GESTION DE LA CONFIDENTIALITÉ DANS LES SGBD

MC Patricia Serrano Alvarado
LINA-Université de Nantes

PLAN

- Introduction
- DAC, MAC, RBAC
- La confidentialité dans Oracle 10g
- Confidentialité dans le WWW

Introduction

- Sécurité
  - Confidentialité
    - L'information ne doit pas être dévoilée aux utilisateurs non-autorisés. Uniquement les utilisateurs autorisés doivent modifier les données
  - Intégrité
    - Les modifications doivent respecter les contraintes d'intégrité des données ainsi que les propriétés transactionnelles
  - Disponibilité
    - Les données doivent être toujours disponibles pour les utilisateurs autorisés. Les données doivent être tolérantes aux pannes et l'accès doit être performant

CONFIDENTIALITÉ

- La confidentialité est le droit des individus de déterminer par eux-mêmes, quand, comment et quelle information sur eux peut être communiquée à d'autres individus

Alain Westin, Professor Emeritus of
Public Law and Government, Columbia University

- Pourquoi a-t-on besoin de gérer la confidentialité des données ?
  - Digitalisation de l'information sans précédentes
  - La quantité d'information double tous les 20 mois et la taille, ainsi que le nombre des bases de données, accroit encore plus
  - Internet facilite la collecte de données

TECHNIQUES POUR ASSURER LA CONFIDENTIALITÉ DES DONNÉES

- Control d'accès (RBAC, DAC, MAC)
- Vues
- Cryptographie
- Bases de données statistiques et résumées
- Mais aussi des législations et des directives on été proposées

QQS DIRECTIVES ET LÉGISLATIONS CONCERNANT LA CONFIDENTIALITÉ

- The OECD Privacy Guidelines in Europe
- The Canadian Privacy Act and the Personal Information Protection and Electronic Documents Act
- The Australian Privacy Amendment Act
- The Japanese Privacy Code
- The Health Insurance Portability and Accountability Act (HIPAA)
- Gramm-Leach-Bliley Consumer Privacy Rule
- CNIL in France

DIRECTIVE OECD

- OECD (Organisation for Economic Cooperation and Development, 1960)
- Membres : Union Européenne, USA, Canada, Mexique, Australie, Nouvelle Zélande, Suisse, Turquie, Japon, Corée, Islande, Norvège
- Basé à Paris
- Définie pour protéger le flux des données personnelles entre les pays
- Document qui a joué un rôle essentiel dans le développement des lois de nombreux pays

PRINCIPES DE L' OECD

- Purpose Specification
- Limited Retention

For personal information stored in the database, the purposes for which the information has been collected shall be associated with that information.

The purposes associated with personal information shall have consent of the donor of the personal information.

The personal information collected shall be limited to the minimum necessary for accomplishing the specified purposes.

The database shall run only those queries that are consistent with the purposes for which the

The personal information stored in the database shall not be communicated outside the database for purposes other than those for which there is consent from the donor of the information.

INTÉGRITÉ

- Une base de données est cohérente si elle respecte les **propriétés transactionnelles**
  - Atomicité. Toutes les modifications faites par une transaction sont validées ou aucune
  - Cohérence. Les contraintes d'intégrité doivent être vérifiées et respectées
  - Isolation. Chaque transaction s'exécute « comme » si elle est la seule à être exécutée sur le système
  - Durabilité. Une fois qu'une transaction est validée, toutes ses modifications doivent persister malgré des éventuelles pannes

INTÉGRITÉ

- Les propriétés transactionnelles sont assurées par le gestionnaire des transactions
  - Atomicité et durabilité. Protocole de validation (e.g., validation à deux phases 2PC)
  - Cohérence. Contraintes d'intégrité sémantiques
  - Isolation. Protocoles de control de concurrence (e.g., verrouillage à deux phases 2PL)

INTÉGRITÉ

- Contraintes d'intégrité
  - Règles qui représentent les propriétés d'une application
    - Contraintes structurelles : ex. clés primaires
    - Contraintes comportementales : ex. rang de valeurs pour un attribut,
- Une base de données doit assurer ces contraintes
- Le gestionnaire de la base de données assure les contraintes
  - Activité complexe et coûteuse lorsque la base est répartie mais aussi lorsque, lors des mises à jour, vérifier les contraintes demande l'accès à un grand nombre de données

GESTION DES CONTRAINTES D'INTÉGRITÉ

- Dans l'idéal la gestion de contraintes devrait
  - Limiter le nombre de contraintes à assurer
  - Réduire le nombre de données accédées lors d'une mise à jour
  - Avoir des stratégies pour détecter des inconsistances entre les contraintes afin d'éviter de défaire les mises à jour
  - Réaliser le moins de control en temps réel

DISPONIBILITÉ

- Disponibilité de données malgré les pannes et les nombreuses demandes d'accès
  - Accès efficace et rapide qui passe à l'échelle
    - Eviter les goulots d'étranglement
    - Faciliter l'accès avec de données plus « proches » de l'utilisateur
  - Type de pannes
    - Système. La mémoire centrale est perdue
    - Disque. La mémoire secondaire est endommagée
    - Transactionnelle. Erreur dans l'exécution des transactions

SOLUTIONS POUR LA DISPONIBILITÉ

- Accès efficace
  - Duplication de données
  - Systèmes pair-à-pair
- Données disponibles malgré les pannes
  - Disque : sauvegardes périodiques, point de reprise
    - Sauvegarde. Copie cohérente d'une base de données effectuée périodiquement alors que cette base est dans un état cohérent
    - Point de reprise. Etat d'avancement du système sauvegardé sur mémoires secondaires à partir duquel il est possible de repartir après un arrêt
  - Transactionnelle et système: journaux des images avant et après utilisés dans les procédures de reprise
    - Ex. Procédure de reprise à chaud. A partir d'un point de reprise, les transactions non validées sont défaites et les validées sont ré exécutées.

CONTROL D' ACCÈS

DAC (Discretionary Access Control)

MAC (Mandatory Access Control)

RBAC (Role-Based Access Control)

CONTROL D'ACCÈS DANS LES SGBD

- Discretionary Access Control (DAC)
  - Basé sur des privilèges ou droits d'accès
  - Gestion de privilèges par les sujets (processus des utilisateurs)
- Mandatory Access Control (MAC)
  - Basé sur des politiques du système
  - Gestion de politiques par le système (pas par les utilisateurs)
- Role-based Access Control (RBAC)
  - Basé sur des rôles
  - Gestion de rôles et leur privilèges par une unité centrale

DAC (Discretionary Access Control)

- Consiste à attribuer aux sujets des privilèges des opérations sur les objets et à les vérifier lors de l'accès
- Un privilège permet à un sujet l'accès à une donnée (e.g., table, vue, objet, procédure)
  - SELECT : privilège pour lire toutes le colonnes de la donnée ainsi que les colonnes insérées plus tard avec ALTER TABLE
  - INSERT/UPDATE : privilège pour insérer/modifier des lignes
  - DELETE : privilège pour supprimer des lignes
  - REFERENCES : privilège pour définir des clés étrangères (référencier d'autres tables ou vues)
  - EXECUTE : privilège pour exécuter une fonction/ procédure

DAC

- Gestion de privilèges
  - Les privilèges peuvent être partagés et révoqués (commandes GRANT et REVOKE)
  - Si un privilège a été donné avec GRANT OPTION, le nouveau sujet peut également partager le privilège
  - Le sujet qui crée une table automatiquement a tous les privilèges sur celle-ci
  - La commande REVOKE peut être utilisée pour annuler un privilège ou l'option GRANT OPTION
  - REVOKE peut être utilisé avec CASCADE

EXEMPLES

- Considérez les tables suivantes et les utilisateurs Joe, Michel, Lea, Yuppy, Bill, Eric, Guppy, Art, Bob, etc.

Sailors	Boats	Reservs
• sid: integer • sname: string • rating: integer • age: real	• bid: integer • bname: string • color: string	• sid: integer • bid: integer • day: date

EXEMPLES

- L'utilisateur Joe crée les tables (Boats, Reservs et Sailors) et exécute les commandes :
  - GRANT INSERT, DELETE ON Reservs TO Yuppy WITH GRANT OPTION
  - GRANT SELECT ON Reservs TO Michel
  - GRANT SELECT ON Sailors TO Michel WITH GRANT OPTION
  - GRANT INSERT ON Sailors TO Michel
  - GRANT REFERENCES (bid) ON Boats TO Bill

- Michel exécute :
  - CREATE VIEW YoungSailors (sid, age, rating)
    AS SELECT sid, age, rating
    FROM Sailors
    WHERE age<18
  - GRANT SELECT ON YoungSailors TO Eric, Guppy

- Peut Michel faire la même chose avec Reservs ?

EXEMPLES

- Si Joe ajoute la colonne « address » à Sailors, peut Michel insérer des valeurs à la nouvelle colonne ?
- Si Joe ajoute la colonne « employe » à Reservs, peut Yuppy insérer des valeurs à la nouvelle colonne ?
- Peut Bill créer une table Reservs de la manière suivante ?
  - CREATE TABLE Reservs (sid INTEGER,
    bid INTEGER,
    day DATE,
    PRIMARY KEY (bid,day),
    FOREIGN KEY (sid) REFERENCES Sailors,
    FOREIGN KEY (bid) REFERENCES Boats)
- Qu'est-ce qui se passe si Joe exécute :
  - REVOKE REFERENCES ON Boats FROM Bill

EXEMPLES

- Considérer les commandes suivantes :
- Joe (qui a crée Sailors) exécute :
  - GRANT SELECT ON Sailors TO Art WITH GRANT OPTION
- Art exécute :
  - GRANT SELECT ON Sailors TO Bob WITH GRANT OPTION
- Qu'est-ce qui se passe si Joe exécute :
  - REVOKE SELECT ON Sailors FROM Art CASCADE

  ou
  - REVOKE GRANT OPTION FOR SELECT ON Sailors FROM Art CASCADE

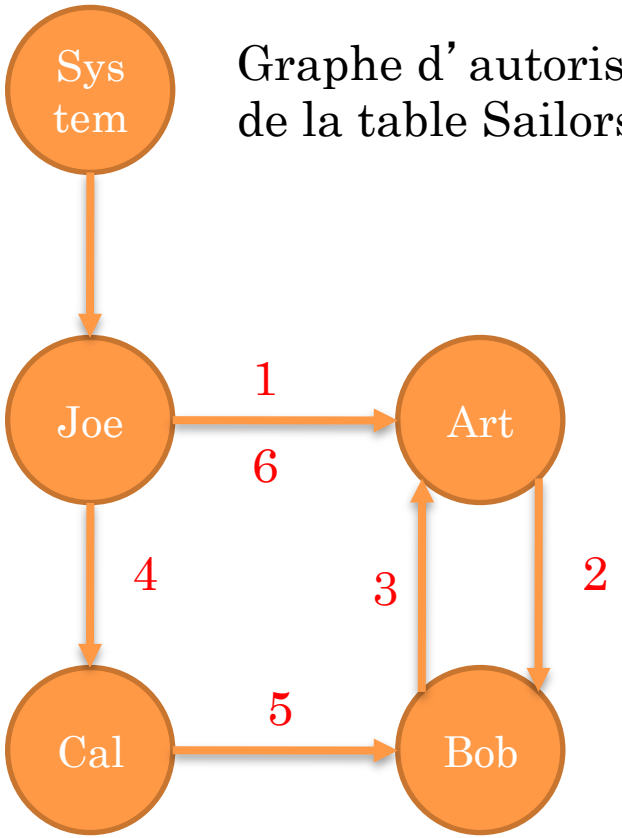
- Avec l'option RESTRICT la commande REVOKE est rejetée si en plus de l'utilisateur spécifié d'autres utilisateurs vont voir ses privilèges affectés

GRAPHES D'AUTORISATION

- Les effets des commandes GRANT et REVOKE peuvent être décrits dans un graphe d'autorisations
  - Un graphe concerne une seule donnée
  - Les nœuds sont les utilisateurs
  - Les arcs indiquent comment les privilèges sont accordés

EXEMPLE DE GRAPHE D'AUTORISATION

- Joe exécute :
  1. GRANT SELECT ON Sailors TO Art WITH GRANT OPTION
- Art exécute :
  2. GRANT SELECT ON Sailors TO Bob WITH GRANT OPTION
- Bob exécute
  3. GRANT SELECT ON Sailors TO Art WITH GRANT OPTION
- Joe exécute
  4. GRANT SELECT ON Sailors TO Cal WITH GRANT OPTION
- Cal exécute
  5. GRANT SELECT ON Sailors TO Bob WITH GRANT OPTION
- Joe exécute
  6. REVOKE SELECT ON Sailors FROM Art CASCADE



Graphe d'autorisation de la table Sailors

PRIVILÈGES SUR LES VUES

- Gestion de privilèges sur les vues
  - L'utilisateur qui crée une vue doit préalablement avoir au moins le privilège de SELECT sur les données concernées (tables/vues utilisées pour la création de la vue)
  - Si la vue peut être mise à jour, l'utilisateur doit également avoir le privilège de INSERT, UPDATE et/ou DELETE sur les données de base
  - Du même avec GRANT OPTION

PRIVILÈGES SUR LES VUES

- Les privilèges du créateur d'une vue dépendent de ses privilèges sur les tables de base
  - Si le créateur perd un privilège obtenu avec l'option GRANT OPTION les utilisateurs auxquels il a donné le privilège le perdent également
  - Une vue peut être supprimée si son créateur perd le privilège de sélection d'une des tables de base de la vue
  - Si le créateur obtient plus de privilèges sur les tables de base d'une vue alors il obtient automatiquement les mêmes privilèges sur la vue

VUES ET CONFIDENTIALITÉ - CAS D'ÉTUDE

Sailors	Boats	Reservs
• sid: integer • sname: string • rating: integer • age: real	• bid: integer • bname: string • color: string	• sid: integer • bid: integer • day: date

- L'utilisateur Joe crée les tables (Boats, Reservs et Sailors). Qu'est-ce qu'il doit faire pour permettre :
  - La lecture et l'écriture des marins jeunes à Michel
  - La lecture et l'écriture des marins âgés à Bill
  - La lecture de son information à chaque marin
- Joe doit exécuter :
  - Création des vues YoungSailors et OldSailors
  - Donner les droits de lecture et d'écriture à Michel et Bill
  - Création des vues de chaque marin et donner les droits de lecture à chacun (ex dans un déclencheur)

MAC (Mandatory Access Control)

- Le control basé sur DAC a une limitation principal qui est la possibilité d'avoir de chevaux de Troie
  - L'utilisateur A a accès aux données R et S
  - L'utilisateur B a accès uniquement à S
  - Si B trafique une application utilisée par A et il écrit les données de R sur S alors B peut lire des données non autorisées sans violer les règles de DAC
- MAC a été proposé pour palier ce type de problèmes
- MAC consiste à associer un niveau d'autorisation aux sujets et aux données

MAC

- Le modèle MAC de Bell-LaPadula
  - Objets
    - Tables, vues, lignes, colonnes, etc.
  - Sujets
    - Utilisateurs, programmes
  - Classes de sécurité
    - Organisées dans un ordre partiel de la plus sécurisée à la moins sécurisée (i.e., *lattice* ou treillis)
    - Ex. 4 clases : *top secret* (TS), *secret* (S), *confidential* (C) et *unclassified* (U) où TS>S>C>U où TS est le plus sensible et U le moins
  - Autorisations (*clearance*)
  - Chaque objet appartient à une classe de sécurité
  - Chaque sujet a une autorisation pour une clase de sécurité

MAC

Les deux règles de Bell-LaPadula

1. Propriété de sécurité simple ou *no read up*
   - Un sujet S peut lire l'objet O uniquement si

     classe(S)>= classe(O)

     Ex. un sujet avec autorisation S ne peut pas lire une donnée avec autorisation TS

2. Propriété *- ou *no write down*
   - Un sujet S peut écrire sur un objet O si

     classe(S)<= classe(O)

     Ex. un sujet avec autorisation S peut écrire sur les objets avec autorisation S ou TS

CLASSIFICATION MULTI-NIVEAUX

- Le MAC dans le modèle relationnel et les tables « multi-niveaux »
- Granularité de la classification (classes de sécurité)
  - Tables
  - Lignes
  - Colonnes
  - Valeur d'une colonne
- Utilisateurs avec différentes autorisations perçoivent différemment une même table
- Voici une table avec une classification par ligne :

bid (PK)	bname	color	classe de sécurité
101	Salsa	Red	S
102	Pinto	Brown	C

La table Boats

CLASSIFICATION MULTI-NIVEAUX

- Table avec une classification par colonne
  - Chaque colonne a en plus son niveau de sécurité
  - Les niveaux de sécurité incrémentent la taille de la base

sid (PK)	SL1	sname	SL2	rating	SL3	age	SL4
AA	**C**	Tim	C	Prof essional	C	45	C
BB	**C**	Bill	C	NULL	S	NULL	S
CC	**S**	Michel	S	Amateur	S	50	S

La table Sailors

  - Un utilisateur perçoit uniquement les attributs qui lui sont autorisés, pour les autres des valeurs nulles seront insérées

POLY-INSTANCIATION

- Problème de la poly-instanciation avec MAC

bid (PK)	bname	color	classe de sécurité
101	Salsa	Red	S
102	Pinto	Brown	C

La table Boats

- Si l'utilisateur U avec autorisation C veut insérer le tuple <101, Picante, Scarlet, C> soit
  - L'insertion est faite et 2 lignes auront comme clé 101
  - L'insertion est rejetée et S déduit qu'il existe un autre bateau avec comme clé 101 ce qui viole la confidentialité des objets de classe supérieure
  - Solution : les classes de sécurité font partie de la clé

RBAC (Role Based Access Control)

- DAC et MAC ne sont pas applicables à des systèmes complexes avec des centaines/milliers d'utilisateurs
  - Inconvénients de DAC
    - Approche basée sur l'attribution de droits par sujet (utilisateur) donc difficulté pour gérer la dynamicité des privilèges
      - Sorties du système (entreprise)
      - Modification des droits (changement de responsabilités)
  - Inconvénients de MAC
    - Approche basée sur une organisation hiérarchique des privilèges
      - Difficilement une grande organisation peut organiser ses privilèges hiérarchiquement

RBAC (Role Based Access Control)

- Proposé pour des systèmes très grands avec un important nombre d'utilisateurs et de données
- Contrôle d'accès à base de rôles
  - Modèle dans lequel les décisions d'accès dépendent du rôle auquel l'utilisateur est attaché
- Concepts de base
  - Utilisateur (pas de processus)
  - Privilège. Concerne un droit (opération) quelconque sur une donnée, plusieurs droits sur une donnée ou plusieurs droits sur plusieurs données
  - Rôles. Entité déterminant une activité d'entreprise (comptable, chef de projet, chef de département, cassier, etc.)
- Deux approches : modèles ANSI et *role graphe*

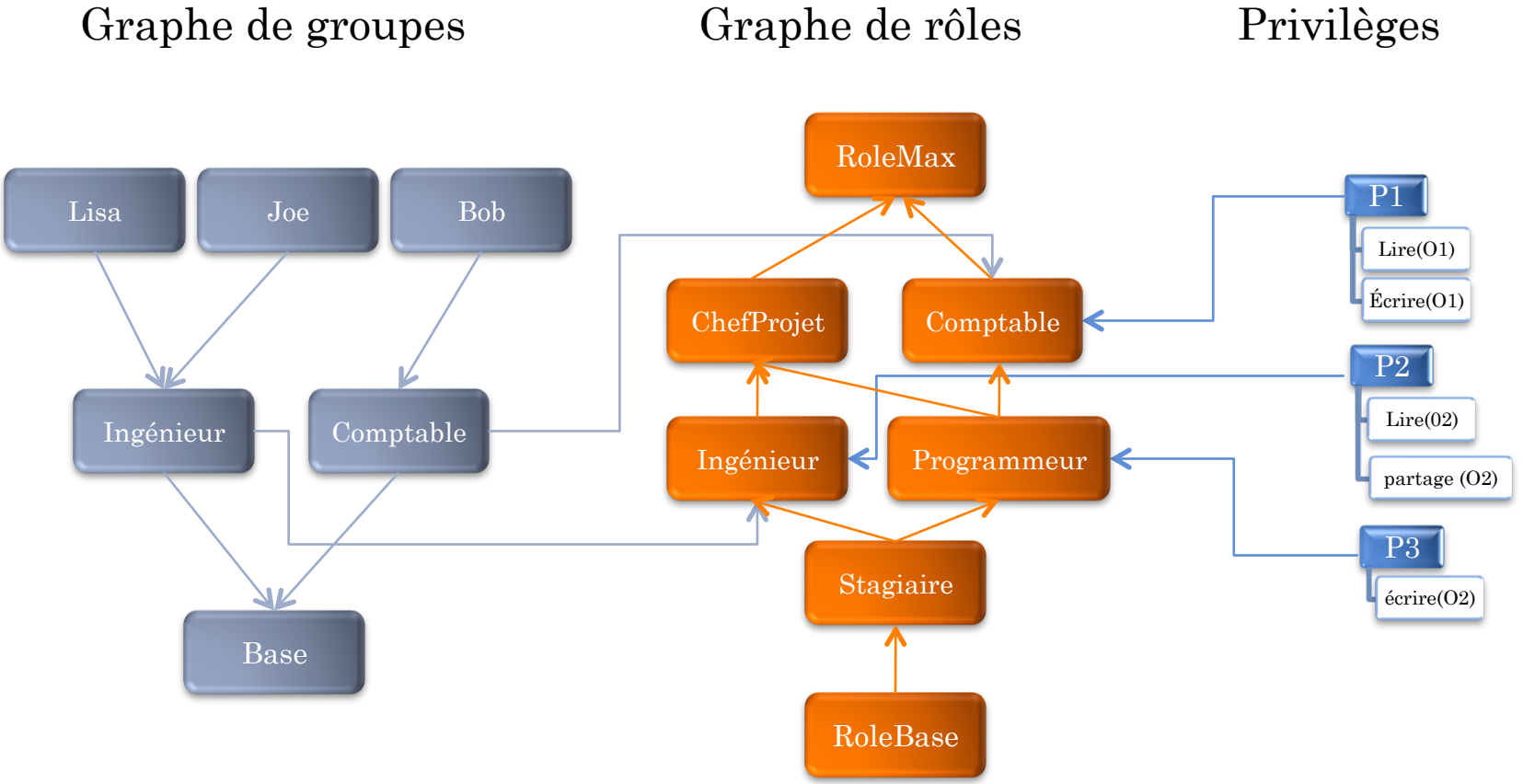
Les rôles

- Sont définis par une autorité centrale
- Sont identifiés par un nom unique
- Peuvent être organisés en une hiérarchie ou un graphe. Attention aux
  - Cycles (redondance)
  - Conflits d'intérêt
- Sont attribués aux utilisateurs ou groups d'utilisateurs
  - Mapping rôle-utilisateur
  - Plusieurs rôles peuvent être attribués à un utilisateur
  - Plusieurs utilisateurs peuvent partager plusieurs rôles
- Facilité de gestion rôle-utilisateur

LES GROUPES D'UTILISATEURS ET LES RÔLES

- Groupes d'utilisateurs
  - Facilite la gestion des utilisateurs
  - Les membres du group peuvent changer fréquemment
  - La gestion des groupes peut se faire par le gestionnaire des ressources humaines
- Rôles
  - les rôles sont définis avant la mise en place du système
  - Les privilèges attribués aux rôles varient très peu
  - La gestion des rôles doit se faire par un utilisateur de confiance

COMPOSANTS DU MODÈLE BASÉ SUR LES RÔLES



PROPRIÉTÉS DU GRAPHE DE RÔLES

- Héritage
  - r1→r2 indique que r1 est junior de r2 et que les privilèges de r1 sont hérités à r2
  - privilèges(r1) ⊆ privilèges(r2)
- Privilèges directs
  - Privilèges attribués directement au rôle par l'administrateur du graphe
- Privilèges effectifs
  - Privilèges directs plus privilèges hérités
- Graphe acyclique
- Pas de redondance de privilèges
  - Un arc r1→r2 doit être ajouté si privilèges(r1) ⊂ privilèges(r2)

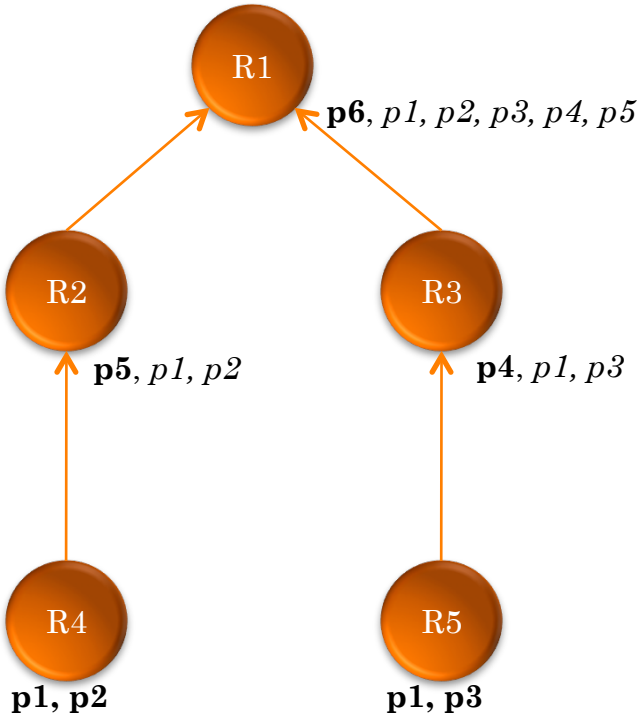
GESTION DU GRAPHE DE RÔLES

- roleAddition
  - Si pas de cycle, le rôle est ajouté
- permitionAddition
  - Un nouveau privilège est ajouté à un rôle
- permitionDeletion
  - Suppression d'un privilège d'un rôle
- roleDeletion
  - Suppression d'un rôle
- edgeInsertion
  - Insertion d'un arc si pas de cycle
- edgeDeletion
  - Suppression d'un arc si pas de cycle

- Dans toutes les fonctions, si nécessaire, les privilèges/arcs sont réorganisés pour éviter les redondances

EXEMPLE 1 DE GESTION DE GRAPHE DE RÔLES

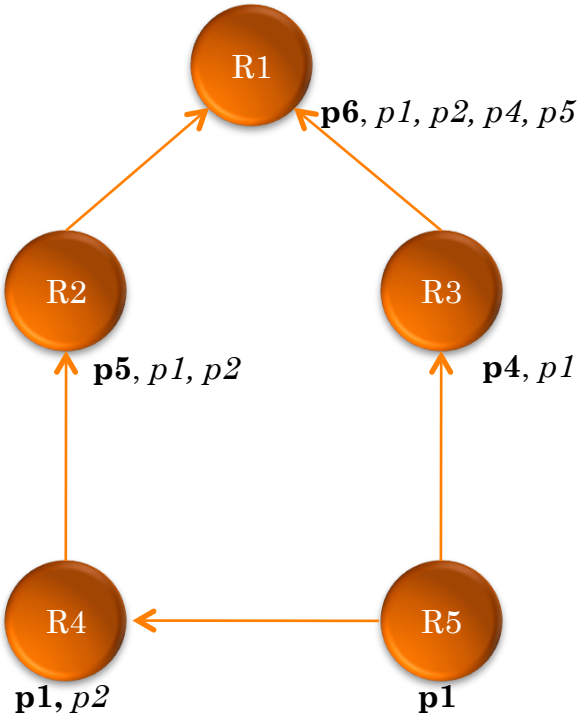
Graphe A



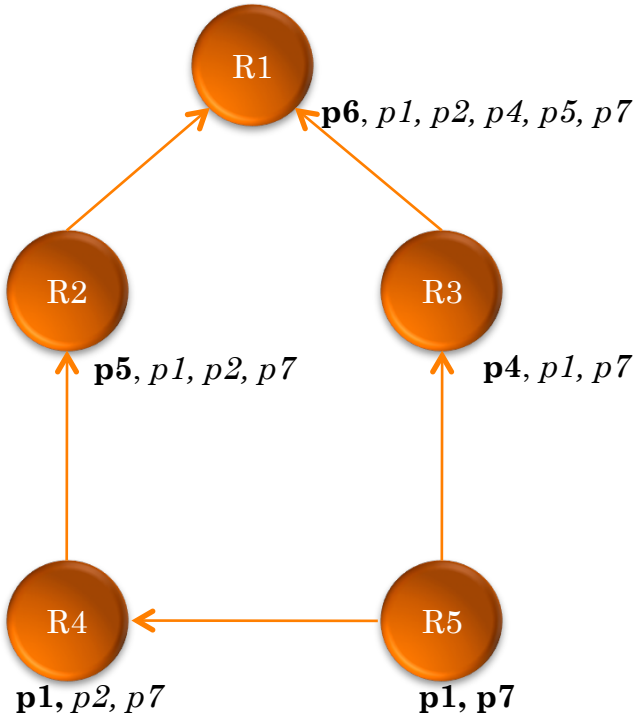
Pas de cycle
Pas de redondance de privilèges i.e., ¬(privilèges(r1) ⊂ privilèges(r2))

EXEMPLE 2 DE GESTION DE GRAPHE DE RÔLES

Graphe A, la suite

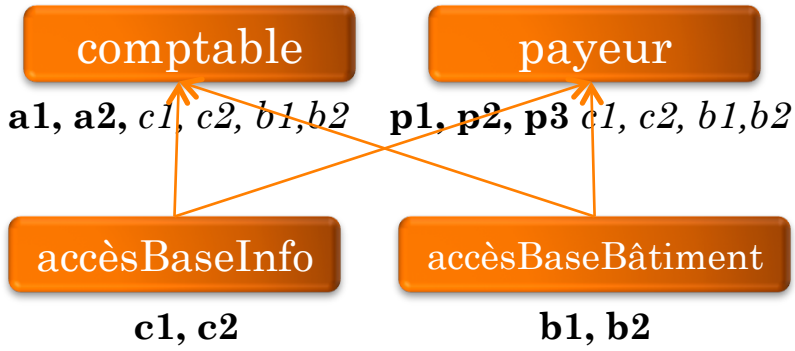


Graphe A avec p7 ajouté à R5

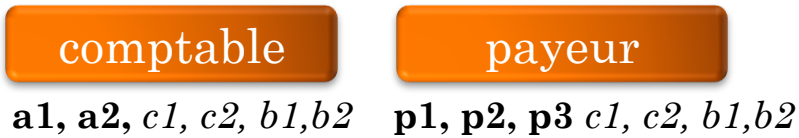


CONSEILS POUR DÉFINIR UNE HIÉRARCHIE DE RÔLES

- Ne pas faire de la hiérarchie de rôles un miroir de la hiérarchie de l'entreprise
- Surveiller les privilèges des rôles
- Considérer les groupes d'utilisateurs
- Considérer de rôles abstraits



Avec des rôles abstraits



Sans des rôles abstraits

Conclusion

- Le control d'accès est vital pour assurer la confidentialité des données
- DAC, MAC, RBAC : les trois grandes approches
  - Chacun a ses avantages et désavantages
  - RBAC se révèle le plus adapté aux systèmes actuels

BIBLIOGRAPHIE

- Security, Privacy, and Trust in Data Management. Milan Petkovic, Willem Jonker Eds. 2007.
- Database Management Systems. Raghu Ramakrishnan, Johannes Gehrke. International 3rd Edition, 2003.
- Internet notamment wikipedia ;)