

# SLAM3

Formes normales

# FN !

- \* Ce n'est pas le nom d'une partie politique !
- \* Non MERISE ne fait pas de la politique
- \* C'est juste l'abréviation de Forme Normale
- \* MERISE propose des modèles normalisés
- \* A condition d'appliquer des règles de normalisation
- \* Il existe au moins 6 niveaux de normalisation pour les modèles relationnel
- \* L'objectif de la normalisation c'est de supprimer les redondances de données et les problèmes d'identification, c'est à dire les dépendances fonctionnelles

# SGBDR

- \* Les SGBDR appartiennent à la famille des système OLTP (Online Transaction Processing)
  - \* Traitement transactionnel en ligne
  - \* Il y a donc modification d'informations en temps réel
  - \* (lecture & écriture)
  - \* La normalisation vient ici perturber le concept de performance dans ce cas
- \* Pour la famille de OLAP (Online Analytical Processing)
  - \* Traitement analytique en ligne
  - \* Il y a ici uniquement de la lecture de données, et donc le concept de performance est plus simple à régler.

# Normalisation

- \* Rôle :
  - \* Ennuyer les étudiants !
  - \* Eviter les anomalies transactionnelles, et donc éviter l'annulation d'une transaction
  - \* La transaction c'est le tendon d'Achille des SGBDR
  - \* Eviter les redondances, et donc les problèmes de performances
  - \* La performance (intégrité référentielle) est aussi le tendon d'Achille des SGBDR

# Normalisation

- \* Permet de vérifier la robustesse de conception du modèle
- \* Permet d'améliorer la représentation du modèle
- \* Chaque forme normale de niveau inférieur s'emboîte dans celle du niveau supérieur
- \* Le dernier niveau inclut donc tous les niveaux inférieurs
- \* Avant d'appliquer la normalisation, faites en sorte que chaque relation possède une clé primaire robuste

# Normalisation

- \* Les avantages :

- \* de limiter les redondances de données (multiples écritures)
- \* donc de limiter l'espace disque nécessaire
- \* de limiter les incohérences de données qui pourraient les rendre inutilisables (multiples écritures)
- \* d'éviter les processus de mise à jour (réécritures)

- \* Les inconvénients :

- \* des temps d'accès potentiellement plus longs si les requêtes sont trop complexes (lectures plus lentes)
- \* Il faut alors recourir aux index
- \* une plus grande fragilité des données étant donné la non redondance (lecture impossible)
- \* la difficulté de réorganiser la structure après normalisation
- \* Il faut alors dénormaliser et posséder le modèle initial pour gagner du temps

# Normalisation

- \* Pour les petites BD limiter la normalisation en 3FN
- \* Si vous écrivez beaucoup dans une BD il est préférable de normaliser
- \* Si vous lisez beaucoup dans un BD il est préférable de ne pas normaliser (limitation des performances)

# Niveaux de normalisation

- \* 1FN : Forme normale la plus simple
  - \* 2FN : Problème d'identifiant composé : dépendance partielle clé multi-attributs
  - \* 3FN : Problème de transitivité des attributs non clés
  - \* La forme de Boyce Codd FNBC (dans le genre de la 3FN)
  - \* 4FN : normalement on s'arrête là en BTS : clé multi-attributs à décomposer
  - \* 5FN : problème de dépendance de jointure ...
  - \* La forme normale domaine clé FNDC
  - \* 6FN : très rarement présentée
- 
- \* Pour toutes les FN il faut raisonner au niveau relationnel voir physique
  - \* On peut raisonner sur un MCD mais il faut imaginer le niveau relationnel en direct.



# 1FN

- \* La première forme normale se caractérise par la notion d'atomicité
- \* Ce qui veut dire qu'un attribut ne peut contenir qu'une seule valeur à la fois
- \* Exemple suivant ->

# 1FN

pas en 1FN

isbn	auteur
2-212-09283-0	Gardarin
2-7117-8645-5	Abiteboul, Hull, Vianu

en 1FN



isbn	auteur
2-212-09283-0	Gardarin
2-7117-8645-5	Abiteboul
2-7117-8645-5	Hull
2-7117-8645-5	Vianu

# 2FN

- \* Il faut déjà que le modèle valide la 1FN
- \* Si la clé est composée d'un seul attribut, alors la relation est directement en 2FN
- \* Si la clé est composée de plusieurs attributs (multi-attribut), il faut que les attributs non clés dépendent totalement de la clé et non partiellement
- \* En clair, un attribut non clé dépend complètement de la clé, et pas uniquement d'une partie
- \* Exemple suivant ->

# 2FN

pas en 2FN

isbn bib → nb\_ex  
isbn → titre

<u>isbn</u>	titre	<u>bib</u>	nb_ex
2-212-09283-0	Bases de données ...	Toulon	3
2-7117-8645-5	Fondements des ...	Toulon	1
2-212-09283-0	Bases de données ...	Aix-Marseille 3	7

en 2FN



en 2FN

<u>isbn</u>	<u>bib</u>	nb_ex
2-212-09283-0	Toulon	3
2-7117-8645-5	Toulon	1
2-212-09283-0	Aix-Marseille 3	7

<u>isbn</u>	titre
2-212-09283-0	Bases de données ...
2-7117-8645-5	Fondements des ...

**Cette décomposition est sans perte et préserve les dépendances fonctionnelles.**

# 2FN

\* En 2FN mais pas en 3FN (transitivité)

en 2FN

isbn → titre éditeur  
éditeur → pays

<u>isbn</u>	titre	éditeur	pays
2-212-09283-0	Bases de données ...	Eyrolles	France
2-7117-8645-5	Fondements des ...	Vuibert	USA
2-212-0969-2	Internet/Intranet ...	Eyrolles	France

**Il reste des redondances :** (... , Eyrolles , France).

# 3FN

- \* Une relation est en 3FN, si elle est déjà en 2FN
- \* Si un attribut non clé dépend fonctionnellement d'un autre attribut non clé, il y a dépendance fonctionnelle transitive
- \* Il ne doit pas exister de DF transitive, uniquement des DF directes

# 3FN

pas en 3FN

isbn bib → nb\_ex  
isbn → titre

<u>isbn</u>	titre	<u>bib</u>	nb_ex
2-212-09283-0	Bases de données ...	Toulon	3
2-7117-8645-5	Fondements des ...	Toulon	1
2-212-09283-0	Bases de données ...	Aix-Marseille 3	7

en 3FN



en 3FN

<u>isbn</u>	<u>bib</u>	nb_ex
2-212-09283-0	Toulon	3
2-7117-8645-5	Toulon	1
2-212-09283-0	Aix-Marseille 3	7

<u>isbn</u>	titre
2-212-09283-0	Bases de données ...
2-7117-8645-5	Fondements des ...

**Cette décomposition est sans perte et préserve les dépendances fonctionnelles.**

# 3FN

en 2FN  
pas en 3FN

isbn → titre éditeur  
éditeur → pays

<u>isbn</u>	titre	éditeur	pays
2-212-09283-0	Bases de données ...	Eyrolles	France
2-7117-8645-5	Fondements des ...	Vuibert	USA
2-212-0969-2	Internet/Intranet ...	Eyrolles	France

en 3FN

<u>isbn</u>	titre	éditeur
2-212-09283-0	Bases de données ...	Eyrolles
2-7117-8645-5	Fondements des ...	Vuibert
2-212-0969-2	Internet/Intranet ...	Eyrolles

en 3FN

<u>éditeur</u>	pays
Eyrolles	France
Vuibert	USA

**Cette décomposition est sans perte et préserve les dépendances fonctionnelles.**



# 3FN

- \* En 3FN mais pas en 4FN (insuffisance de la 3FN)

rue ville  $\rightarrow$  cp  
cp  $\rightarrow$  ville

en 3FN car :  
{rue, ville} est une super clé  
ville appartient à une clé candidate

clés candidates :  
{rue, ville}  
{rue, cp}

rue	cp	ville
Rue des Lilas	83100	Toulon
Rue des Pins	83100	Toulon
Rue des Palmiers	13008	Marseille
Rue des Lauriers	13013	Marseille

**Il reste des redondances :** (... , 83100 , Toulon) .

# 4FN

- \* Elle s'occupe des dépendances multivaluées,
- \*  $X \twoheadrightarrow Y$  X multidétermine Y
- \*  $R(X,Y,Z)$
- \* Pour chaque valeur de X il correspond toujours le même ensemble de valeurs de Y et cet ensemble de valeurs ne dépend pas des valeurs de Z.
- \* C'est le cas d'un livre avec plusieurs auteurs

# 4FN

- \* Soit la relation livre (isbn, mots-clé, auteur)
- \* La clé est l'ensemble des trois attributs : clé composée de plusieurs attributs (multi-attributs)
- \* Si un livre peut avoir plusieurs auteurs, la relation livre possède la dépendance multivaluée :
- \*  $isbn \twoheadrightarrow auteur$
- \*  $isbn \twoheadrightarrow (DM) \rightarrow auteur$
- \* Une relation  $R$  est en 4e forme normale, si pour chaque dépendance multivaluée  $X \twoheadrightarrow Y$ ,  $X$  est une super-clé de  $R$ .

# 4FN

\* On s'attaque ici à démonter des relations dynamiques

pas en 4FN

$isbn \twoheadrightarrow auteur$   
 $isbn \twoheadrightarrow mot\_clé$

<u>isbn</u>	<u>auteur</u>	<u>mot clé</u>
2-86601-368-9	adiba	bd
2-86601-368-9	collet	bd
2-86601-368-9	adiba	objet
2-86601-368-9	collet	objet

en 4FN

<u>isbn</u>	<u>auteur</u>
2-86601-368-9	adiba
2-86601-368-9	collet



en 4FN

<u>isbn</u>	<u>mot clé</u>
2-86601-368-9	bd
2-86601-368-9	objet

**Cette décomposition est sans perte et préserve les dépendances fonctionnelles.**

# 4FN

- \* Si on coupe la relation précédente en deux :
  - \* isbn  $\rightarrow$  auteur
  - \* isbn  $\rightarrow$  mot-clé
- \* La jointure de ces deux relations reproduit la relation initiale

# 4FN

- \* Normalement les modèles sont normalisés en 3FN, ce qui limite la complexité par une décomposition également extrême
- \* Une relation est en 4FN si elle est déjà en 3FN
- \* La 4FN comme la 5FN vise à supprimer toutes les redondances de valeurs dans les colonnes d'une table

# 5FN

- \* On est toujours dans le démontage des relations dynamiques
- \* Une relation peut être en 5FN si elle est déjà en 4FN !
- \* La 5FN s'occupe des dépendances de jointures (DJ)
- \* Attention certaines relations n'ont pas de décompositions possibles en 5FN
- \* Ce dernier niveau est donc rarement mis en pratique !
- \* C'est le niveau ultime !

# 5FN

- \* De manière très triviale :
- \* La 4FN découpe une relation multi-attributs en « n » parties
- \* La 5FN découpe une relation multi-attributs en « n + m » parties !
- \* Si le découpage de la 4FN ne suffit pas à reproduire la relation initiale, c'est qu'il faut passer en 5FN !



# 5FN




- \* Une DM est un cas particulier de DJ, voici un cas inbuvable !

VINS	BUVEUR	CRU	PRODUCTEUR
	Pierre	Chablis	Claude
	Pierre	Chablis	Nicolas
	Pierre	Volnay	Nicolas
	Paul	Chablis	Nicolas

Si on décompose VINS en 2 relations à 2 attributs chacune, on ne peut plus retrouver la relation initiale, et ce quelle que soit la décomposition choisie :

R1	BUVEUR	CRU	R2	BUVEUR	PRODUCTEUR
	Pierre	Chablis		Pierre	Claude
	Pierre	Volnay		Pierre	Nicolas
	Paul	Chablis		Paul	Nicolas

R3	CRU	PRODUCTEUR
	Chablis	Claude
	Chablis	Nicolas
	Volnay	Nicolas

VINS  $\circ$  R1  R2  
 VINS  $\circ$  R1  R3  
 VINS  $\circ$  R2  R3

Pourtant, VINS contient des redondances :

2 fois Pierre Chablis

2 fois Chablis Nicolas

=> Il existe des relations non décomposables en 2 relations, mais en 3 ou plus.

# 4FN & 5FN

- \* Avec les remarques qui viennent d'être faites, il faut se méfier des multipattes (n-aires) -> ternaires
- \* Il est parfois nécessaire de les décomposer en relations plus simples
- \* Voir même d'ajouter une notion de porteuse de données

# FNBC FNDC 6FN

- \* FNBC (Boyce Codd) FNDC (domaine clé) et 6FN ne sont pas présentées ici.
- \* Elles ne présentent pas toujours un intérêt dans la portée de nos travaux.