



4^{èmes} Journées francophones sur les **Entrepôts de Données et l'Analyse en ligne**
5 et 6 Juin 2008 Toulouse - France

EDA'08

Impact de l'évolution de nomenclature sur l'entrepôt de données

Inès Zouari Turki
Ines.Zouari@isimsf.rnu.tn

Faiza Ghozzi Jedidi
Jedidi_Faiza@yahoo.fr

Rafik Bouaziz
Raf.Bouaziz@fsegs.rnu.tn

**Laboratoire MIRACL,
Faculté des Sciences Économiques et de Gestion de Sfax - Tunisie**



Contexte et motivations (1/2)

■ Entrepôt de données (DW)

- Une collection de données **orientées sujet, intégrées, non volatiles** et **historisées**, organisées pour un support d'aide à la décision (Inmon, 1996)

■ Évolution au niveau d'un DW

- Évolution des sources de données
- Évolution des besoins d'analyse des utilisateurs

Nécessité de faire **évoluer le schéma** du DW afin de répondre à ces besoins



Contexte et motivations (2/2)

- Évolution des sources de données
 - Évolution influant sur le schéma du DW
 - Évolution influant sur les données du DW
 - Évolution de la **nomenclature des membres**
 - influe sur le **schéma** et sur les **données** de DW
 - **Exemple** : DW SONEDE



Plan

- 1- État de l'art sur l'évolution au niveau des DW
- 2- Évolution de nomenclature : Problèmes et solutions
- 3- Prototypage
- 4- Conclusion et perspectives



Plan

1- État de l'art sur l'évolution au niveau des DW

2- Évolution de nomenclature : Problèmes et solutions

3- Prototypage

4- Conclusion et perspectives



État de l'art

Approches d'évolution

■ Plusieurs travaux traitant des problèmes d'évolution de DW

■ Deux approches à distinguer :

Changement structurel de DW	<ul style="list-style-type: none">- Évolution du schéma et/ou des instances de DW- Aucune historisation <p><i>(Blaschka et al., 1999, Hurtado et al., 1999, Favre et al., 2007)</i></p>
Historisation et/ou versionnement de DW	<ul style="list-style-type: none">- Historisation des MemD et/ou du schéma de DW <p><i>(Bliujute et al. 1998, Mendelzon et al., 2000)</i></p> <ul style="list-style-type: none">- Versionnement des schémas de DW <p><i>(Body et al., 2002, Bebel et al., 2004, Bebel et al., 2006, Eder et al., 2001, Eder et al., 2002, ...)</i></p>



État de l'art

Types d'évolution

- Évolution du schéma de DW (*Eder et al. 2001*)
 - Insertion/Suppression de dimension, de niveau, de hiérarchie, de mesure
- Évolution des membres de dimension
 - *Opérations simples sur les membres (Eder et al. 2001)*
 - Création, Suppression, Modification (de valeur de clé ou d'attribut)
 - Subdivision, Fusion, ...
 - *Opérations complexes sur les membres (Body et al. 2003)*
 - Diminution (Subdivision + Suppression), Augmentation, ...

État de l'art & Problématique

- Peu de travaux traitent de l'évolution de nomenclature
 - Évolution simple de valeur de clé de membre (*Eder et al., 2002a*)
- Des cas complexes d'évolution de nomenclature

*Différents types
d'évolution de
nomenclature ?*

*Perturbations
engendrées sur
les analyses ?*

*Solutions à ces
perturbations ?*





Plan

- 1- État de l'art sur l'évolution au niveau des DW
- 2- Évolution de nomenclature : Problèmes et solutions**
- 3- Prototypage
- 4- Conclusion et perspectives



Notre approche

Systeme de gestion de DW multiversions

■ Principes de base :

■ Versionnement

Cubes
Dimensions
Niveaux
Membres de dimension

■ Historisation

Attributs de niveau
Mesures
Relations de hiérarchie entre niveaux
Relations de hiérarchie entre membres

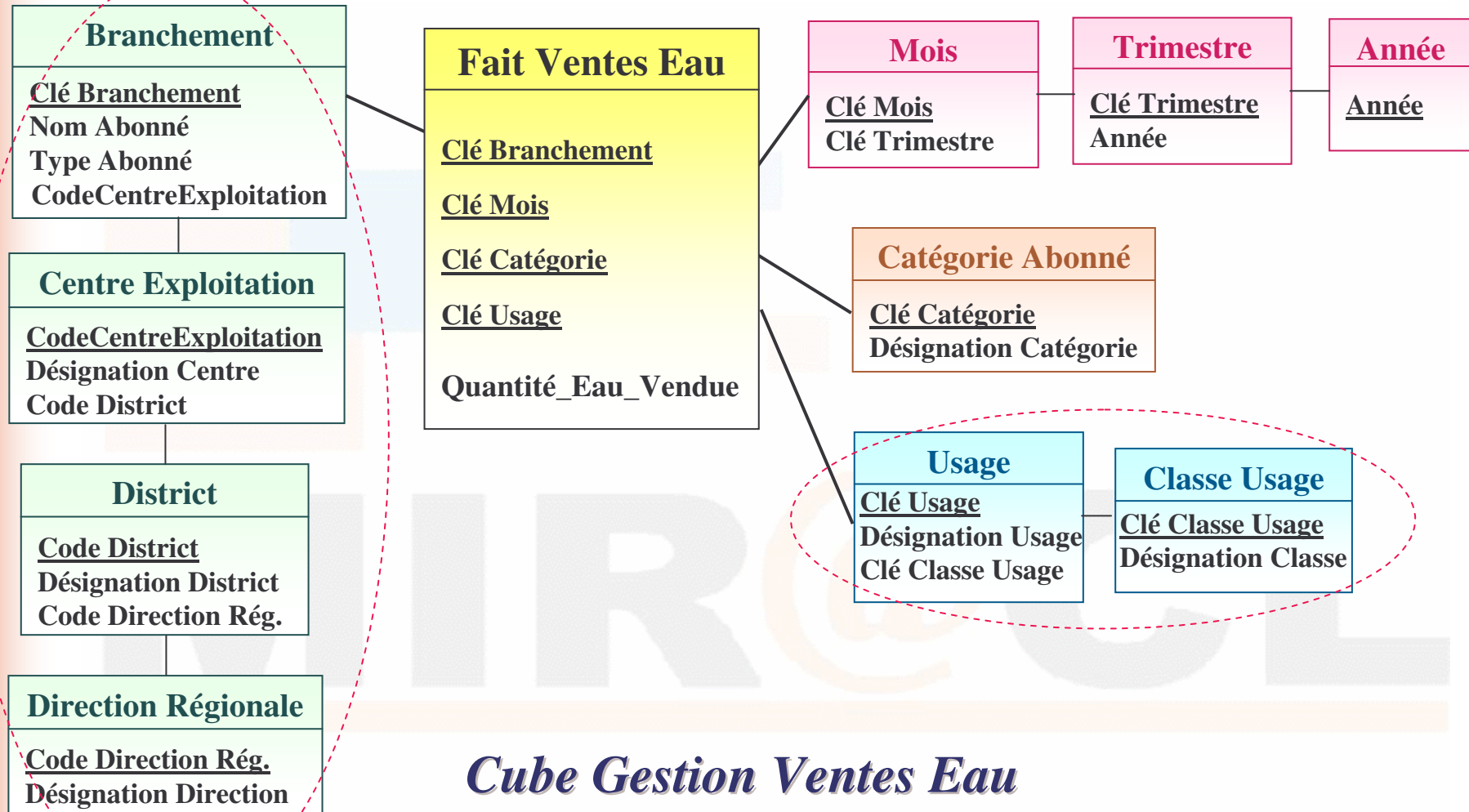
■ Estampillage par les temps d'application

Temps de Début d'Application (TDA)
Temps de Fin d'Application (TFA)



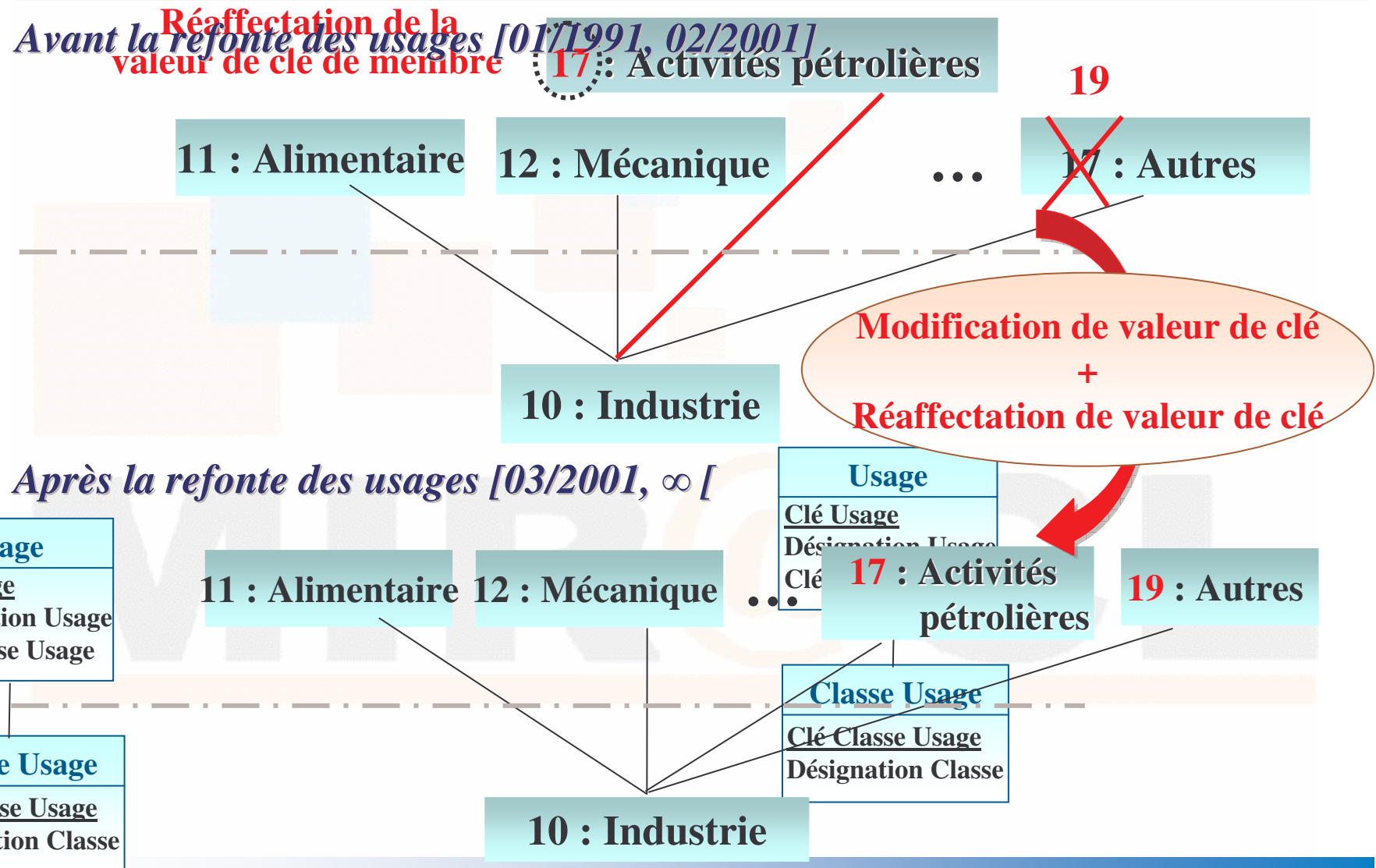
Exemple motivant

SONEDE



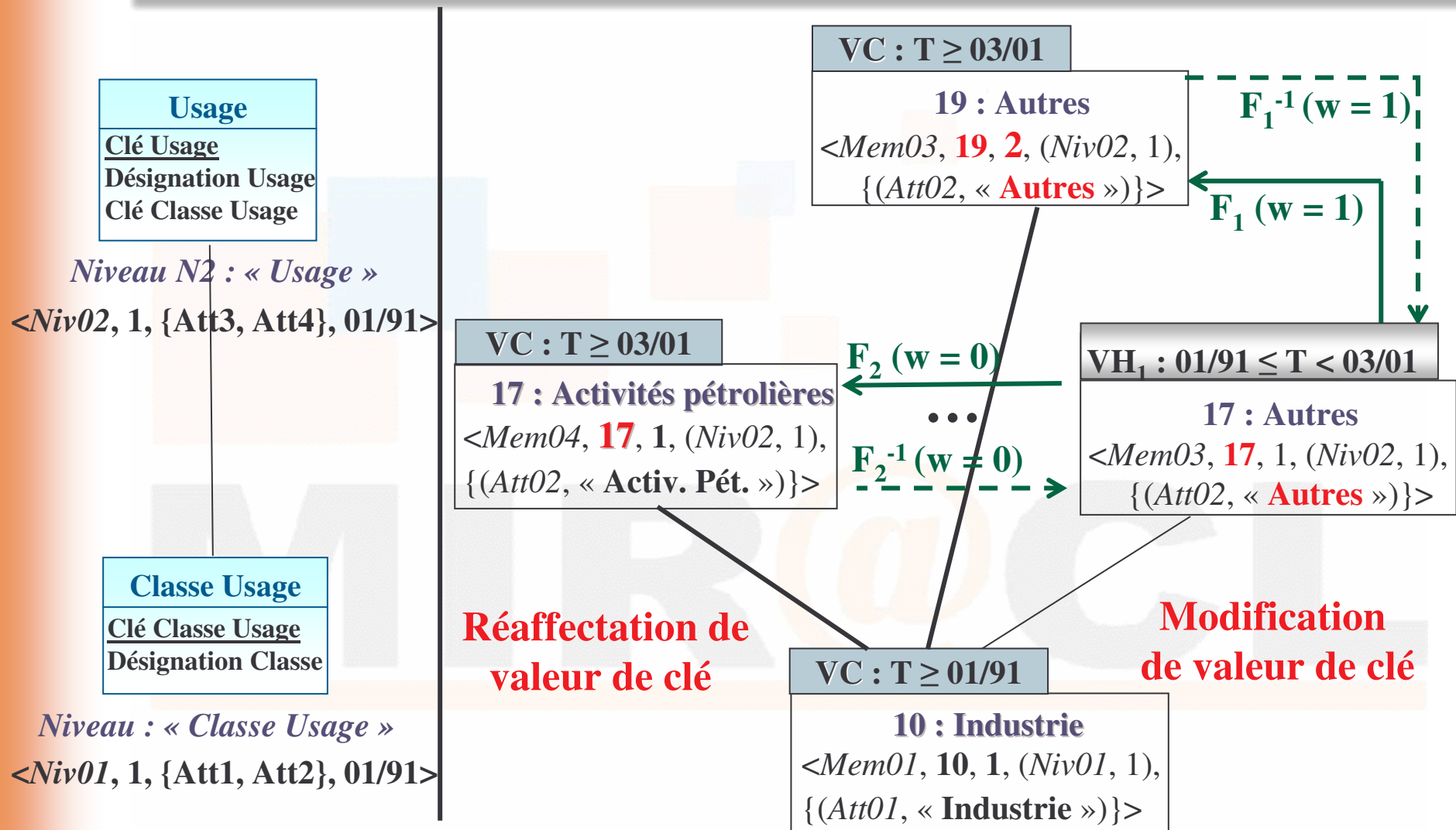
Évolution de nomenclature

Modification/Réaffectation de valeur de clé (1/2)



Évolution de nomenclature

Modification/Réaffectation de valeur de clé (2/2)





Évolution de nomenclature

Subdivision perturbatrice de membre (1/2)

Avant la refonte des usages [01/1991, 02/2001]

11 : Alimentaire 18 : Huilerie

Usage
Clé Usage
Désignation Usage
Clé Classe Usage

11 : Alimentaire

12 : Mécanique

...

17 : Autres

Classe Usage
Clé Classe Usage
Désignation Classe

10 : Industrie

Subdivision perturbatrice de membre

Après la refonte des usages [03/2001, ∞ [

11 : Alimentaire

18 : Huilerie

12 : Mécanique

...

19 : Autres

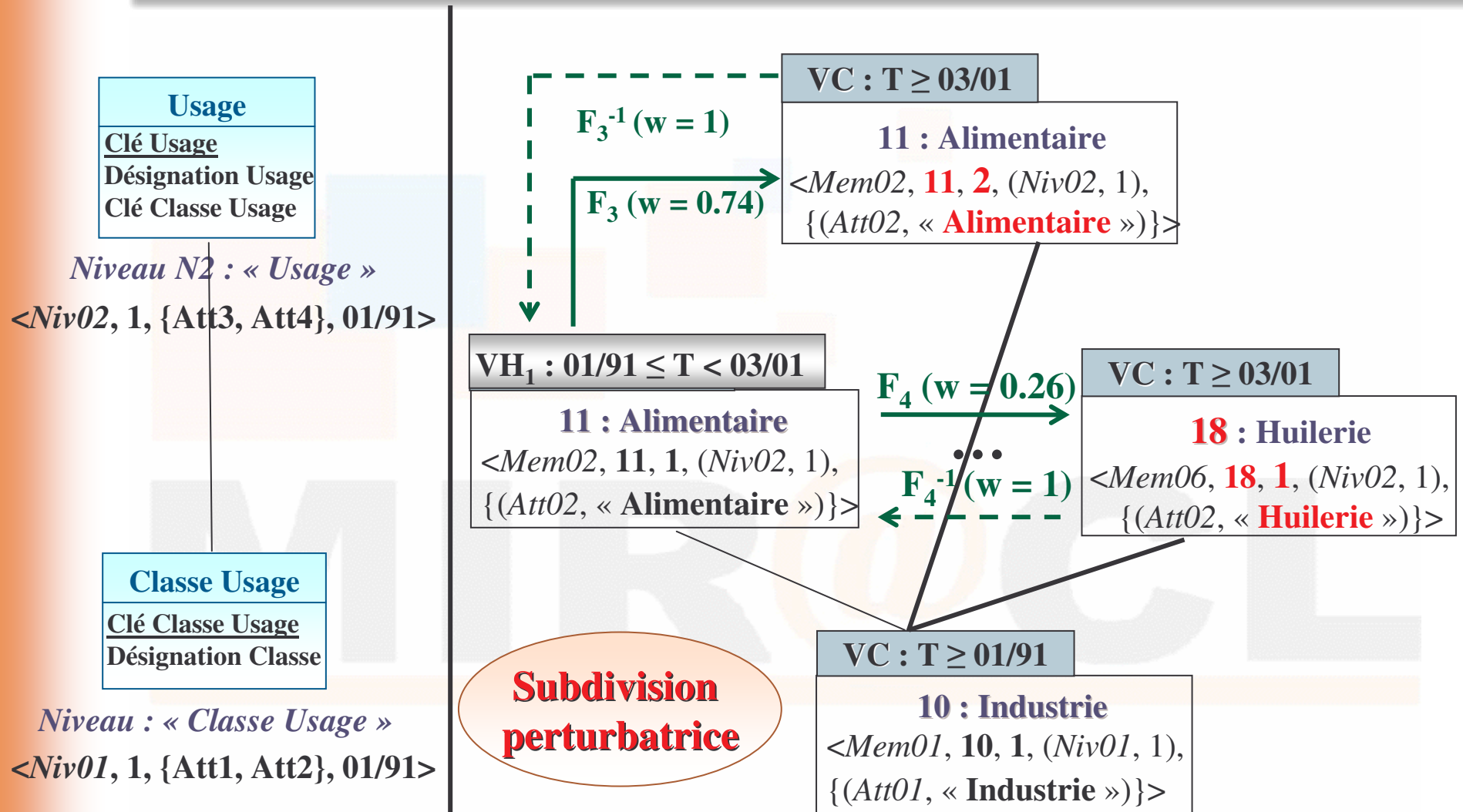
Usage
Clé Usage
Désignation Usage
Clé Classe Usage

10 : Industrie



Évolution de nomenclature

Subdivision perturbatrice de membre (2/2)

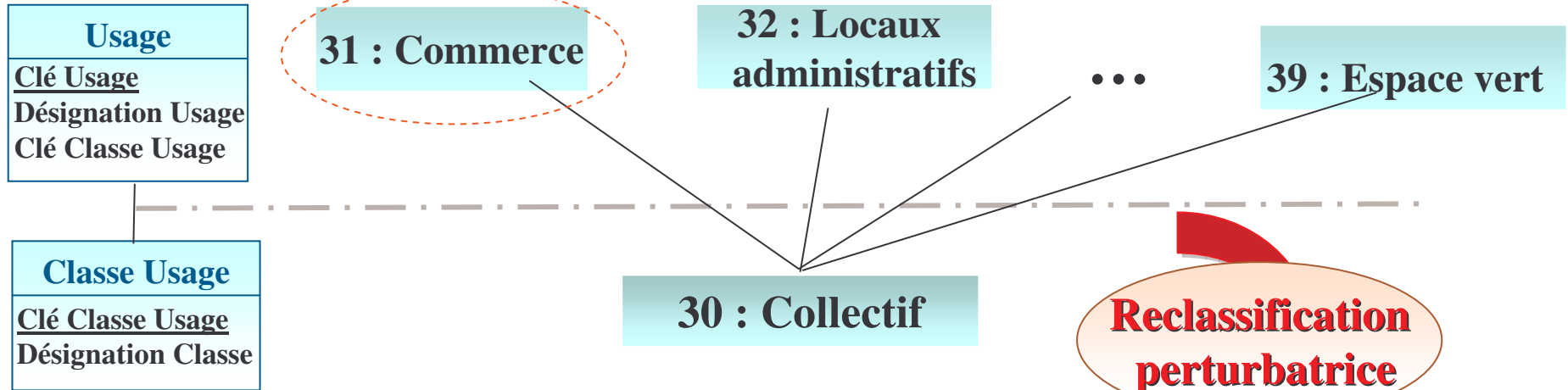




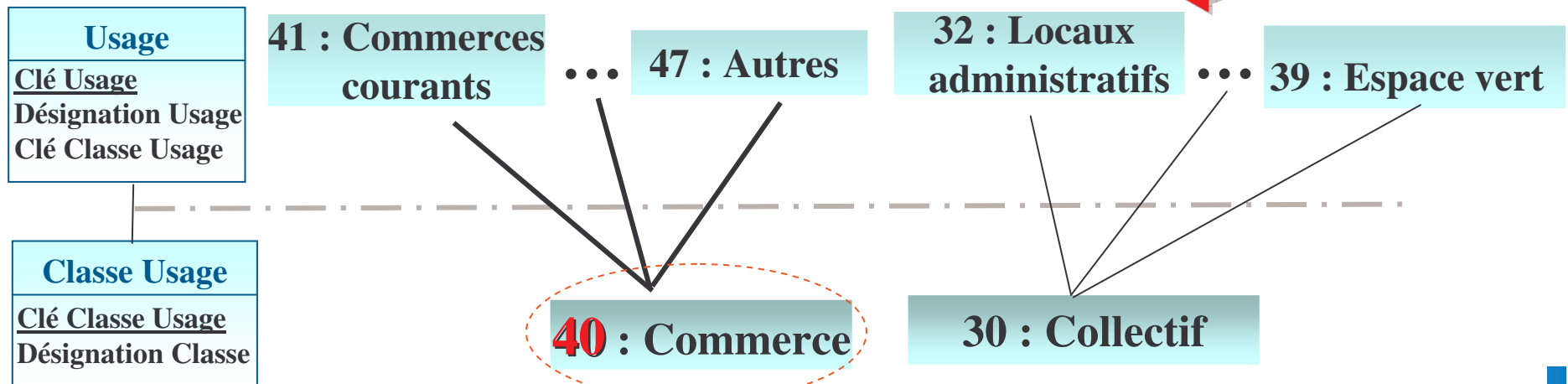
Évolution de nomenclature

Reclassification perturbatrice de membre (1/2)

Avant la refonte des usages [01/1991, 02/2001]

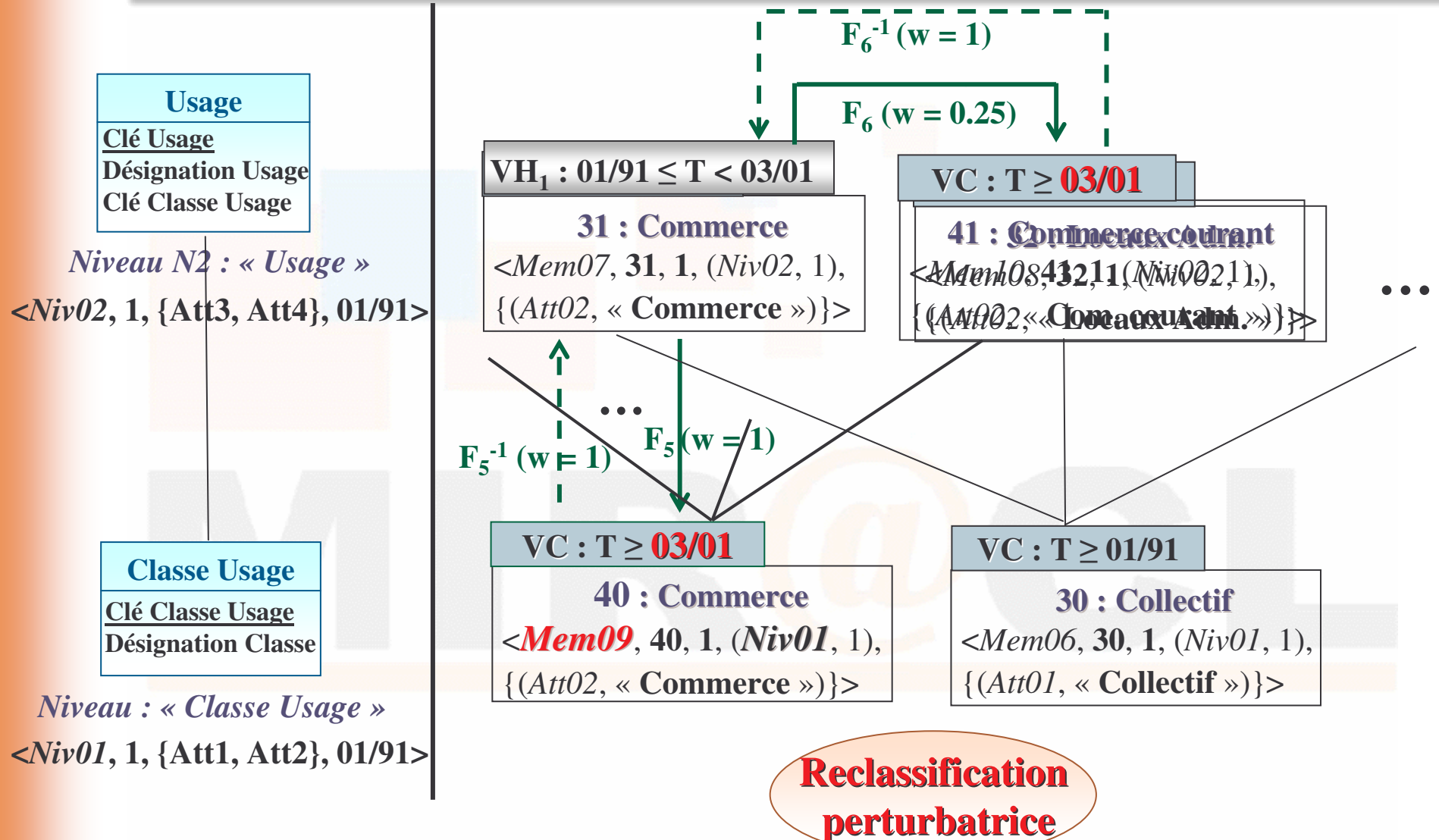


Après la refonte des usages [03/2001, ∞ [



Évolution de nomenclature

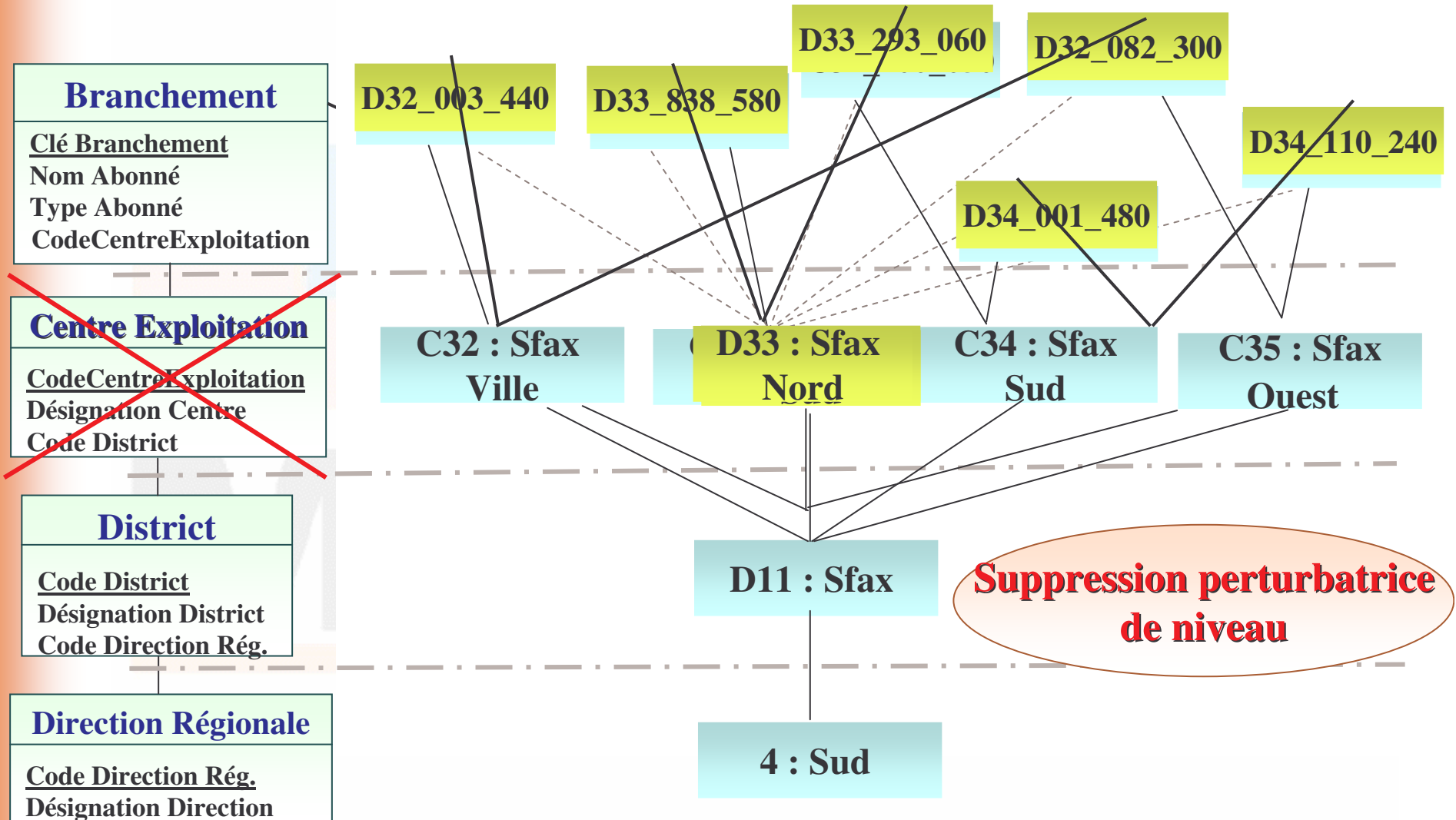
Reclassification perturbatrice de membre (2/2)





Évolution de nomenclature

Suppression perturbatrice de niveau (1/2)





Évolution de nomenclature

Suppression perturbatrice de niveau (2/2)

■ Solutions :

- Nouvelle version pour chaque membre du niveau « *Branchement* »
- Historisation des anciennes VC de ces membres
- Transformation entre versions de membres :
→ Table de correspondance

Ancienne clé Branchement	Nouvelle clé Branchement
C32_004_880	D32_003_440
C33_668_410	D33_838_580
C34_400_050	D33_293_060



Évolution de nomenclature

Bilan des évolutions perturbatrices

■ **Problèmes**

- Modification de la valeur de la clé d'un membre
- Réaffectation d'une valeur de la clé d'un membre
- Subdivision perturbatrice de membre
- Fusion perturbatrice de membre
- Reclassification perturbatrice de membre
- Suppression perturbatrice d'un niveau de hiérarchie

■ **Solution**

→ Système de gestion de DW multiversions



Plan

- 1- État de l'art sur l'évolution au niveau des DW
- 2- Évolution de nomenclature : Problèmes et solutions
- 3- Prototypage**
- 4- Conclusion et perspectives



Extrait des données du Cube Gestion Ventes Eaux_TMP (1/2)

*Avant la subdivision perturbatrice
de l'usage « 16 : Chimique et activités pétrolières »*

Usage_Eau_Tmp 16 : Chimique + pétrolière

- [-] Tous les Usages
 - [-] 01/01/1991
 - [+] 00 : Domestique
 - [-] 10 : Industrie
 - 11 : Alimentaire + huilerie
 - 12 : Mécanique
 - 13 : Textile
 - 14 : Mines
 - 15 : Matériaux de constructio
 - 16 : Chimique + pétrolière**
 - 17 : Autres
 - [+] 20 : Tourisme
 - [+] 30 : Collectif
 - [+] 50 : Divers
 - [+] 01/03/2001

Explorateur de cube - Gestion Ventes Eaux_Tmp

Categorie	Tous les Categories
Branchement_Tmp	Tous les Branchements
Usage_Eau_Tmp	16 : Chimique + pétrolière

- Annee	- Periodicite	Periode	Measures
	2001 Total		Quantite_Eau_Vendue 42 497
		Abonnes Mensuels Total	42 497
		janvier - 2001	20 347
		mars - 2001	
		avril - 2001	



Extrait des données du Cube Gestion Ventes Eaux_TMP (2/2)

*Après la subdivision perturbatrice
de l'usage « 16 : Chimique et activités pétrolières »*

Usage_Eau_Tmp 16 : Chimique

- Tous les Usages
- 01/01/1991

16: Chimique	janvier - 2001	19 271
+		
17: Activités pétrolière	janvier - 2001	1 076
<hr/>		
=		
16: Chimique et activités pétrolières	janvier - 2001	20 347
- 2001	mai - 2001	686



Plan

- 1- État de l'art sur l'évolution au niveau des DW
- 2- Évolution de nomenclature : Problèmes et solutions
- 3- Prototypage
- 4- Conclusion et perspectives



Conclusion

- Recensement des différents aspects d'évolution de nomenclature
- Notre approche :
 - Versionnement des éléments de DW
 - Définition de fonctions de transformation entre les versions de membres
 - Facteur poids
 - Tables de correspondance et requêtes de transformation



Perspectives

- Définition formelle des opérations d'évolution de nomenclature
- Validation formelle de ces opérations
- Proposition d'un langage d'interrogation OLAP multiversions (Mastère en cours)



Merci pour votre attention

MIR@CL