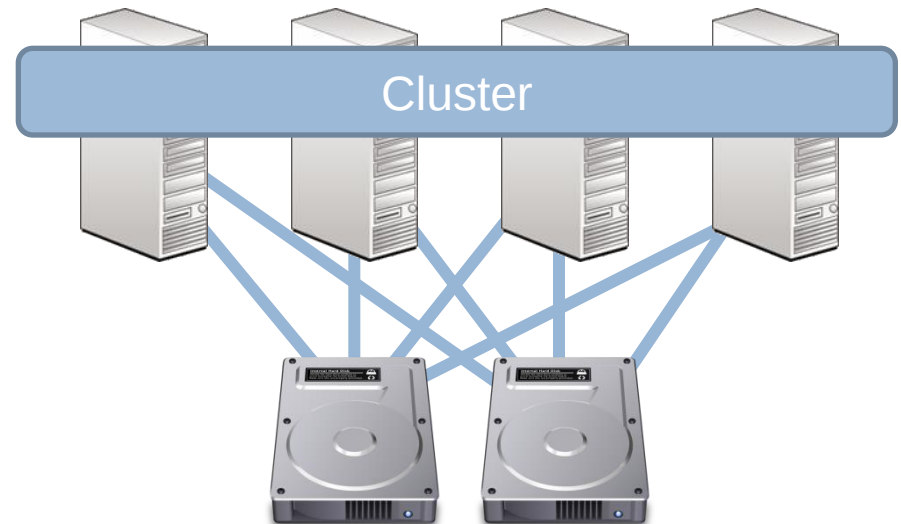
■ Shared-nothing

- ▶ Chaque nœud ne partage pas de ressource de manière commune avec les autres nœuds
 - Dans la réalisation de sa tâche



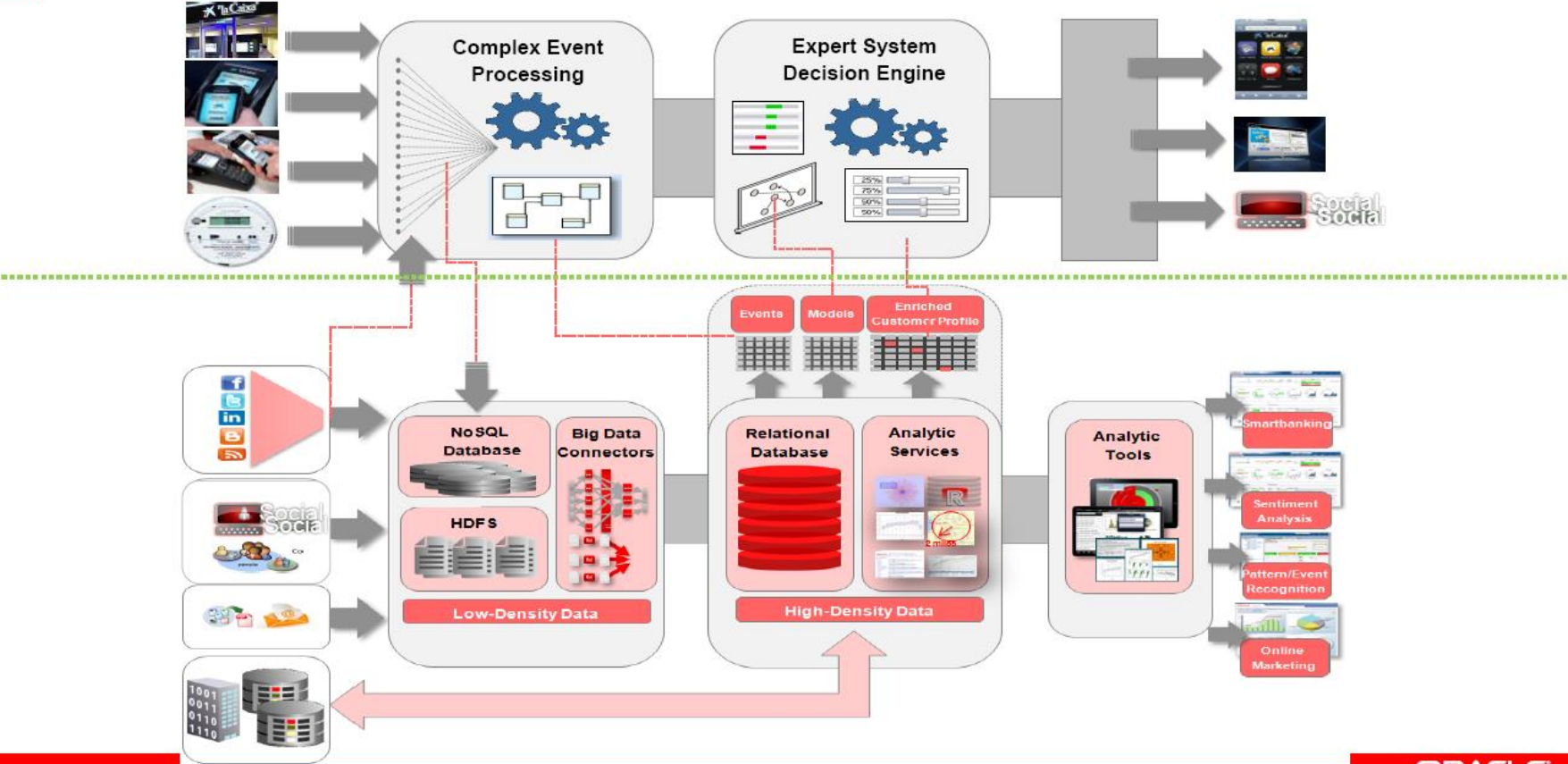
■ Shared-everything

- ▶ Chaque nœud peut accéder à l'ensemble des ressources du cluster en même temps



Use case Oracle

Use Case: Financial Services – Customer Loyalty



ORACLE

Agenda

1. Hype Cycle
2. BigData
- ▶ 3. **Entreprise 2.0**
4. BYOD
5. Business Intelligence

Objectifs



- ☐ Enjeux des entreprises face au 2.0
- ☐ Concepts
- ☐ Outils
- ☐ Démarche
- ☐ Retours d'expérience

Une vue historique : L'évolution des outils intranet et de communication



D'autres facteurs d'influence importants

- La démographie : Les attentes de la génération Y



**3000 collaborateurs
nouveaux cette année**

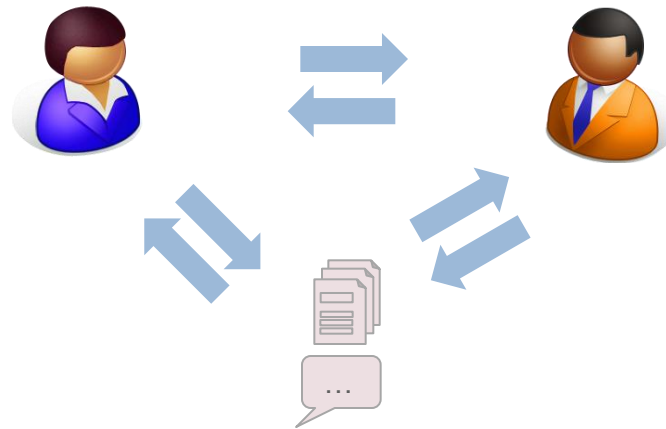
- Des évolutions sur les organisations et les modes de fonctionnement
 - « Globalisation »
 - Entreprise étendue
 - Accélération de la mise sur le marché d'innovations
 - Recherche de performance opérationnelle

Les entreprises face au travail collaboratif 2.0

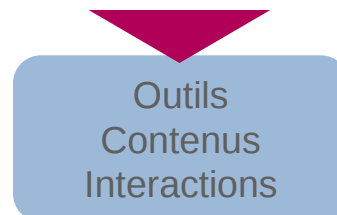
- A quoi ça sert ?
 - Veille et innovation
 - Mise en œuvre de projets / de plans d'action
 - Capitalisation de pratiques et d'expertise
 - Mise en relation d'individus

- Comment abordent-elles le sujet ?
 - **La DSI** : évolution de l'offre de services et du poste de travail
 - **Les métiers** : réponse à des enjeux stratégiques et de performance
 - **La comm'** : évolution des intranets et des outils de communication

Les concepts autour du travail collaboratif



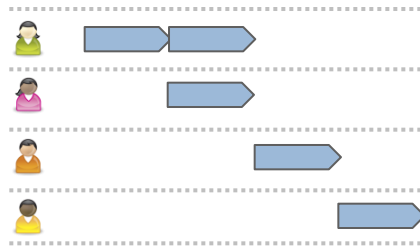
- Des individus
- Des échanges
- Des contenus informationnels
- Des règles d'accès et d'interaction



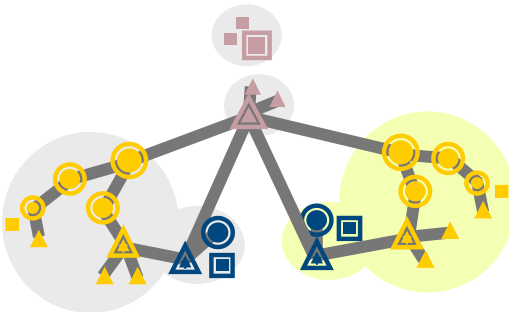
Les concepts : travail coopératif et travail collaboratif

Travail coopératif

Découpage
organisé
a priori

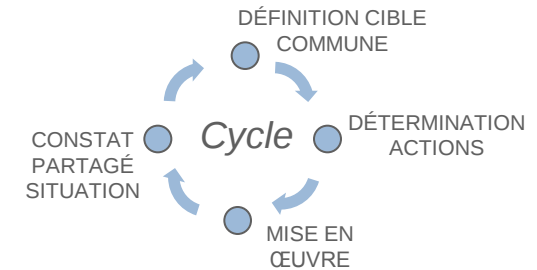


Interactions
reflétant la
structure de
l'organisation
du travail



Travail collaboratif

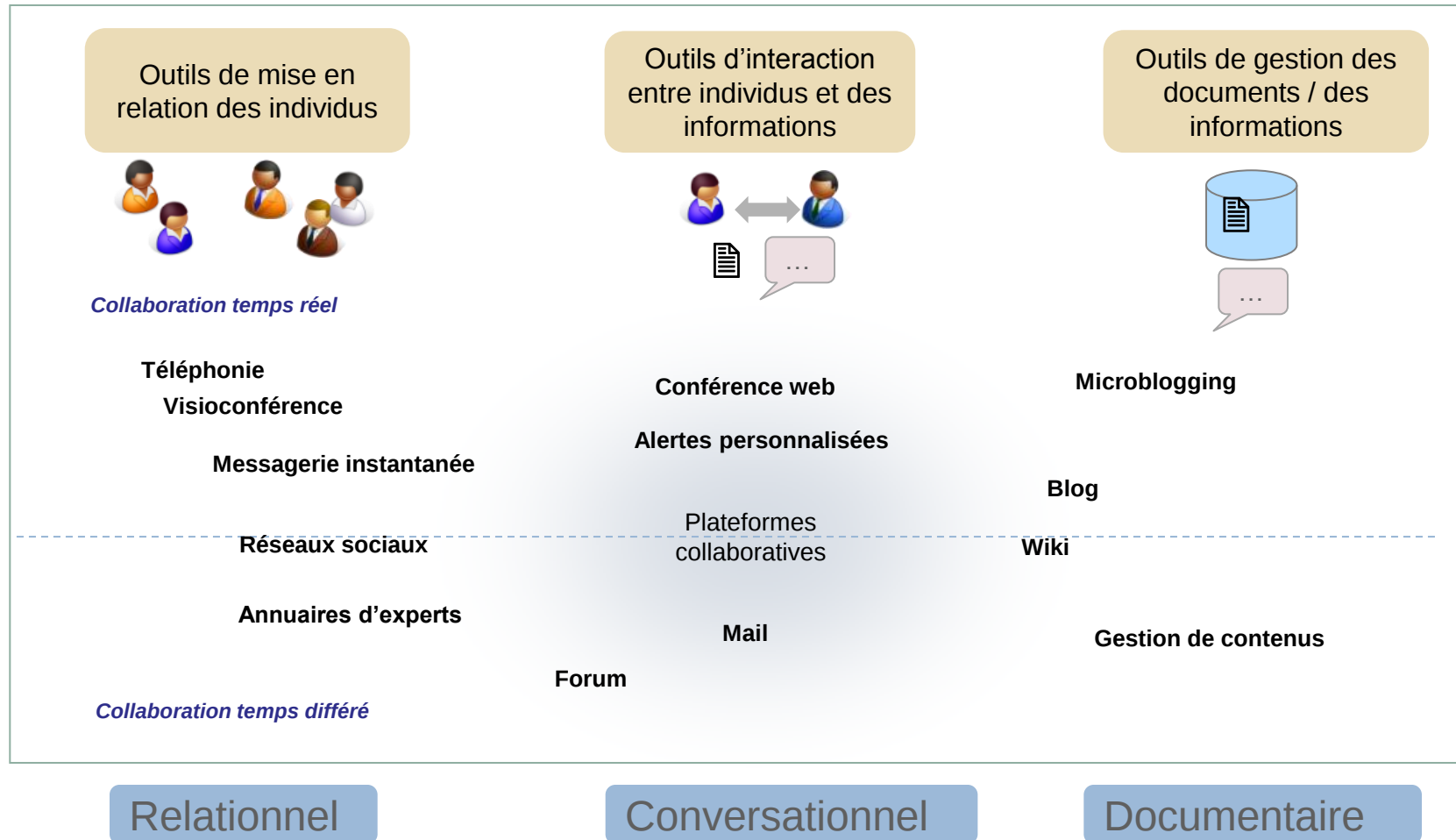
Intelligence
collective



Interactions
ad hoc :
- Connaissance
- Object de
l'interaction



Une cartographie fonctionnelle pour se repérer



Agenda

1. Hype Cycle
2. BigData
3. Entreprise 2.0
- ▶ 4. **BYOD**
5. Business Intelligence

Notre vision du BYOD

Qu'est-ce que le BYOD

Le BYOD (*Bring Your Own Device*) désigne le fait d'utiliser des terminaux personnels pour des besoins professionnels



Notre vision du BYOD

Un concept à la mode, mais qui existe depuis des années

ZDNet UK / News and Analysis / Infrastructure / Syst

BP employees get DIY IT

By Andy McCue, silicon.com, 12 October, 2006 09:25

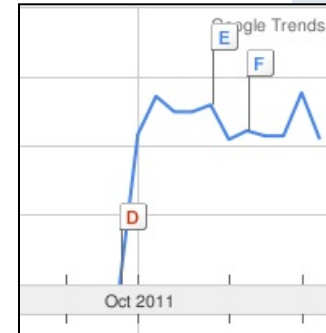


BYODKM:
Bring your own
display, keyboard,
and mouse.

2005...
BYODKM

2006...
DIY IT

2011...
BYOD



20th
Century...

BYOB

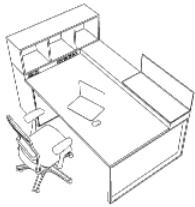
Please feel free
to bring your
own bottle



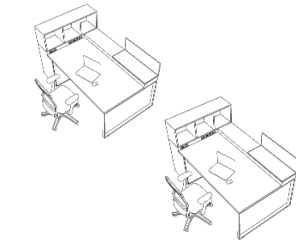
Notre vision du BYOD

Tout commence par une évolution des usages

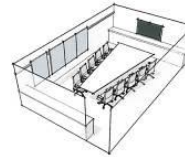
Hier...



Bureau dédié



Dans un ou plusieurs bureaux



Dans des espaces collaboratifs



A distance



Sur la route



Téléphone portable



PDA



Internet mobile



Musique



Applications personnelles



Vidéos

Ces changements profonds ont amené à une redéfinition de l'espace de travail

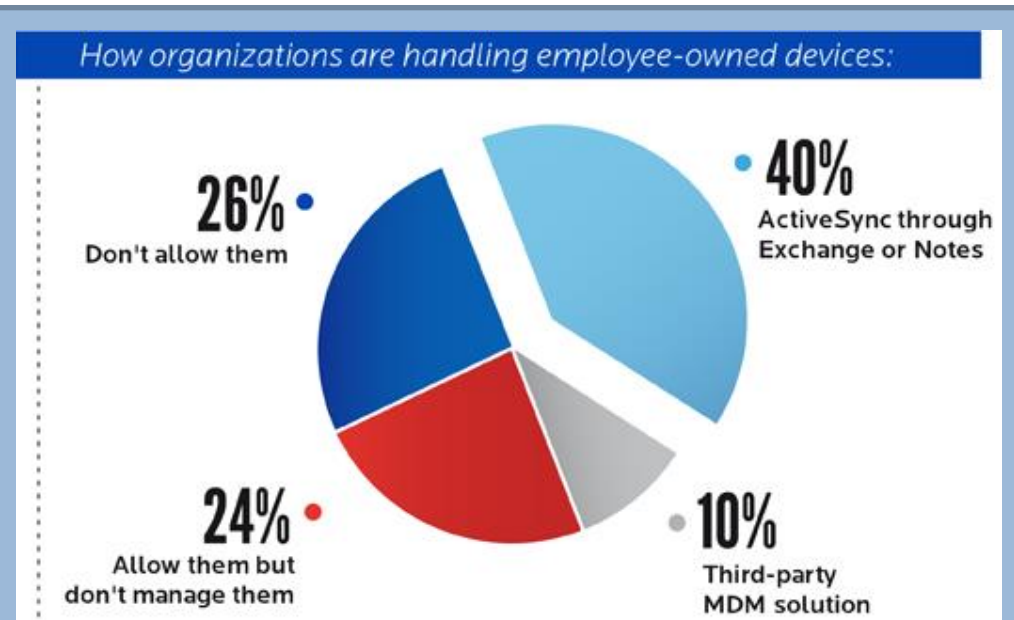
Notre vision du BYOD

... Mais plusieurs contraintes se sont tout de suite posées !

- **Les DSI** : “Le BYOD n’est pas dans mes priorités”
 - ▶ Cohérence du SI
 - ▶ Gestion et coût des licences
 - ▶ ...
- **Les RSSI** : “Comment protéger les données de l’entreprise ?”
 - ▶ Pas de contrôle du terminal
 - ▶ Pas de traçabilité
 - ▶ ...

Des approches variées chez les Grands Comptes

- Interdiction
 - Tolérance
 - Sensibilisation
- ... et des usages “sauvages” presque partout !



Sources: <http://blog.maas360.com> ; Fiberlink ; Techland Time

Notre vision du BYOD

Gains attendus



Pour les (futurs) collaborateurs

- Souplesse
- Innovation
- Motivation

Il s'agit aujourd'hui d'un élément majeur d'attraction/de motivation (pour les digital natives, mais pas uniquement)



Pour l'entreprise

- Productivité
- Réduction des coûts
- Satisfaction des collaborateurs

*De premiers retours concluants dans des entreprises du domaine IT
Toujours à confirmer cependant dans divers environnements*

Notre vision du BYOD

Les besoins (ou envies ?) qui tirent le BYOD

Priorités des utilisateurs (décroissantes)

Quelles
fonctions ?



Messagerie



Agenda



Contacts



Intranet



Documents



Applis Métier

Quels
terminaux ?



Smartphone



Tablette



Laptop

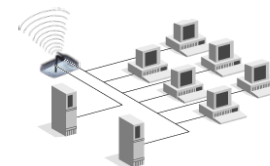


Quelle
connectivité ?



3G

Connexion
distante



Connexion directe au
réseau de l'entreprise

Niveau de risque sécurité (croissant)

Mais avant tout, l'ergonomie !

Contraintes et risques associés

Les risques doivent être évalués de bout en bout



**Risques sur les
Données**

**Risques sur le Système
d'Information de
l'Entreprise**

**Risques juridiques et
réglementaires**

Contraintes et risques associés

Scénarios de risque liés aux données

	Scénarios de risque	Criticité
Vol/fuite de données	... par un vol du terminal personnel	H
	... par une application malveillante	M
	... par une interception des communications	H
	... par une intrusion sur un terminal personnel	H
	... par une fuite sur l'infrastructure du fournisseur/Cloud	M
Indisponibilité des données	... due à une absence prolongée du collaborateur	L
	... due à la destruction du terminal personnel	H
	... due à une erreur d'administration sur le terminal personnel	M
Perte d'intégrité des données	... due à une modification illégitime des données sur le terminal	M

Contraintes et risques associés

Risques liés au SI de l'Entreprise

	Scénario de risque	Criticité
Intrusion sur le SI	... due to à la compromission d'un terminal	H
	... due à une compromission de la plateforme de gestion des terminaux	M
	... due à la compromission d'une plateforme externe de connexion	L
	... due à des mécanismes d'authentification trop faibles	M
Propagation virale	... due à un terminal compromis	H
Responsabilité de l'entreprise suite à un incident	... manque de traçabilité sur les flux	M
	... compromission d'un terminal personnel depuis le SI de l'Entreprise	L

Contraintes et risques associés

Risques juridiques et réglementaires



Agenda

1. Hype Cycle

2. BigData

3. Entreprise 2.0

4. BYOD

► 5. Business Intelligence

5. 1 Introduction

5. 2 Cas d'usage

5. 3 Architecture

5. 4 Glossaire

Les deux mondes : décisionnel et opérationnel

SI Opérationnel (SIO)

▼

Ce sont les tâches, quotidiennes, répétitives et atomiques qui sont effectuées par les employés de l'entreprise pour lui permettre d'avoir une activité. Les systèmes informatiques opérationnels sont faits pour assister les opérations d'une entreprise.

SI Décisionnel (SID)

▼

C'est l'ensemble des méthodes, moyens et outils informatiques utilisés pour piloter une entreprise et **aider à la décision**. Le monde décisionnel analyse, prédit, conseille, regarde de haut les données de l'entreprise pour mieux apprécier l'ensemble de l'activité.

Points à retenir

- ▼
- ▶ Les systèmes opérationnels sont au cœur des activités de l'entreprise, ils assistent la production.
 - ▶ Les systèmes décisionnels apportent une vision transverse de l'entreprise. Ils permettent de générer de la connaissance à partir des données, et donc, d'aider à faire des décisions stratégiques.

Qu'est-ce que la Business Intelligence ?



Définition

La **Business Intelligence (BI)** ou **Informatique Décisionnelle** désigne les moyens, les outils et les méthodes qui permettent de collecter les données internes et externes, et les transformer en informations analysées par les utilisateurs pour **prendre de meilleures décisions** et améliorer la performance de l'entreprise

Objectifs généraux

- Améliorer la **pertinence** et la **réactivité** du pilotage de la stratégie
- Donner une **vision cohérente** de l'activité de l'entreprise
- Aligner objectifs stratégiques et **actions du management opérationnel**
- **Modéliser le business** (centralisation et historisation des données)

BI stratégique

Promouvoir la performance de l'entreprise dans son entreprise



BI tactique

Permet d'identifier la source du problème une fois celui-ci détecté



BI opérationnelle

Facilite la prise de décision quotidienne aux niveaux inférieurs de l'organisation



Les différents types de la BI

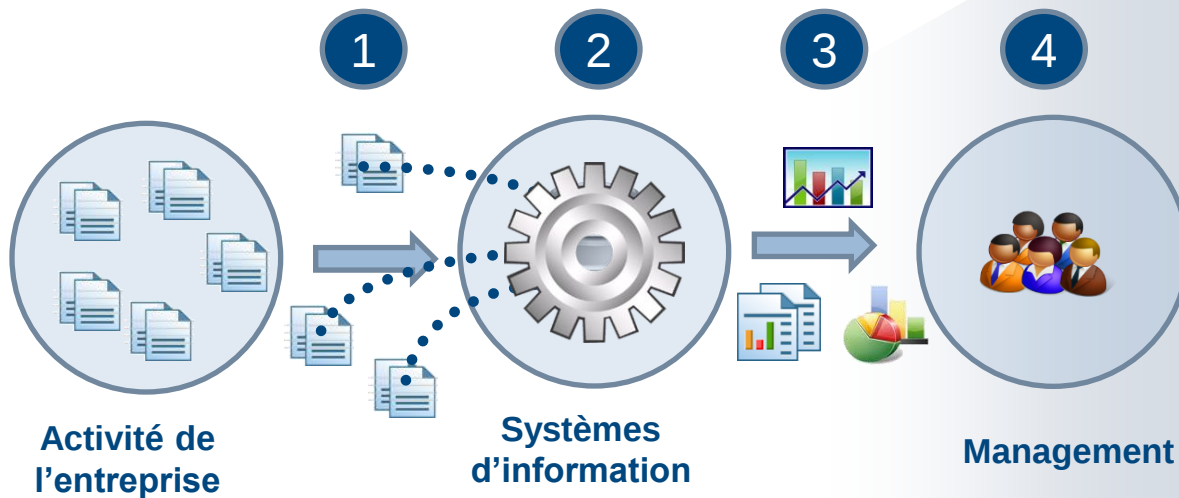
Exemple de cas d'usage de la Business Intelligence :

- Donner l'état courant de l'activité
- Analyser les comportements (client, employé, fournisseur...)
- Identifier des coûts inacceptables
- Réaliser des simulations
- Identifier des tendances
- ... et bien d'autres usages

Les 4 fonctions de la chaine décisionnelle

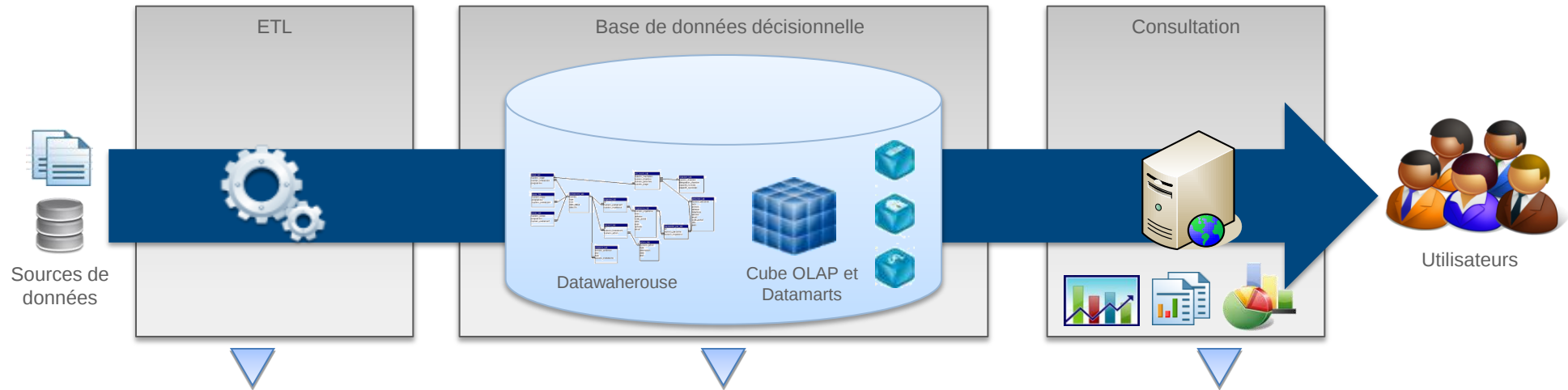
4 grandes étapes

La chaine décisionnelle



- ① Étape 1 : **Collecter** les données des systèmes de production, les nettoyer et les consolider
- ② Étape 2 : **Stocker**, centraliser les données structurées et traitées
- ③ Étape 3 : **Distribuer** ou plutôt faciliter l'accessibilité des informations selon les fonctions et les types d'utilisation
- ④ Étape 4 : **Restituer** ou comment assister du mieux possible l'utilisateur afin qu'il puisse exploiter l'information des données stockées à cet usage.

Architecture d'un Système d'Information Décisionnelle (SID)



COLLECTER / TRANSFORMER

- Outils de transfert et réception de données (Type plate-forme d'échanges)
- Outils de transformation (Ex : ETL, ELT)
- Outils de contrôle et de cohérence de la donnée

STOCKER / DISTRIBUER

- Solution de stockage logique (SGBD)
- Solution de stockage physique (Type SAN)
- Solution intégrée (Type Appliance) pour les hautes volumétrie

RESTITUER / EXPLOITER

- Outils de reporting généraliste ou spécialisé (Ex: BO, Hyperion...)
- Outils de type statistiques (Ex: SAS)
- Portail de publication de rapports
- Diffusion de données valorisées



Sur le plan pratique et technique, la Business Intelligence se compose d'une famille d'outils informatiques et de progiciels assurant le fonctionnement de la chaîne de traitement de l'information formant un SID.

Agenda

1. Hype Cycle

2. BigData

3. Entreprise 2.0

4. BYOD

► 5. Business Intelligence

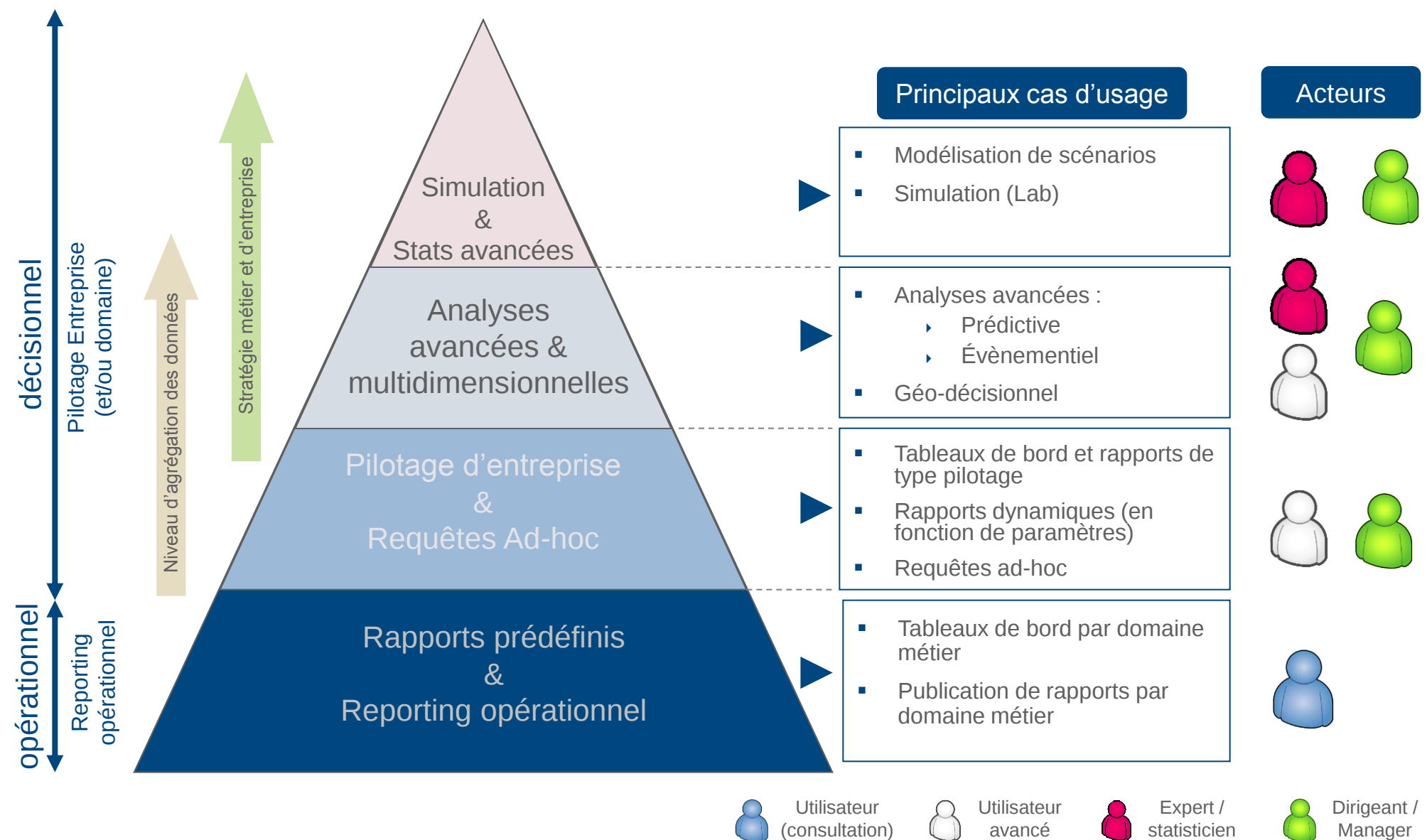
5. 1 Introduction

5. 2 Cas d'usage

5. 3 Architecture

5. 4 Glossaire

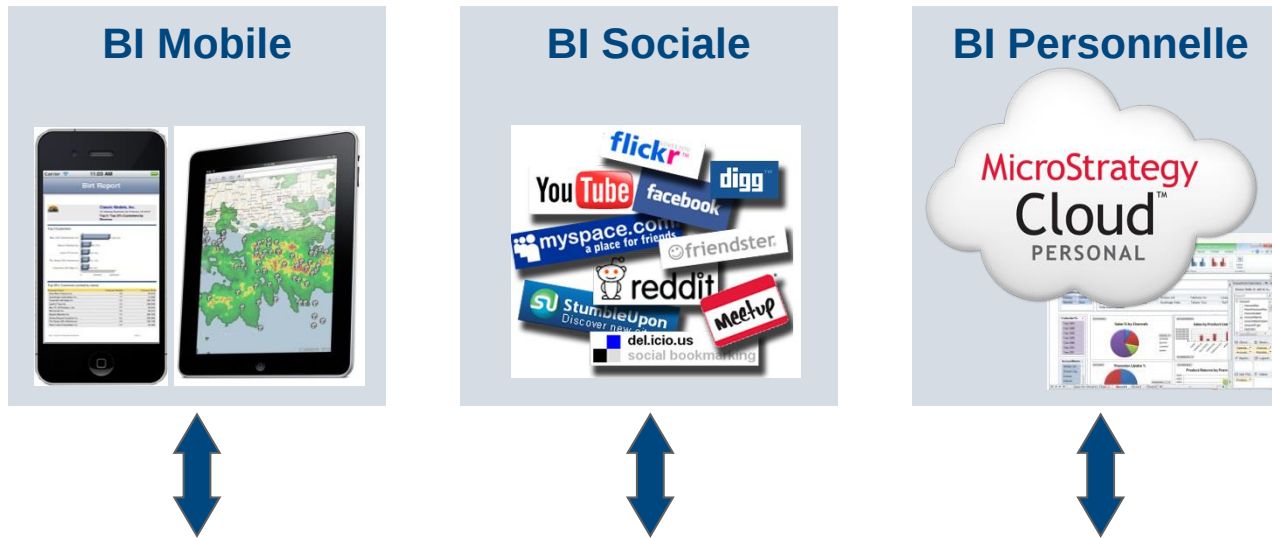
La pyramide des grands cas d'usage du décisionnel



Cas d'usage

Nouveaux usages

→ Émergence de nouveaux usages portés par les tendances technologiques

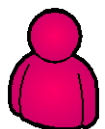


Vers un usage toujours + mobile + rapide + simple + collaboratif

Accessibilité accrue de l'information « Anywhere, Anytime » (ex. démarchage des commerciaux en situation de mobilité)

Élargissement de l'analyse décisionnelle aux réseaux sociaux (ex. complétion d'information pour une division marketing)

Consumer Intelligence (ex. Lloyds Banking Group permet à ses clients d'utiliser la BI pour leurs décisions financières)



Utilisateur
simple



Utilisateur
avancé



Dirigeant/ Manager

+



Particulier



PME

Cible

Agenda

1. Hype Cycle

2. BigData

3. Entreprise 2.0

4. BYOD

► **5. Business Intelligence**

5. 1 Introduction

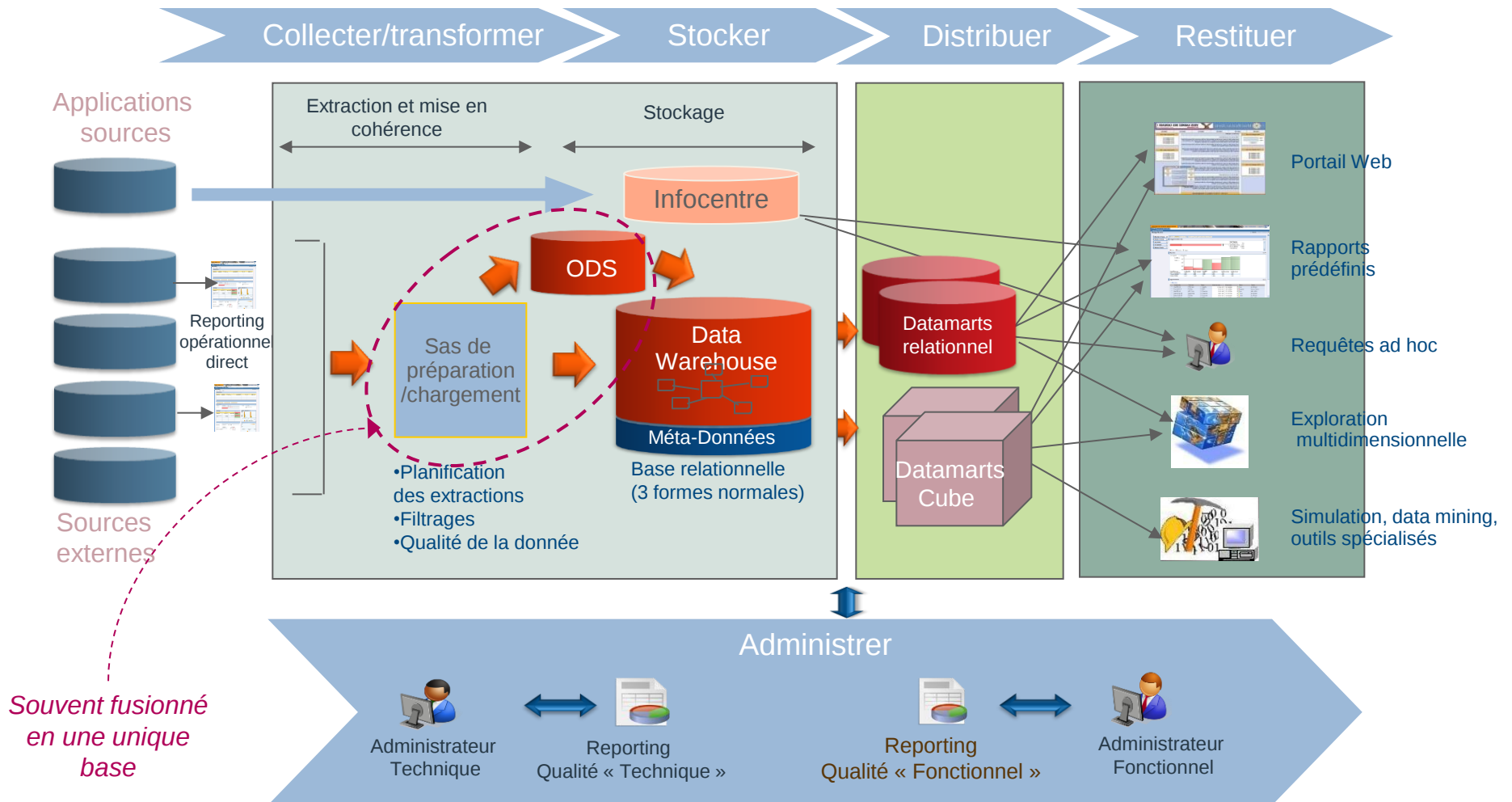
5. 2 Cas d'usage

5. 3 Architecture

5. 4 Glossaire

Architecture type d'un système décisionnel

Le modèle Hub & Spoke (Top-Down) :

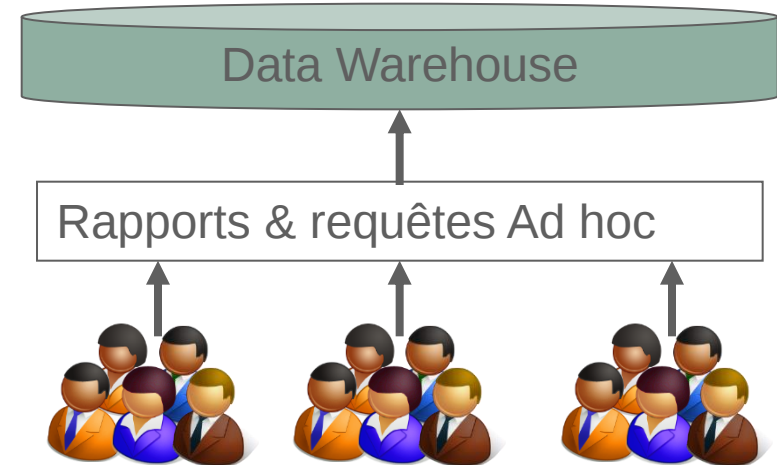


Patterns d'architecture

Les modèles d'architecture (1/2)

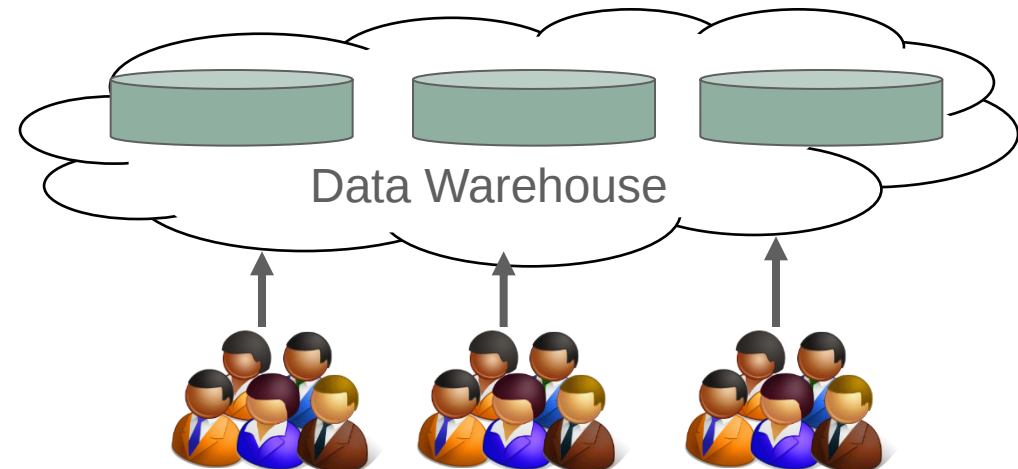
Centralisé

- Toutes les données sont dans un système central dont la responsabilité incombe à une unique entité de l'entreprise
- Pour autant, il est possible d'avoir plusieurs entrepôts par domaine métier pour respecter une isolation des types de données
- Point essentiel : l'environnement est géré comme une seule entité intégrée.
- Ce modèle est peu répandu dans les entreprises (absence de datamarts)



Distribué

- Dans ce modèle, l'entrepôt de données est distribué, que ce soit avec ou sans datamarts. Cela peut impliquer deux implémentations différentes :
 - L'entrepôt de données peut résider sur de multiples environnements. L'essentiel est que ces entrepôts soient conformes au même modèle de données, et gérées comme une seule entité.
 - L'entrepôt de données peut résider sur divers environnements, mais comme des entités distinctes et indépendantes. Elles n'ont généralement pas un modèle de données unique et sont gérées de façon indépendante.

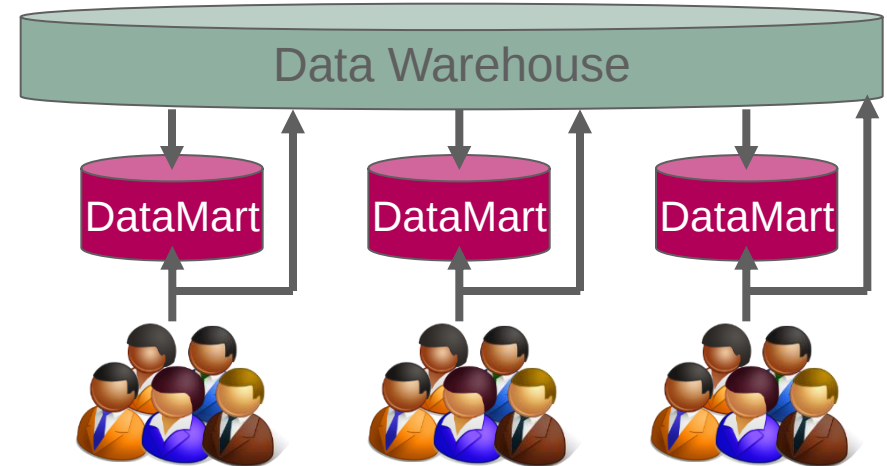


Patterns d'architecture

Les modèles d'architecture (2/2)

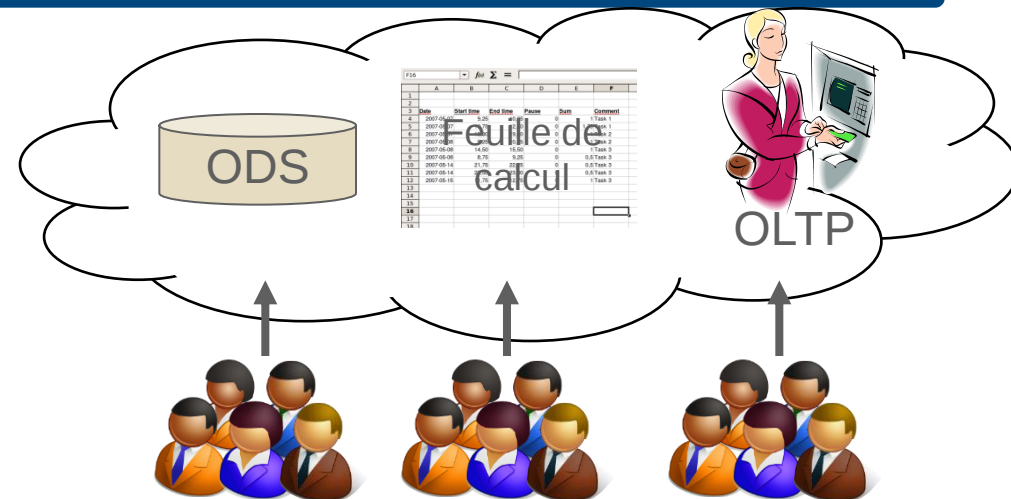
Hub & Spoke

- Ce modèle représente une combinaison des modèles centralisé et distribué.
- Il implique un entrepôt de données central, qui est la « plaque tournante », et des implémentations de datamarts distribués, qui sont les « rayons »
- Point essentiel : l'environnement est géré comme une seule entité intégrée.
- **Ce modèle est de plus en plus répandu dans les grandes entreprises**



Fédéré

- Ce modèle est basé sur les systèmes opérants. Il existe quand il n'est pas nécessaire de déplacer, intégrer, et consolider les données de l'entreprise
- Les données, même hétérogènes, de plusieurs sources de données sont accessibles directement pour l'analyse
- Ce modèle a ses limites en termes de performance et de cas d'usages (souvent réduit à du reporting opérationnel)



Les principaux acteurs du marché de la Business Intelligence opérationnelle et leurs principales offres

Les problématiques de Business Intelligence opérationnelle sont souvent, et de plus en plus fréquemment **intégrées aux solutions des différents éditeurs**. Toutefois, il y a des solutions qui ressortent du lot :

QlikView de QlikTech



- plateforme très orientée utilisateur (très intuitive)
- produit BI ayant le plus haut taux de satisfaction client (96%*)

Actuate One de Actuate



- tarification très compétitive (produit basé sur BIRT, solution open source)

BI Cognos de IBM



la plateforme la plus complète en termes de couverture des besoins décisionnels (Gartner 2012)

Business Objects BI suite de SAP

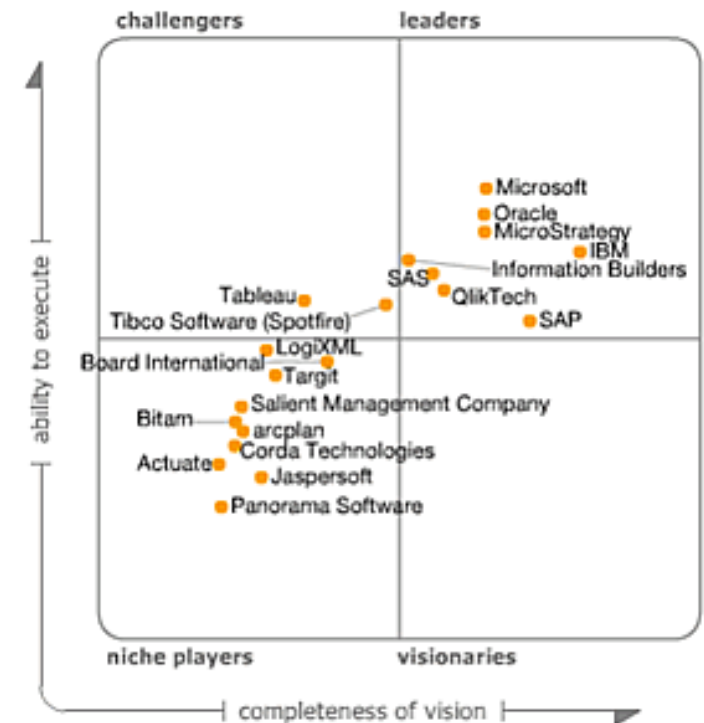


- très performant sur de grands volumes de données (tout particulièrement si intégré avec l'appliance HANA)
- positionnement historique sur le reporting de masse

Spotfire DecisionSite Platform de TIBCO



- élu meilleur outil BI opérationnelle (étude Yphise, 2009)



Source : Magic Quadrant for Business Intelligence Platforms - Gartner, Janvier 2011

Agenda

1. Hype Cycle

2. BigData

3. Entreprise 2.0

4. BYOD

► 5. Business Intelligence

5. 1 Introduction

5. 2 Cas d'usage

5. 3 Architecture

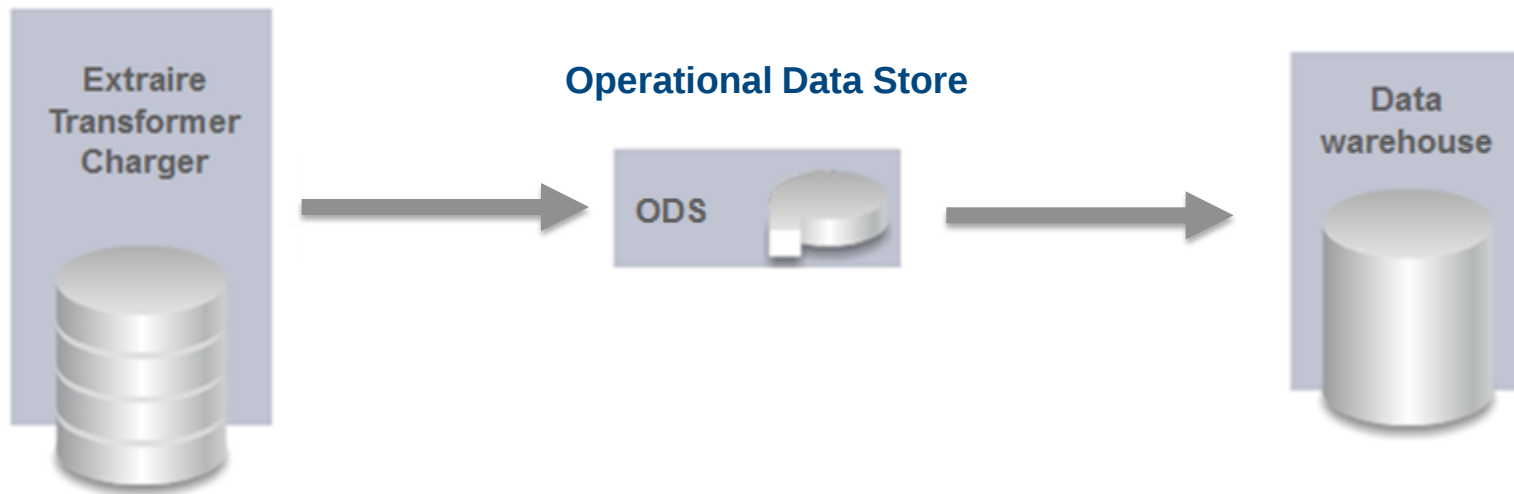
5. 4 Glossaire

Architecture globale du SID : le stockage des données

Operational Data Store (ODS) : Définition

Définition

Base de données conçue pour **centraliser et vérifier la cohérence des données issues de sources hétérogènes afin de faciliter leur intégration dans un Data Warehouse**. Ce processus implique souvent **une purge des informations redondantes**. Un ODS est généralement destiné à contenir des données quotidiennes (type événements du jour) de niveau fin contrairement au Data Warehouse qui contient l'ensemble des données historisées.



L'ODS est un élément facultatif dans une architecture BI. En fonction des besoins fonctionnels et du budget du projet l'ODS n'est pas systématiquement implémenté.

Architecture globale du SID : le stockage des données

Data Warehouse : Définition

Définition

Base de données dédiée au stockage de **l'ensemble des données utilisées dans le cadre de la prise de décision et de l'analyse décisionnelle**. Il est alimenté en données depuis les bases de production grâce notamment aux outils d'ETL. Il n'est pas une simple copie des données de production. Il est **organisé et structuré**. Son objectif est de stocker ces données de façon **intégrée** et **historisée**.



Orienté sujet

- Données organisées par thème
- Consolidation par thèmes des données des différents bases opérationnelles



Intégré

- Intégration et consolidation de données hétérogènes



Non volatile

- Traitements décisionnels uniquement en lecture (non modification des données)



Historisé

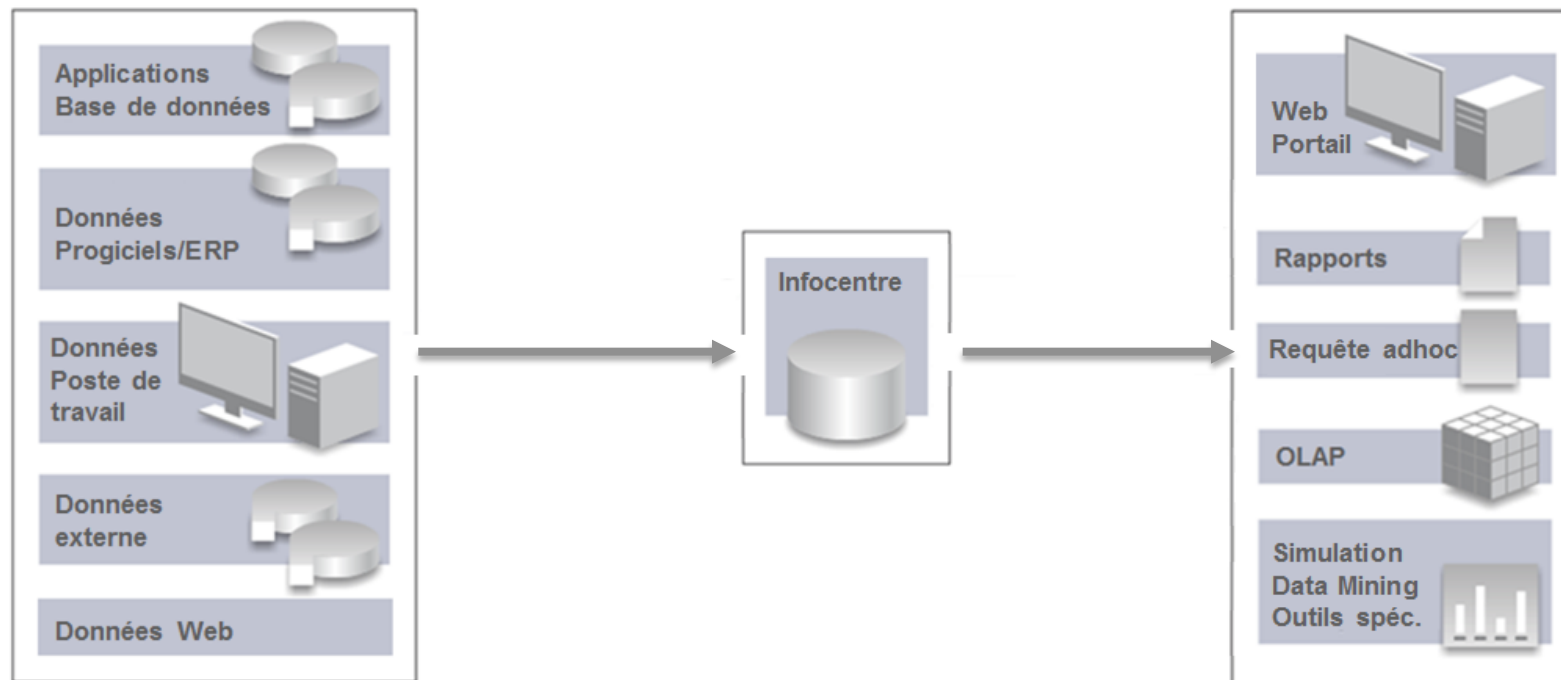
- Horodatage des données
- Visualisation de l'évolution des données dans le temps
- Sélection des données à archiver

Architecture globale du SID

L'infocentre

Définition

L'infocentre est une réplique de la base de données de production. Il permet l'analyse décisionnel uniquement sur le périmètre des données de l'application dont la base a été répliquée (BI en silos). Les analystes peuvent ainsi exécuter directement des requêtes sur cette base de données sans porter atteinte à l'intégrité et à la performance des bases de données de production. Il utilise le modèle de données de la base source, avec parfois quelques aménagements (addition d'index, tri différent, etc.) contrairement au Data Warehouse qui utilise généralement des modèles spécifiques plus adaptés à de l'analyse par les outils cibles.

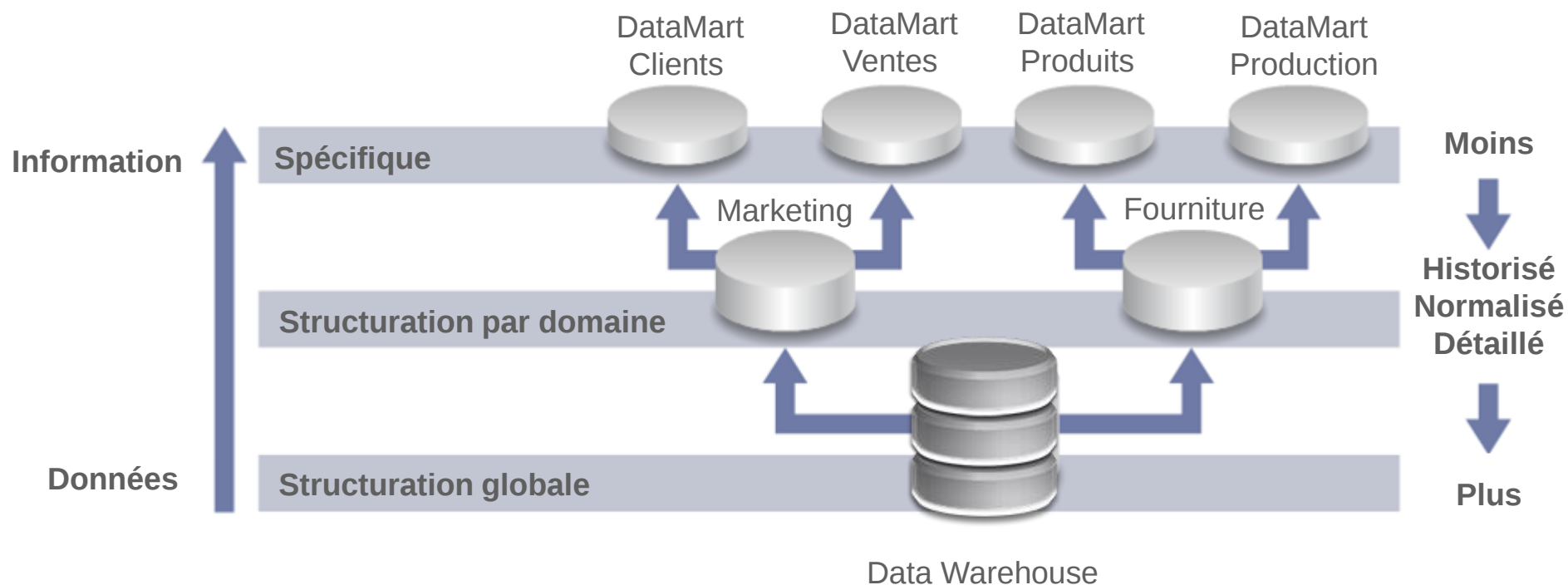


Architecture globale du SID : le stockage des données

Data Mart : Définition

Définition

Le DataMart est un **sous-ensemble d'un Data Warehouse** qui regroupe un **ensemble de données ciblées**, organisées, regroupées et agrégées pour répondre à un **besoin spécifique à un métier ou un domaine donné**. Techniquement, il peut être sous forme relationnelle (étoile, flocon) ou multidimensionnelle (hypercube).



Architecture globale du SID : l'exploitation des données

Reporting : Définition

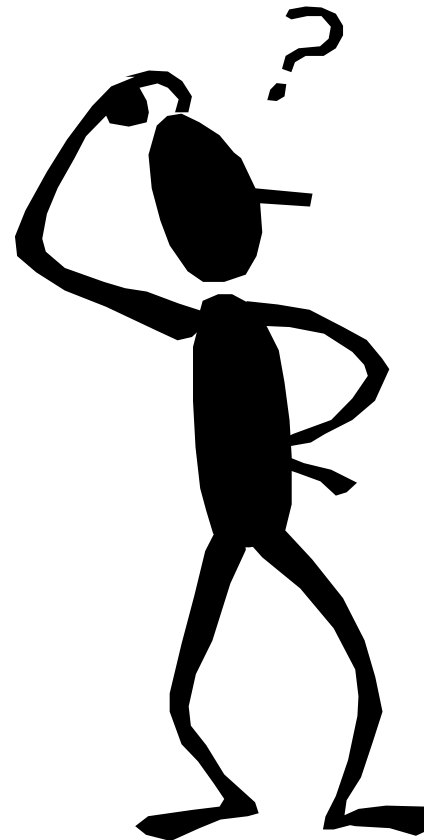
Définition

L'**outil de reporting** assure l'interrogation des bases de données selon les requêtes préparées lors de l'élaboration du modèle. L'outil assure aussi la **génération de rapports**. Le rapport peut ensuite être publié, périodiquement de manière automatique ou ponctuellement à la demande. L'outil offre des fonctions spécifiques pour **l'élaboration du modèle du rapport**, des modules de calcul et de présentation (graphiques) afin de concevoir des comptes rendus standardisés.



Exemples de rapports

Questions



The power of simplicity
«Ce qui est simple est fort»



www.solucom.fr

Contact

Olivier BESNARD
Consultant sénior

Mobile : +33 (0)6 25 36 16 25
Mail : olivier.besnard@solucom.fr