

Base de données distribuées

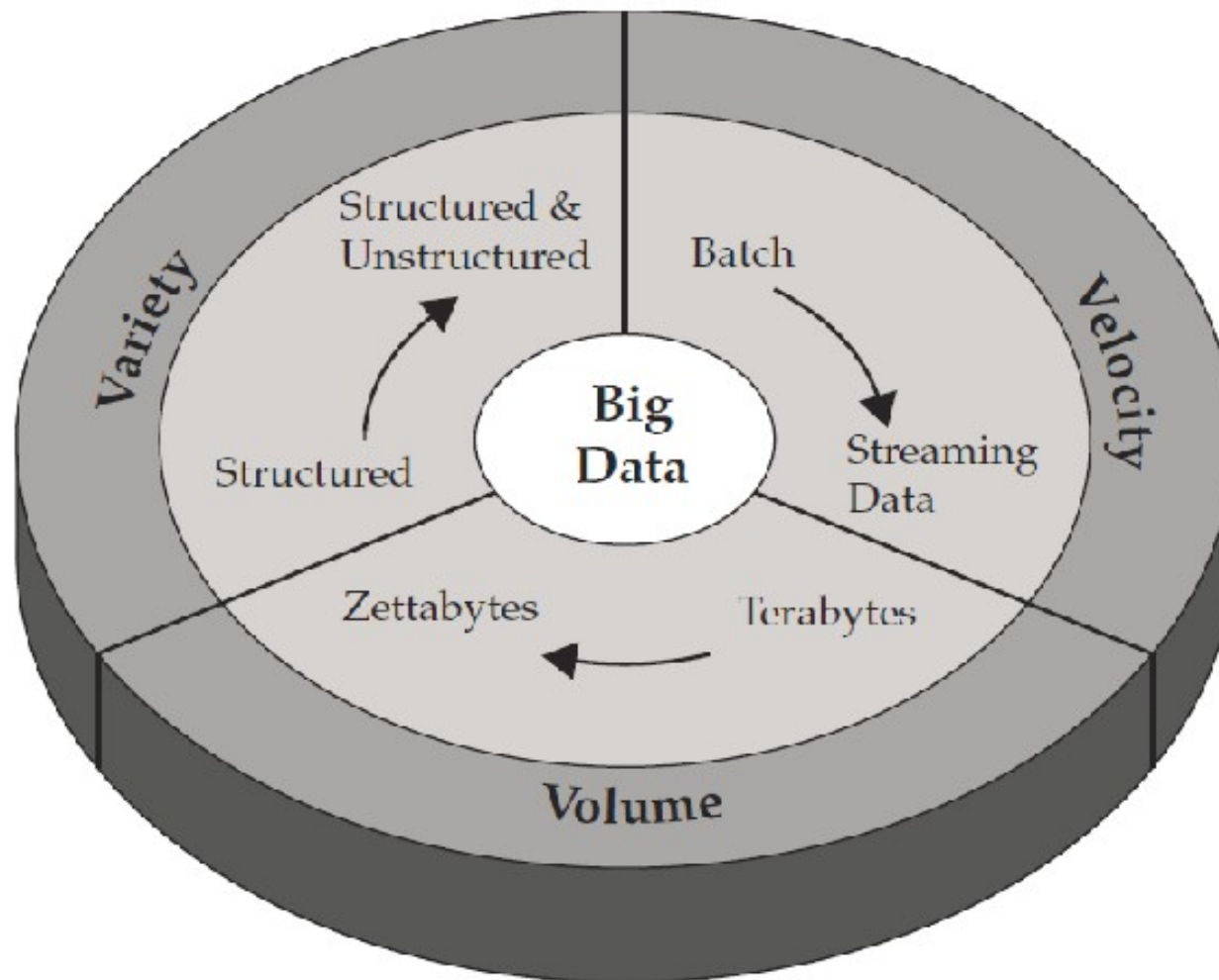
Plan

- **Motivation**
- Système distribué
- Architecture
- Réplication
- Fragmentation
- Optimisation de requêtes
- Concurrency
- Systèmes

Big data

- Big data (déluge de données)
 - Définition de Wikipedia
 - “Big data usually includes data sets with sizes beyond the ability of commonly-used software tools to capture, curate, manage, and process the data within a tolerable elapsed time.”
- Les 3 ‘V’: volume, vitesse et variété

Big data



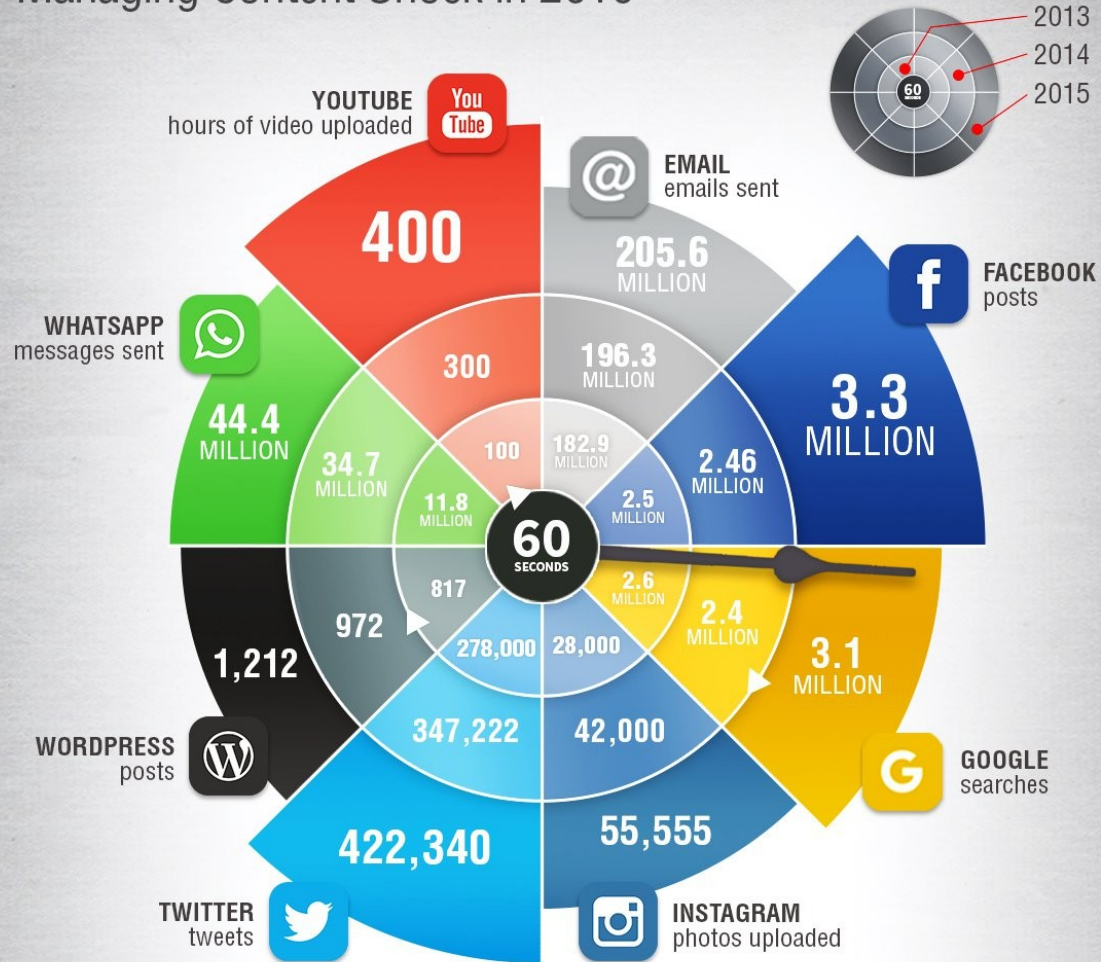
d'autres 'V'

- ♦ On retrouve des 'définitions' avec d'autres V:
 - ♦ Véracité
 - ♦ relatif à la qualité des données
 - ♦ Vocabulaire
 - ♦ Considération de la sémantique
 - ♦ Venue (location)
 - ♦ Localisation des données

Big data

What Happens Online in 60 Seconds?

Managing Content Shock in 2016



The world has fallen in love with social media and now automatically turns to online platforms to research and buy products and services. This gives fantastic opportunities for marketers to engage audiences and encourage content sharing, but also gives huge challenges of getting cut-through and keeping up-to-date ourselves!

At Smart Insights, we look to help by focusing on the 'Must-know' platform developments and developing mind tools to help businesses review how they can best Plan, Manage and Optimize their digital marketing – see our <http://bit.ly/smartlibrary>

Brought to you by:



www.smartinsights.com

univ-mlv.fr

2021 *This Is What Happens In An Internet Minute*

Big data



Big data

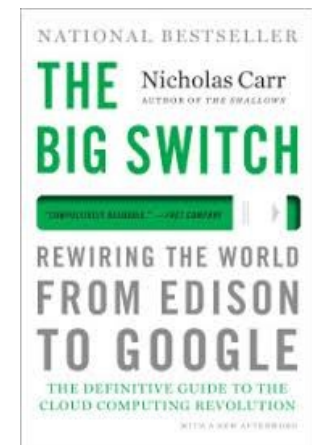
- D'autres 'V'
 - Véracité: qualité des données
 - Vocabulaire: sémantique
 - Venue (en anglais) : localisation

Autres motivations

- Loi de Moore
 - Vitesse d'horloge des ordinateurs ne croit plus
 - Plus de coeurs
 - Augmentation des perf. passe par le parallélisme

Autres motivations

- Loi de Moore
 - Vitesse d'horloge des ordinateurs ne croit plus
 - Plus de coeurs
 - Augmentation des perf. passe par le parallélisme
- Cloud computing
 - l'accès à de large clusters de machines devient un commodité
 - AWS, Azure, GCP, etc.



DATAAllegro
Microsoft



Amazon Redshift

ORACLE

VERTICA

SAP HANA

SYBASE



MarkLogic

Microsoft
SQL Server

GEMSTONE

teradata.



MySQL

Objectivity
Innovate with Confidence

Google
BigQuery

HyperSQL

xindice
TECHNOLOGY



SQLite

ORACLE
BERKELEY DB

VoltDB

VERTICA

Tamino
eXist

snowflake

Greenplum

INGRES

Apache Derby

VERSANT
THE DATABASE FOR DEVELOPERS

aster data
big data. fast insights.

MONETDB

ScaleDB

NETEZZA



calpont

TREMEDATA

db4objects

PARACCEL

INFOBRIGHT

vectorwise

MariaDB

Relational

Object Relational



Embedded



Memory centric



Column stores / data warehouse



Object



XML



Network

IDMS

Hierarchy

IBM IMS

Key Value stores



MEMBASE.ORG

Voldemort

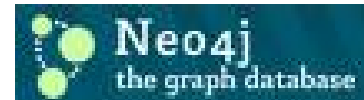
TokyoCabinet

Scalaris

Amazon Dynamo & SimpleDB

Dynomite

Graph databases



AllegroGraph



Document databases

Terrastore



Column family databases

Google BigTable



Plan

- Motivation
- **Système distribué**
- Architecture
- Réplication
- Fragmentation
- Optimisation de requêtes
- Concurrency
- Systèmes

Systeme distribue

- “A distributed system is one in which the failure of a computer you didn’t even know existed can render your own computer unusable.” Leslie Lamport
- Un système coordonnant les actions de plusieurs ordinateurs dans l’objectif de réaliser une tâche particulière

Système distribué

- ♦ A distributed system is a collection of autonomous computing elements that appears to its users as a single coherent system.
- ♦ Limitations:
 - ♦ No global clock
 - ♦ Gérer l'appartenance à un groupe (group membership)
 - ♦ Le réseau n'est pas fiable

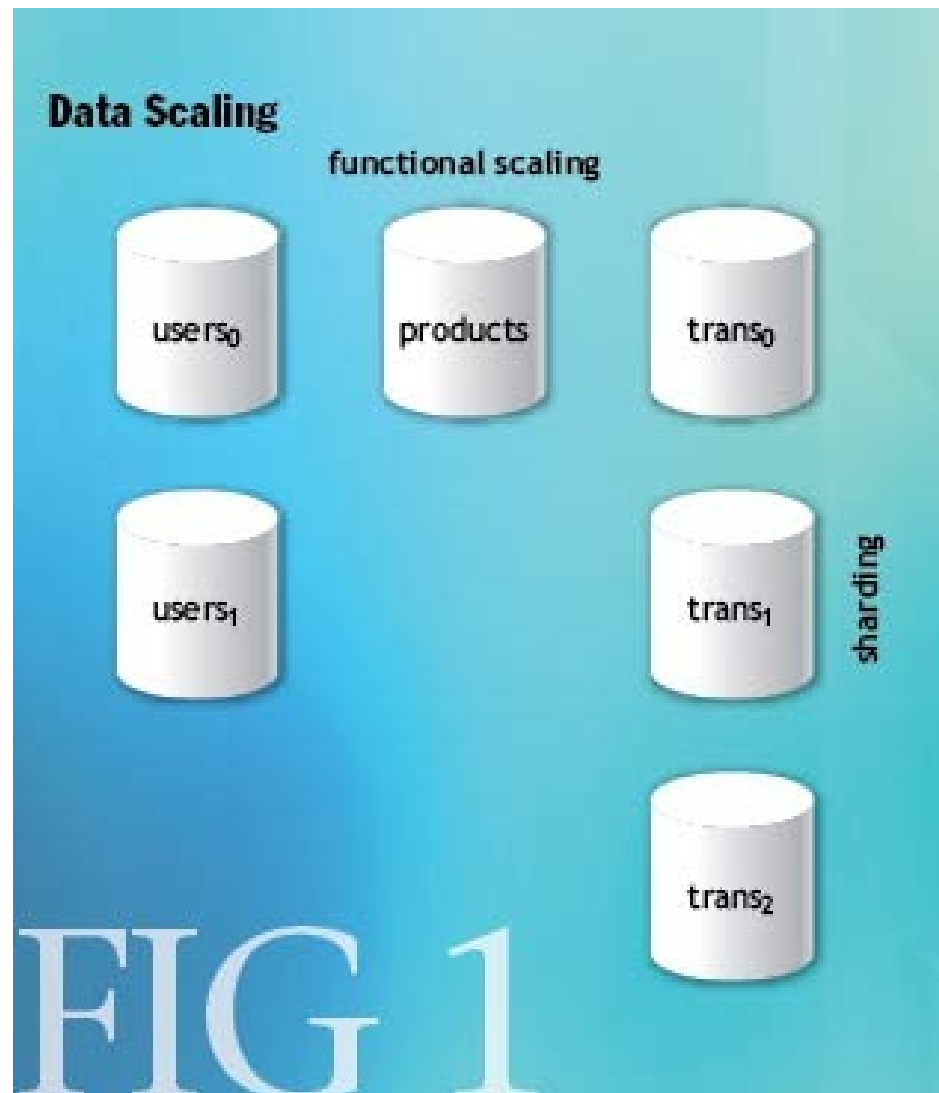
Caractéristiques d'un syst. dist.

- Concurrence, parce que plusieurs traitements sont effectués en parallèle
- Communication entre ordinateurs
- Tolérance aux pannes, la panne d'une machine ne doit pas empêcher les traitements des autres machines
- Passage à l'échelle
- Autonomie et hétérogénéité des machines

Passage à l'échelle

- ♦ Scalability
 - ♦ La capacité d'un système à gérer une quantité grandissante de tâches d'une manière élégante ou sa capacité à s'accroître pour accommoder la croissance
 - ♦ 2 approches pour le passage à l'échelle
 - ♦ Vertical (vertical scaling, scaling up)
 - ♦ On ajoute de la mémoire, des disques, des CPUs
 - ♦ Horizontal (horizontal scaling, scaling out)
 - ♦ Distribution des données/traitements sur un plus grand nombre de machine :
 - ♦ Functional scaling, sharding : distribution des données

Passage à l'échelle



2 approches pour le traitement en parallèle

- Système de gestion de base de données distribuées (SGBDD)
 - Premiers systèmes dans les années 80
 - OLTP (Online Transaction Processing) pour la gestion d'un grand nombre de transactions
 - OLAP (Online Analytic Processing) pour la prise de décision
- MapReduce (Google 2004) => cours en M2 Logiciel et Ingénierie des données

Base de données

BD centralisée :

=> une machine, un SGBD

Vulnérabilité

Coût d'infrastructure

Inadaptation aux structures d'entreprises décentralisées

BD distribuée :

=> Plusieurs machines, plusieurs SGBD

Coopération entre sites informatiques

Partage de données

Interopérabilité des données

SGBD parallèle vs distribué

- ♦ SGBD parallèle
 - ♦ Les noeuds sont physiquement proches
 - ♦ Noeuds connectés par LAN avec haut débit
 - ♦ Coût de communication est considéré comme faible
- ♦ SGBD distribué
 - ♦ Noeuds peuvent être éloignés les uns des autres
 - ♦ Noeuds connectés par un réseau public
 - ♦ Coût de communication n'est pas négligeable et les problèmes associés sont importants

Parallélisme inter et intra requête

- ♦ Inter-requête:
 - ♦ Différents requêtes sont exécutées de manière concurrente
 - ♦ Augmentation du débit et réduction de la latence
- ♦ Intra-query
 - ♦ Exécution des opérations d'une simple requête en parallèle
 - ♦ Diminution de la latence pour des requêtes dont l'exécution peut être longue

Base de données distribuées

- Une grande quantité de données résidant sur plusieurs machines
- Système de gestion de base de données distribuées est un ensemble de logiciels offrant un point d'accès unique à une base de données distribuées

SGBDD

- Objectif
 - Améliorer les perf en exécutant plusieurs opérations en parallèle
- Bénéfice
 - Passage à l'échelle (scalabilité) "sans limite", moins onéreux qu'avec une unique machine
 - Tolérance aux pannes et haute disponibilité
 - Réduction du temps de latence
- Défi
 - Garantir la perf en présence d'une augmentation des échanges entre machines et accès croissant aux ressources

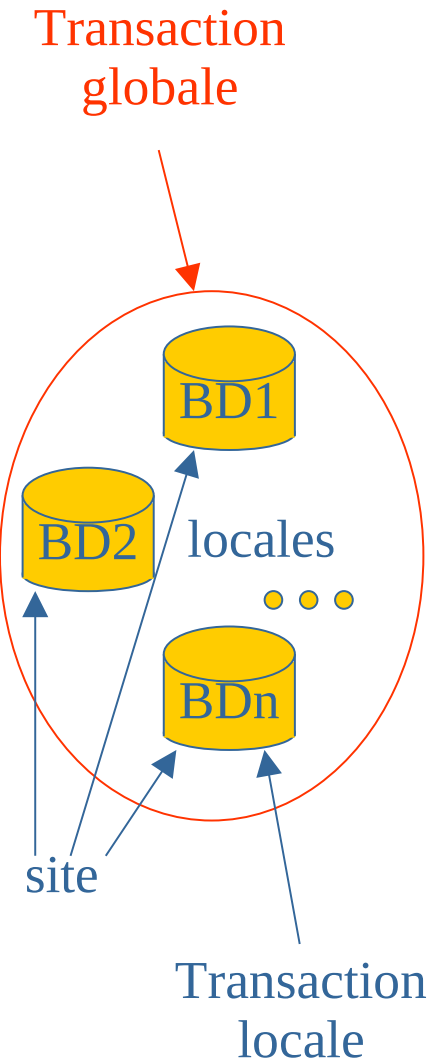
SGBDD - objectifs

- Concilier des systèmes existants (fédération),
- Partager des données, passage à l'échelle (scalabilité)
- Limiter les transferts de données,
- Répartir la charge de calcul (parallélisme),
- Augmenter pour les données :
 - La fiabilité,
 - La disponibilité,
- Permettre l'évolution d'une structure d'entreprise
 - Extension d'un système existant,
 - Fusion de systèmes existants.

SGBDD- défis

- Administration et maintenance difficiles,
- Distribution optimale des données entre les sites.
- Distribution des contrôles
- Prise en compte des pannes possibles.
- Conversation de la cohérence des données du système
- Coût de mise en place élevé,
- Garantir la perf en présence d'une augmentation des échanges entre machines et accès croissant aux ressources

SGBDD - définition



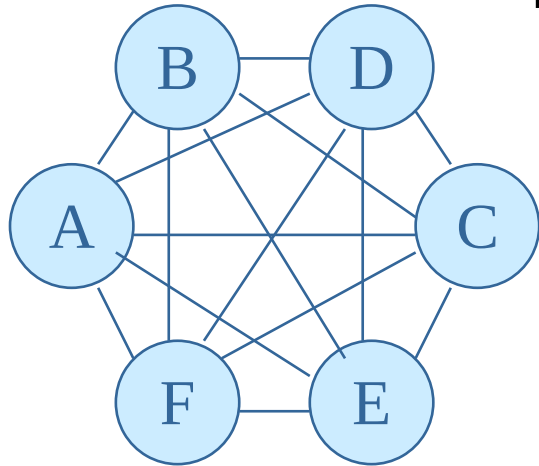
- Base de données distribuées : BD dont les données sont disséminées entre plusieurs sites.
- Partage des ressources du réseau constitué par ces sites.
- 1 site = 1 BD
- Chaque site peut gérer des transactions **locales** (données sur site),
- + Doit savoir exécuter des transactions **globales** (données sur plusieurs sites).

Structure (1/2)

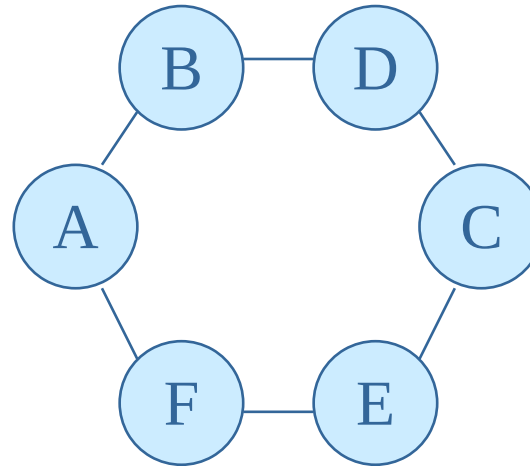
- Connexions inter-sites représentées sous forme de graphes
 - Nœud = site
 - Arc = connexion directe
- Choix de configuration surtout en fonction de :
 - Coût d'installation : frais d'installation d'une liaison physique entre 2 sites,
 - Coût d'exploitation : frais et durée de transmission,
 - Fiabilité : fréquence de pannes (liaison ou site),
 - Disponibilité : accessibilité au système en cas de pannes.

Structure (2/2)

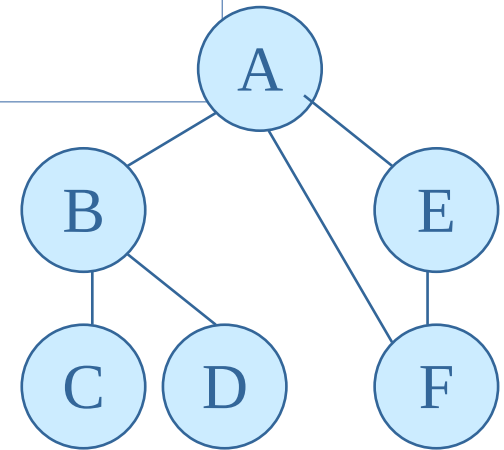
Configurations (logiques) les plus fréquentes :



Réseau totalement connecté

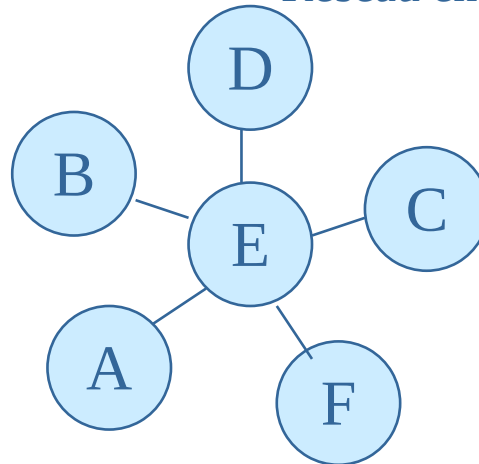


Réseau en anneau

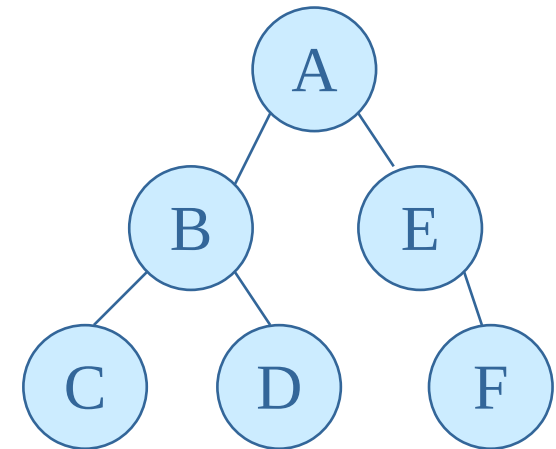


Réseau partiellement connecté

A rapprocher des structures du réseaux physique



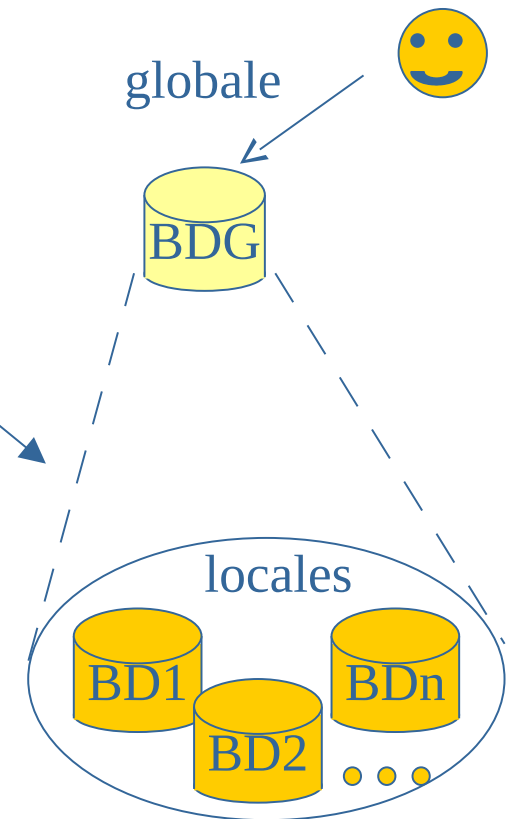
Réseau en étoile



Réseau arborescent

Principes des BD distribuées

- Une seule base de données vue par l'utilisateur
- Il ne voit pas les Pbs
 - de fiabilité
 - d'optimisation
 - de transactions réparties
 - de localisation
- Transparence totale :
 - accès via un schéma conceptuel
 - accès via une vue



Avantages de la distribution des données.

- Efficacité et fiabilité d'accès à des données partagées.
- Partage des données et gestion répartie
 - Accès à partir d'un site à l'ensemble des données (Évite la concentration de données),
 - Gestion locale (maîtrise de ses propres données pour chaque site) / mais autonomie relative.
- Fiabilité et disponibilité
 - Maintien de l'exploitation en cas de panne par substitution des sites,
 - Duplication des données.

Inconvénients de la répartition des données.

- Complexité pour coordonner :
 - Coût de développement logiciel,
 - Potentialité plus forte d'erreurs logicielles (bugs cachés, algo. Fragilisés par parallélisme),
 - Servitudes système accrues pour la coordination :
 - Échange de messages,
 - Calculs supplémentaires,
 - Récupération du système plus complexe après pannes (réintégration des sites ou liaisons en panne).

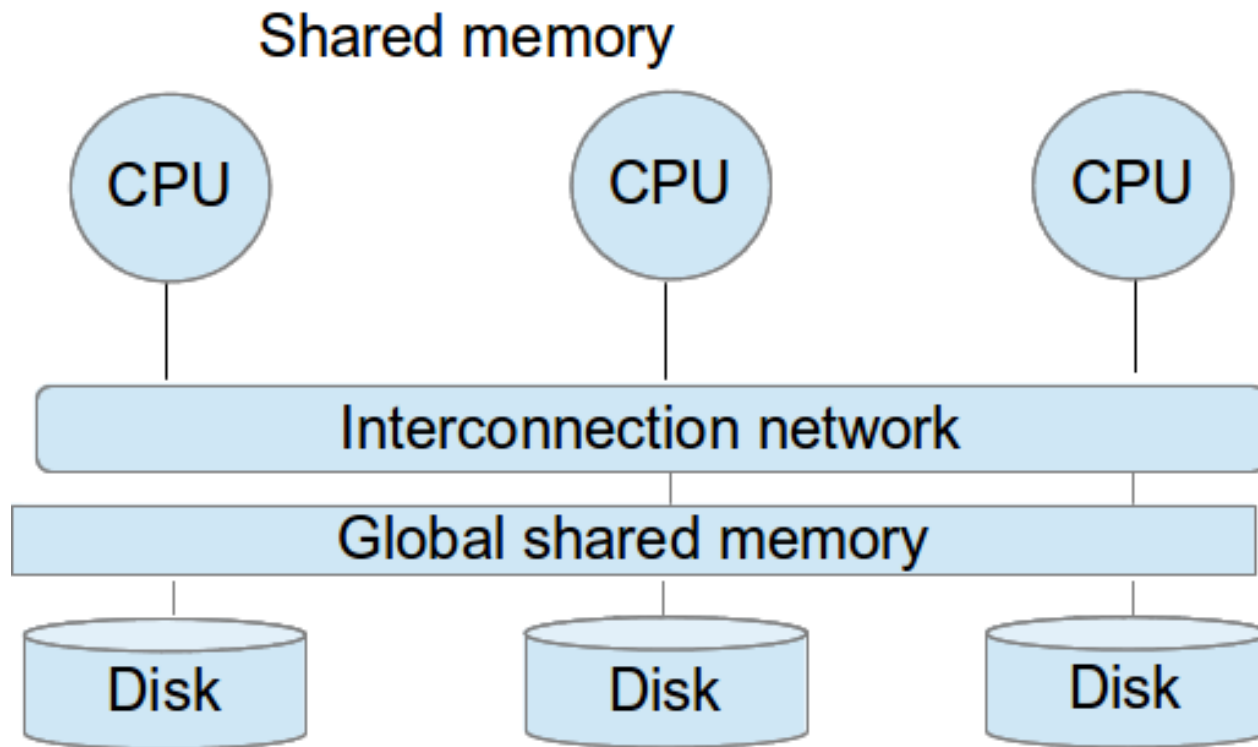
Plan

- Motivation
- Système distribué
- **Architecture**
- Réplication
- Fragmentation
- Optimisation de requêtes
- Concurrency
- Systèmes

Architecture

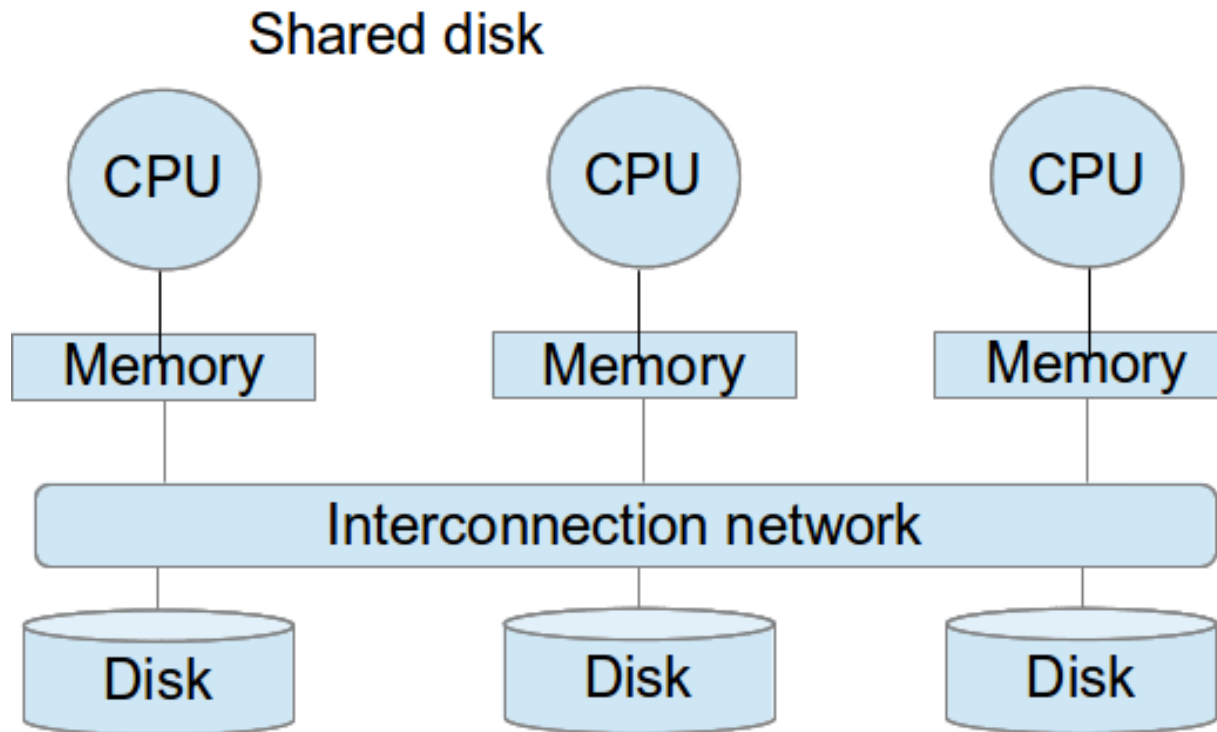
- Shared memory
- Shared Disk
- Shared nothing

Architecture - Shared memory

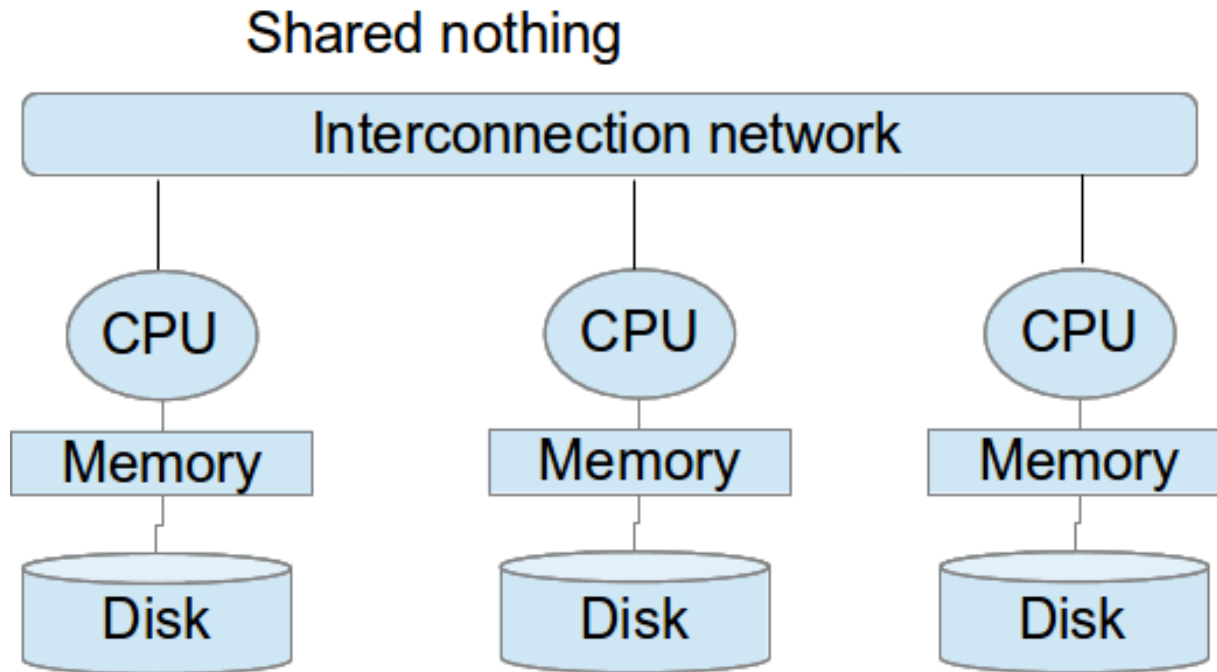


eXtremeDB[®]

Architecture - Shared Disk



Architecture - Shared nothing



TERADATA.

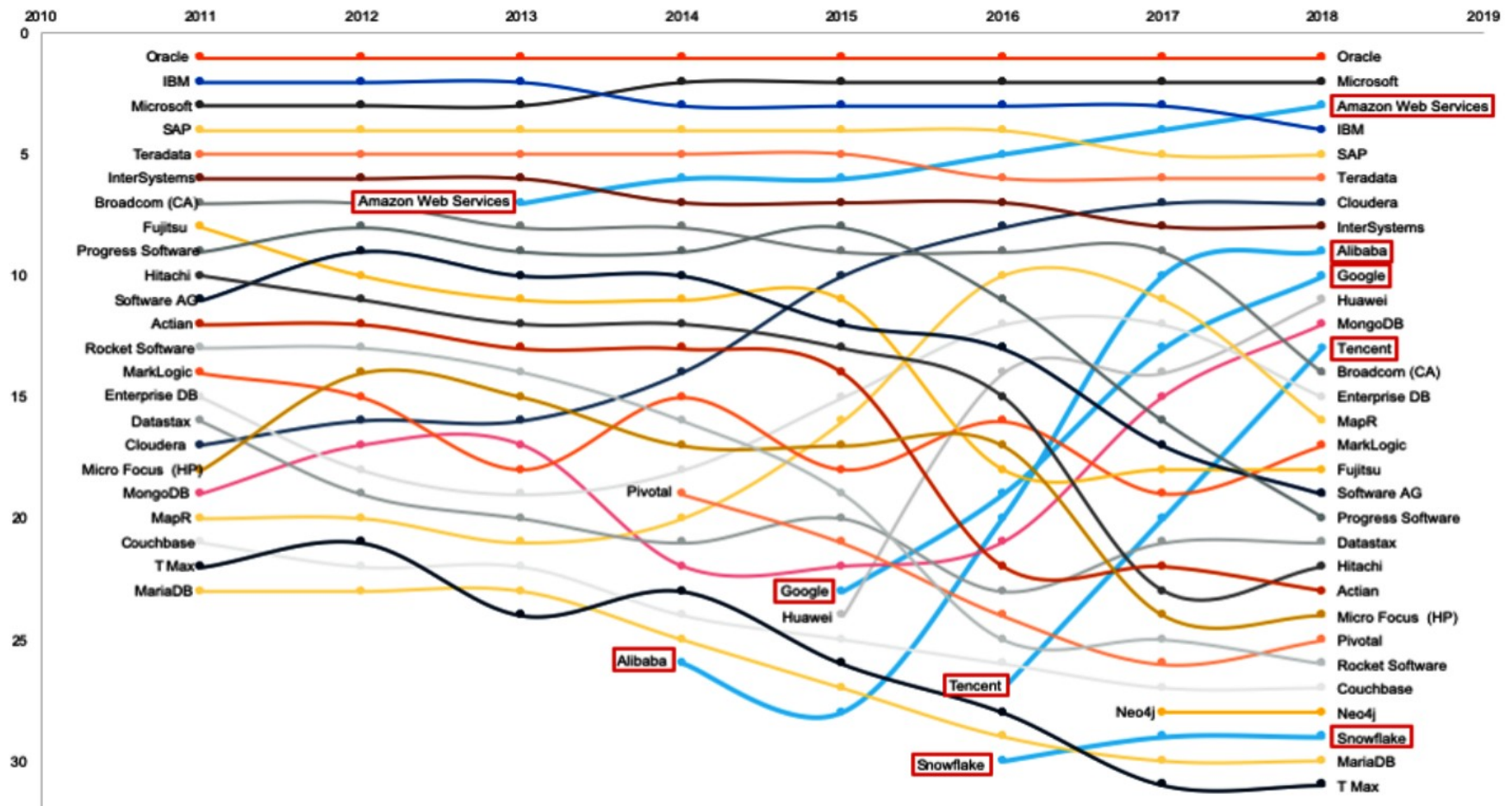
Architecture - bilan

- Shared memory
 - Coût de coordination entre processus est élevé
 - Passage à l'échelle difficile
- Shared Disk
 - Coût intéressant pour un cluster de petite et moyenne taille (pas plus de 10 machines)
- Shared nothing
 - Adopté dans le Cloud computing
 - Efficace pour le passage à l'échelle mais un coût élevé de partitionnement

SGBD et Cloud

Gartner Market Share Ranking, 2011-2018

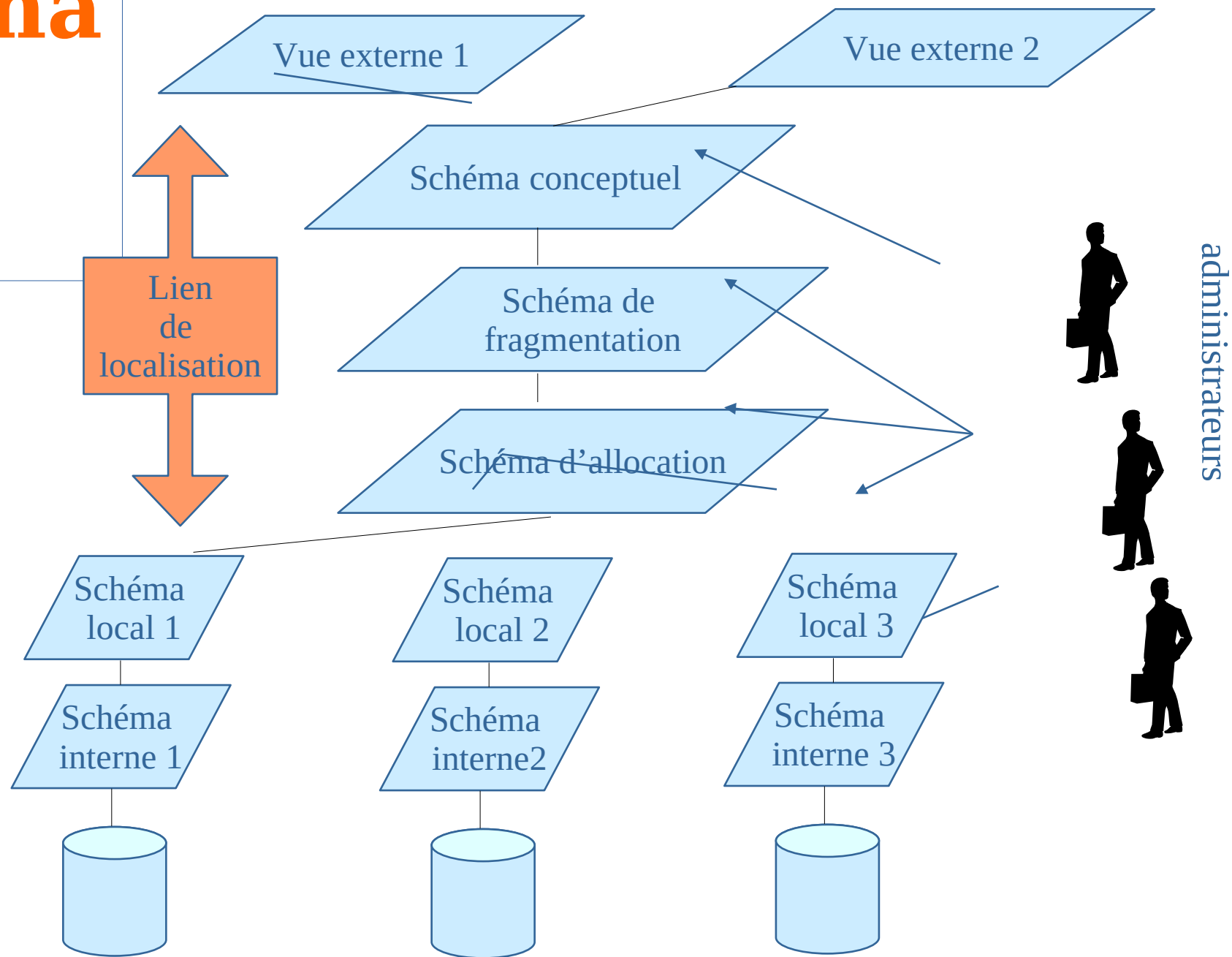
Rank



Principes

- Principes (propriétés) des SGBD centralisés,
- Plus les options suivantes, soit une relation R
 - **Réplication** : R copiée sur plusieurs sites,
 - **Fragmentation** : R découpée en plusieurs fragments, chacun hébergé par un site,
 - **Réplication et fragmentation** : R découpée en plusieurs fragments copiés sur divers sites.

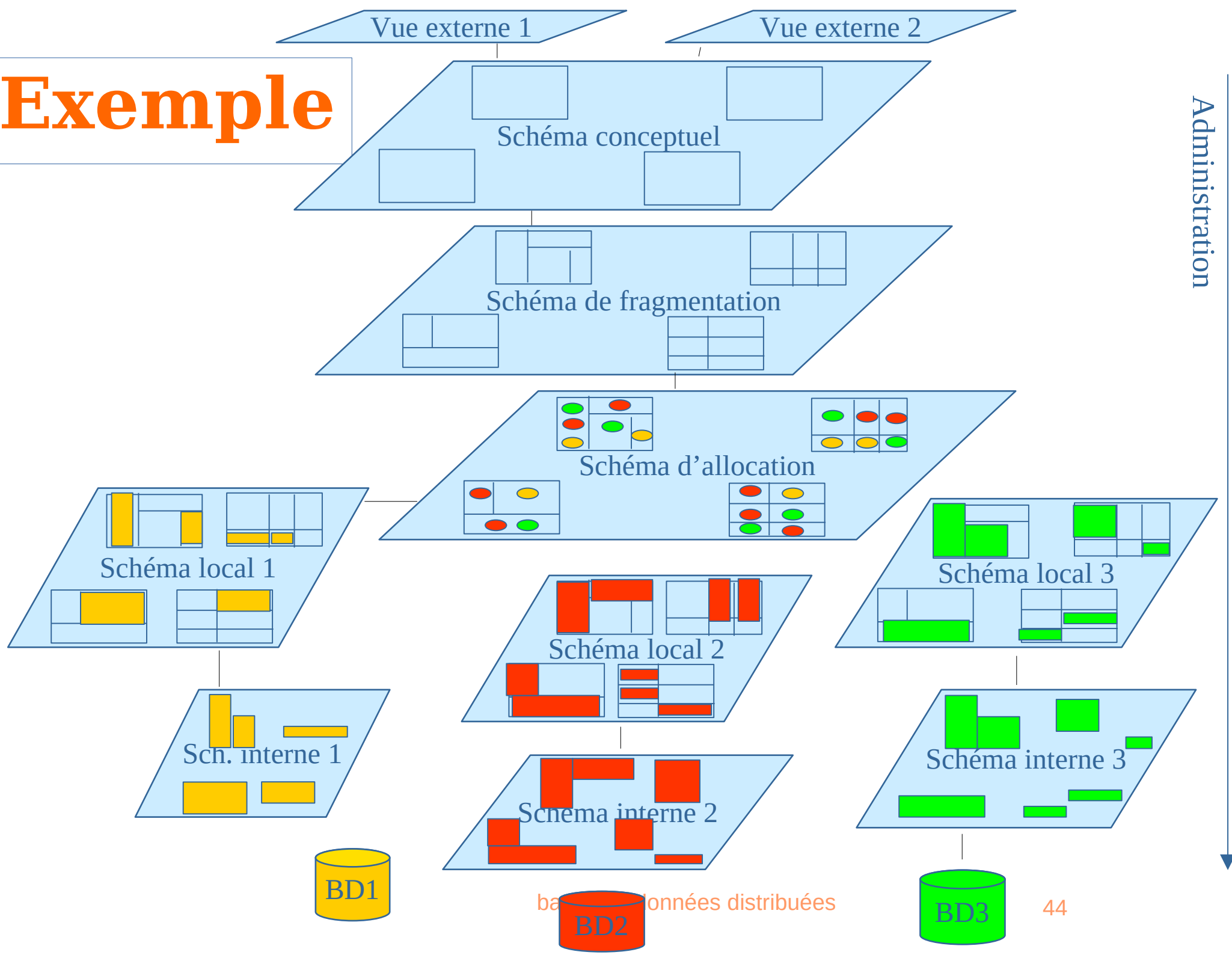
Schéma d'une BDD



Définitions des niveaux de schémas

- les besoins des utilisateurs sont intégrés dans un schéma conceptuel (unique).
- des vues peuvent être dérivées du schéma conceptuel.
- l'administrateur construit le schéma de fragmentation (découpage du schéma conceptuel en plusieurs sous-schémas avec des liens entre les deux)
- l'administrateur définit le schéma d'allocation (localisation des fragments sur les différents sites)
- Des schémas logiques sont créés automatiquement sur chaque site en fonction des fragments alloués.
- L'administrateur doit finalement décrire les schémas internes.

Exemple



Résumé sur l'approche SGBDD

- Transparence pour l'utilisateur
 - Réplication (*duplication*)
 - Partitionnement (*fragmentation*)
 - Localisation
- Conception sur plusieurs niveaux
 - Comme BD classique
 - Association des vues et schémas locaux
 - Lien de localisation (difficile à établir)
 - groupe d'utilisateur symbolisée par une vue \neq un site
 - + Définition des fragments
 - Schéma de fragmentation
 - + Allocation des fragments à un site
 - Schéma d'allocation

Stockage des données

- Fragmentation
 - Les relations sont partitionnées en différents fragments, chacun sur un site distinct.
- Réplication
 - Le système maintient différentes copies des données sur les sites.
- Fragmentation et réplication
 - Combinaison de la réplication et de la fragmentation.