

# Administration des SGBD

Sécurisation / Scripting

M1 Informatique
Damien Ploix

#### Scripting

- Toutes les opérations réalisées via les interfaces IDE sont réalisables via des commandes (scripts)
  - SQL pour les opérations dans la base
  - Shell / PowerShell pour les opérations depuis les OS
    - Le cours utilisera Shell (bash).
    - Pour PowerShell, lien vers Oracle Data Provider pour .NET : <a href="https://www.oracle.com/database/technologies/appdev/dotnet/odp.html">https://www.oracle.com/database/technologies/appdev/dotnet/odp.html</a>

# Types de changements

#### Autour de la base

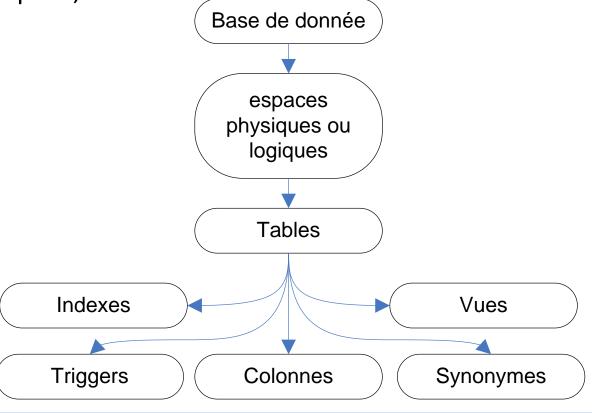
- Modification du matériel,
- Modification/cohérence de l'analyse (MCD, MLD, MPD),
- Modification de l'application utilisatrice...

## Modification d'une base de donnée

Toutes les modifications ne sont pas possibles ... sans destruction/reconstruction...

- Ajout de colonne dans une table,
- Modification de structure de tablespace,

-



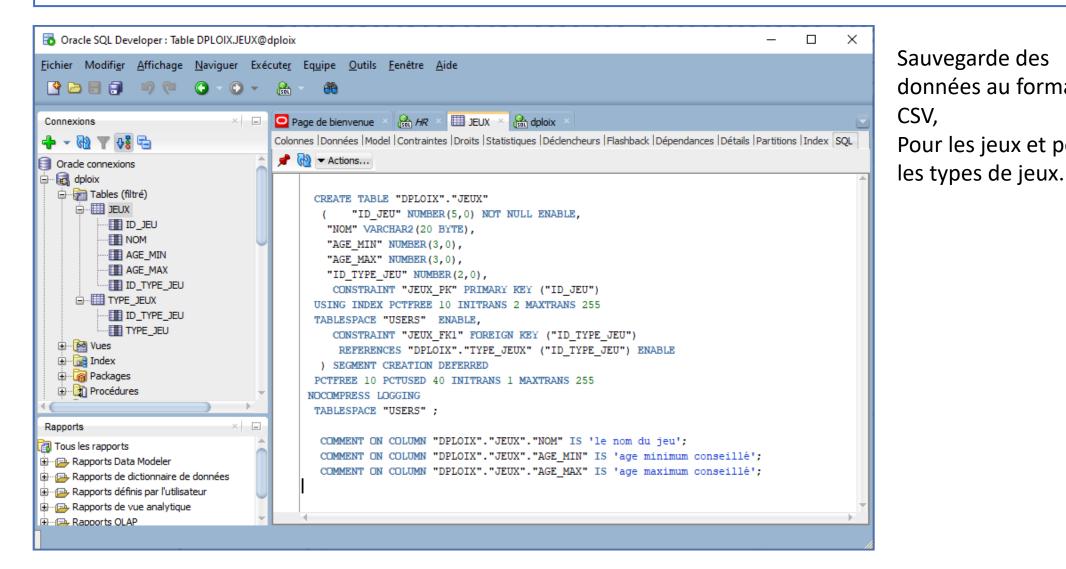
#### Modification d'une base de données Ce que ALTER ne peut généralement pas...

- Pour la base :
  - Changer son nom,
  - Déplacer un objet d'une base à un autre,
- Pour sa structure physique :
  - Changer le nombre de partitions ou de fichiers de données d'un tablespace,
- Pour les tables :
  - Déplacer d'un tablespace à un autre,
  - Réarranger l'ordre des colonnes,
  - Effacer une colonne, en changer le type ou la taille,
  - Ajouter une colonne non nulle ou entre deux existantes,
- Pour les vues :
  - Ajouter ou effacer une colonne, changer la commande SELECT,
- Pour les indexes :
  - L'unicité, le clustering, l'ordre
- Modifier le contenu d'un trigger
- Changer une clé de Hash.

#### Modification d'une base Exemple de l'ajout d'une colonne entre deux existantes

- 1. Retrouver la définition actuelle de la table,
- 2. Retrouver la définition de toute vue qui la référence,
- 3. Retrouver la définition de tout index qui la référence,
- 4. Retrouver la définition de tout trigger sur la table,
- 5. Capturer toutes les contraintes appliquées sur la table et mesurer l'impact de l'effacement de la table,
- 6. Retrouver tous les grants sur la table,
- 7. Retrouver tous les programmes qui la référence,
- 8. Sauvegarder les données de la base,
- 9. Effacer la table, ce qui inclus également tous les indexes, trigger et vues liés,
- 10. Recréer la table avec la nouvelle colonne,
- 11. Recharger les données,
- 12. Recréer les contraintes (clés étrangères) effacées,
- 13. Recréer les triggers, indexes et vues,
- 14. Recréer les grants,
- 15. Examiner en détail toute requête applicative qui référence la table.

#### Récupération de la définition des tables



Sauvegarde des données au format CSV. Pour les jeux et pour

# Scripting : ne pas réinventer la roue !

- Base des scripts sql : <a href="https://oracle-base.com/dba/scripts">https://oracle-base.com/dba/scripts</a>
- Parcours des scripts...

#### Récupération des données d'une table

- Extraction « CSV » (valeurs séparées par des virgules).
- Script disponible : csv.sql :

- Va :
  - Créer un package définissant des fonctions d'extraction
  - Demande le droit de créer un objet 'DIRECTORY'

#### Récupération des données d'une table

#### Cas d'usages :

- Lié à une application et limité à l'utilisateur applicatif correspondant :
  - Le package est définit dans le périmètre de l'utilisateur,
  - Il doit avoir le droit de créer l'objet Oracle 'Directory',
  - Similaire en restriction d'accès à l'utilisateur applicatif.
- Générique à une/plusieurs instances et lié à un outillage d'administration
  - Définition d'un utilisateur dédié à extraction de données
  - Ayant le droit de créer l'objet Oracle 'Directory' et des packages
  - Ayant le droit en lecture des données sur le/les/tous les utilisateurs applicatifs.
  - Similaire en restriction d'accès à ce qui est nécessaire à la sauvegarde.

## Récupération des données d'une table

- Exemple : cas d'usage utilisateur HR pour l'extraction des données de la table EMPLOYEES
- 1) Modification de l'utilisateur HR ajout du droit : GRANT CREATE ANY DIRECTORY TO "HR" ;
- 2) Integration du package "sql.sql":

```
$ cd <répertoire_où_est_csv.sql>
```

\$ sqlplus HR/HR@pdb1 @csv.sql

SQL> exit

3) Extraction de la table « employees » :

```
$ cd <répertoire_où_extraire_les_données>
```

\$ sqlplus HR/<mot\_de\_passe\_hr>@pdb1

SQL> create or replace directory ext\_dir as './';

SQL> exec csv.generate('EXT\_DIR', 'generate.csv', p\_query => 'SELECT \* FROM employees');

## Modification/rechargement de la table

- Commandes : drop, puis recréation de la table avec les nouveaux champs.
- Recharge des données :
  - Si elle a été effacée, en premier lieu la table type\_jeux
  - Puis la table jeux

#### Utilisation de sqlldr : chargement de données dans Oracle

- Définition d'un fichier de contrôle « jeux.ctl » qui indique les opérations :

```
-- on charge des données
LOAD DATA
infile jeux.csv
                                 -- qui sont dans le fichier jeux.csv
badfile jeux.bad
                                 -- contiendra les lignes rejetées (format ou erreur Oracle)
Discardfile jeux.dis
                                 -- contiendra les lignes qui ne matchent pas la clause WHEN
Logfile jeux.log
                                 -- log de l'import
Replace
                                 -- on remplace si existant
into table jeux
                                 -- dans la table jeux
When nom = "toto"
                                -- clause de filtrage des lignes acceptées
fields terminated by ';'
                                -- le séparateur de champ est le ;
trailing nullcols
                                 -- une colonne sans valeur est traité en 'null'
( id jeu,
                                 -- en premier l'id du jeu
                                 -- puis son nom
nom,
age min,
                                 -- age min
                                 -- age max
age max,
id type jeu CHAR "(select id type jeu from type jeux where type jeu like :id type jeu ) "
```

Et on lance l'intégration des données :

```
$ sqlldr dploix/dploix@dploix jeux
```

#### Rappel sur le Shell/Script

#### • Comment:

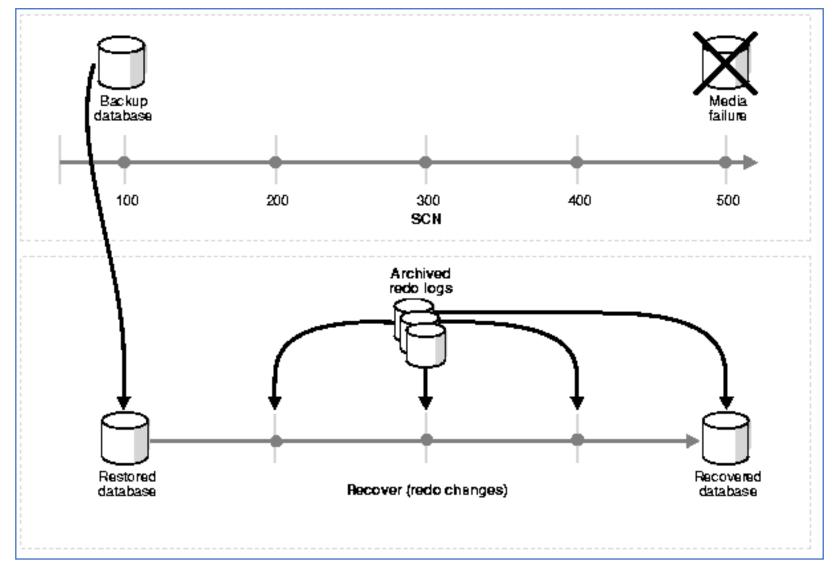
- Programmer l'exécution d'une tâche à intervalle régulier ?
- Faire une boucle sur un ensemble de fichiers ?
- Tester une valeur?
- Écrire dans un fichier ?
- Lire le contenu d'un fichier et le mettre dans une variable ?
- Faire une boucle à partir du contenu de la 3<sup>ème</sup> colonne d'un fichier CSV ?

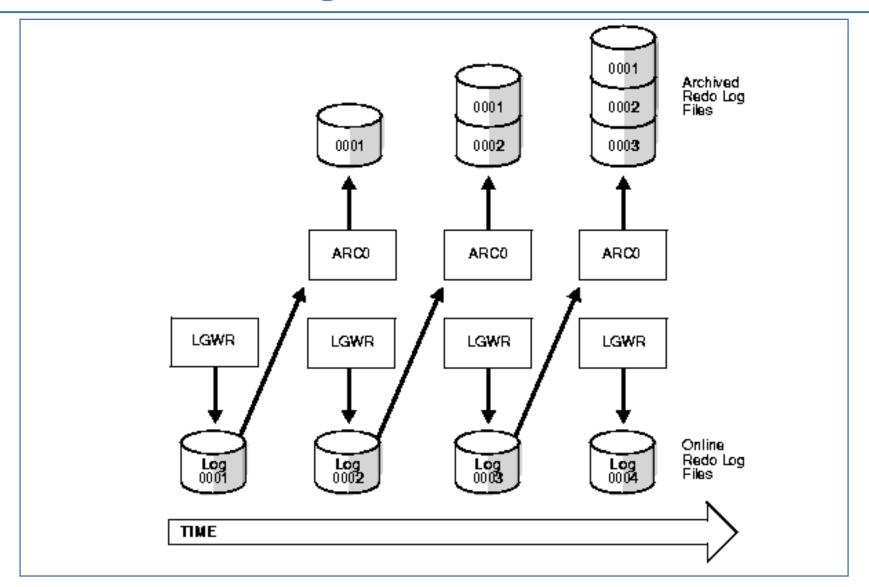
#### Problèmes sur une base?

- Disfonctionnement d'un média physique
  - Un disque casse,
  - Une CPU casse, ...
- Erreur de manipulation
  - Un utilisateur effectue une opération malheureuse,
- Corruption de l'état de l'instance
  - Suite à un arrêt brutal
- Erreur de paramétrage
  - Toutes les extensions (extens) sont utilisées
- Cryptolockage
  - Suite à une attaque réussie
- Problèmes réseaux...

#### Problèmes sur une base!

- Chaque type de problème amène sa solution préventive :
  - Matériel : Sauvegarde + Dataguard + Cluster
  - Manipulation : Sauvegarde + instances de développement
  - Corruption de l'état de l'instance : Sauvegarde + gestion de l'instance par le DBA
  - Erreur de paramétrage : actions ponctuelles
  - Cryptolockage :
    - Capacité à reconstruire entièrement le serveur (OS, logiciels, ...) et sauvegarde.





#### Méthodologie générale de la sauvegarde

- À chaud (on-line) :
  - Minimise la durée de la sauvegarde
  - Peut provoquer des lock selon les DBMS
  - Nécessiter la mise en œuvre de méthodologie spécifiques à chaque DBMS
- À froid (off-line) :
  - Minimise la durée de la restauration
  - Nécessite un arrêt de la base
  - Est effectuée par copie de l'ensemble des éléments physiques de la base
  - Est possible avec tous les DBMS

- Stratégie de sauvegarde :
  - À chaud (aujourd'hui systématique),
  - Sur VTL:
    - En prévention de corruption des données et/ou de perte de média de stockage/serveur.
  - Externalisée (bande/média physique externe) :
    - En prévention d'attaque en cryptolockage.
    - Après la sauvegarde le média doit être physiquement déconnecté.
- On joue alors sur le mode et la durée de rétention :
  - Redondance : combien de copie à un temps t donné,
  - Durée : pendant combien de temps une sauvegarde doit-elle être conservée.
  - Sur la base d'une étude du besoin *métier* et des exigences de performance/durée/reconstructibilité.
- Problème actuel:
  - Des bases (bigdata) deviennent trop volumineuses pour être sauvegardées
  - La sauvegarde est alors une étude métier *des données qu'on ne peut pas perdre* sauf à mettre en risque la vie de l'entreprise.

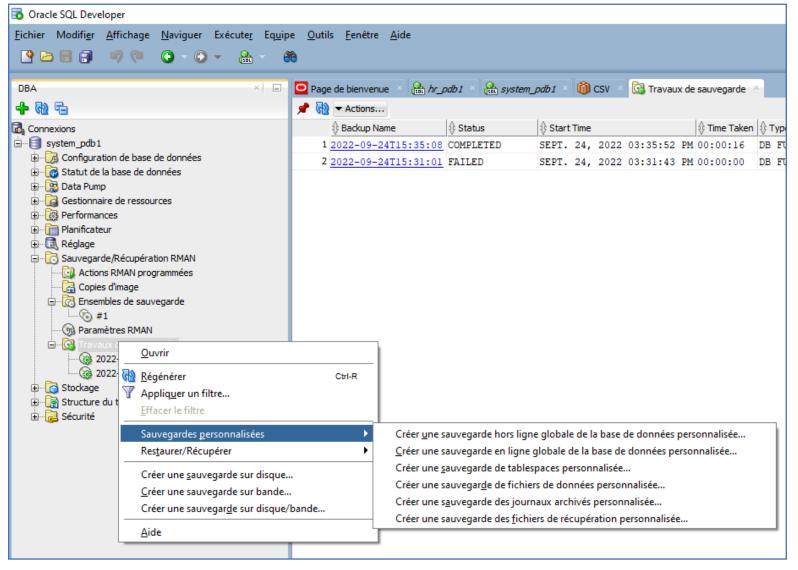
- Nombre de bandes nécessaires ?
  - Méthode GFS : Grand-père, Père, Fils :
    - Fils = Une sauvegarde différentielle ou FULL journalière
    - Père = Une sauvegarde FULL par semaine
    - Grand-père = Une sauvegarde FULL (père) par mois
  - Rotation des bandes (sur 1 ans) :
    - Un jeu de bande est utilisé pour le fils (4, 5 ou 6),
    - Un jeu de bande est utilisé pour le père (4),
    - Un jeu de bandes est utilisée pour le grand-père (12)

#### Selon le problème :

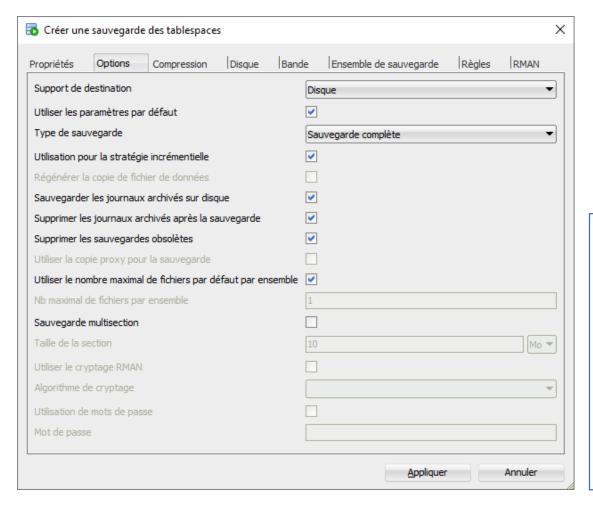
- Revenir au dernier état disponible
  - Restauration de l'ensemble des sauvegardes disponibles (FULL + Archives Log)
- Revenir au dernier état stable
  - Restauration du dernier FULL disponible
  - Analyse des Archives Log pour ne pas reproduire l'erreur.

- Des solutions intégrées existent :
  - Intégrée (RMAN) pour Oracle,
  - Intégrée pour SQL Server,
  - Plus à la main pour MySQL,
- Solutions de sauvegardes « outillées » en entreprise :
  - Outils du marcher OnPremice
  - Solution générique adaptée à un ensemble générique de type de bases via du scripting
  - Via des services câblés dans le Cloud

#### Exemple: RMAN



#### Sauvegarde d'un tablespace

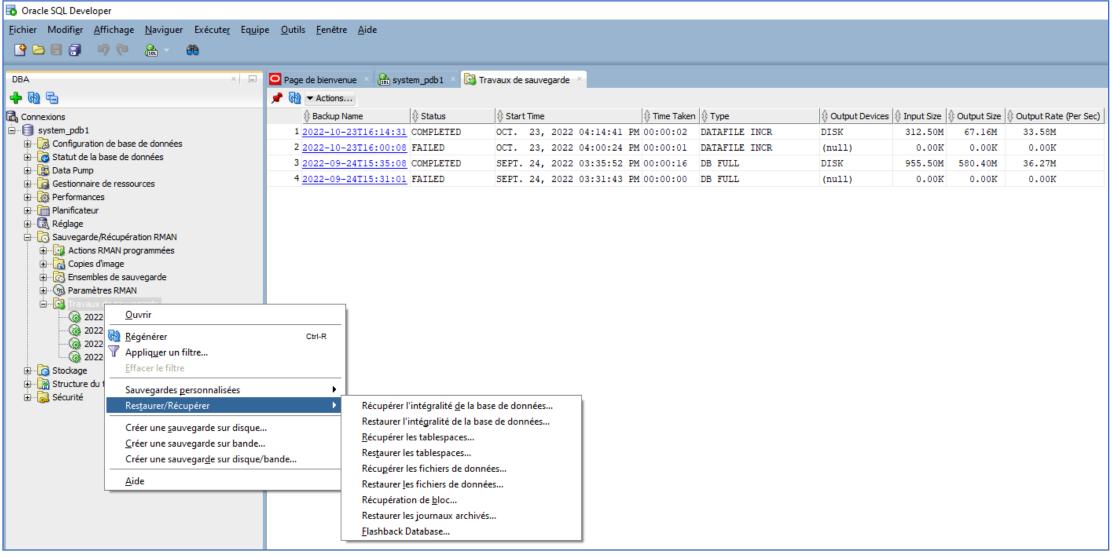




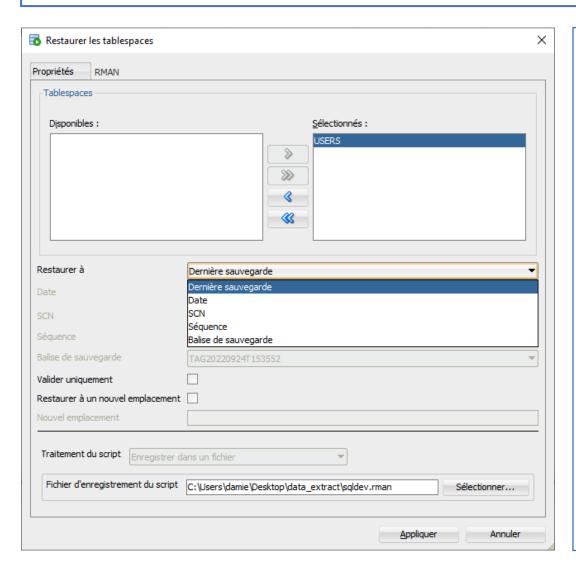
durée de la sauvegarde, il n'y aura pas de commit des archiveLog.

24

#### Restauration d'un tablespace



#### Restauration d'un tablespace



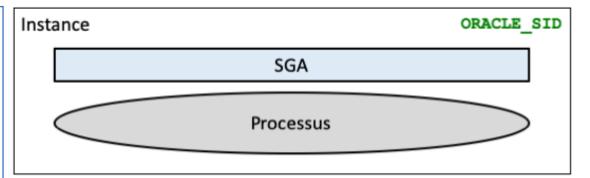
RUN {
SQL 'ALTER TABLESPACE "USERS" OFFLINE IMMEDIATE';
RESTORE TABLESPACE 'USERS' FROM TAG TAG20220924T153552;
}

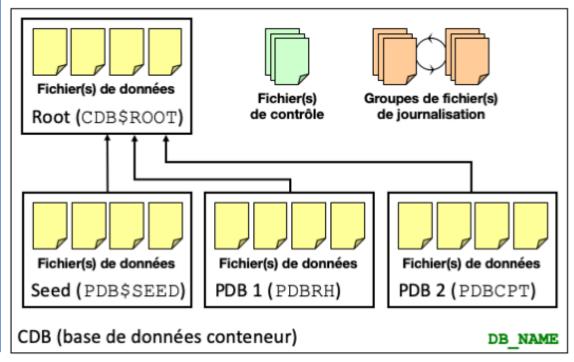
→ La restauration implique l'arrêt du travail sur le tablespace et donc l'arrêt des traitements applicatifs en cours.

La mise en place du mode archivelog ne permet pas de palier mais il permet de « choisir » le moment (date, séquence, tag, ...) correspondant aux données restaurées.

#### Conteneurisation des db : CDB et PDB

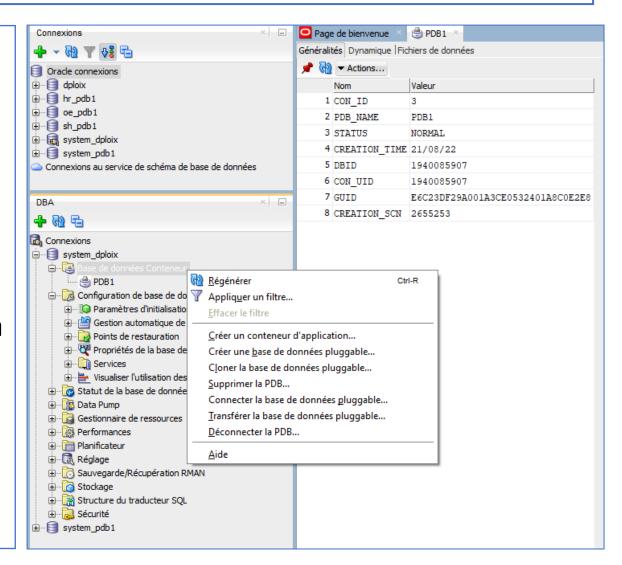
- Une base conteneur gère N bases
   « Pluggable » :
  - Une d'amorçage (Seed),
  - N correspondant aux applications.
- Objectif:
  - Faciliter la migration Cloud :
    - Plug / unplug et migration facilité.
  - Faciliter la sécurité :
    - Ensemble d'utilisateurs applicatifs au sein d'une PDB.





#### Administration PDB

- Création / mouvement des BD : proche des fonctions de déplacement des VM d'un VMWare.
- Gestion de différentiée des ressources :
  - Création/gestion de modèle de ressource dans la CDB affectées à une PDB
  - Création/gestion de modèles de ressource dans les PDB pour des utilisateurs en fonctions d'usage.

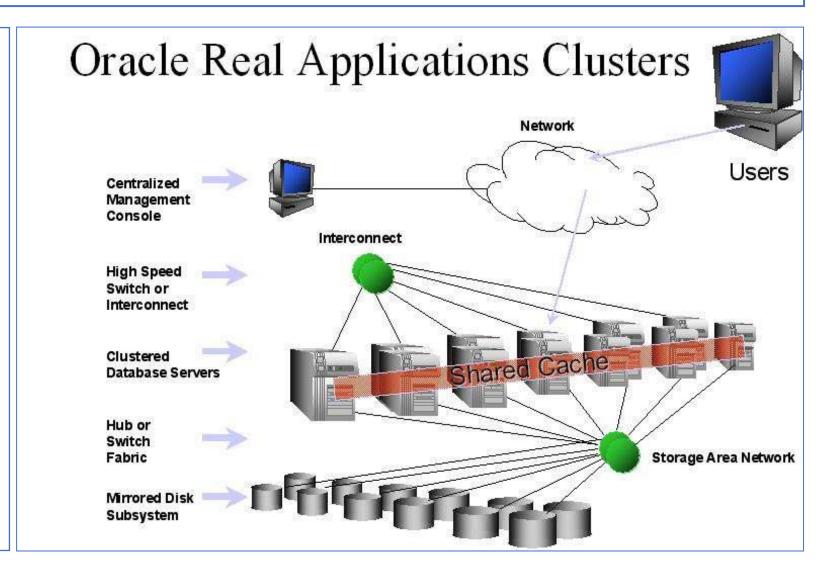


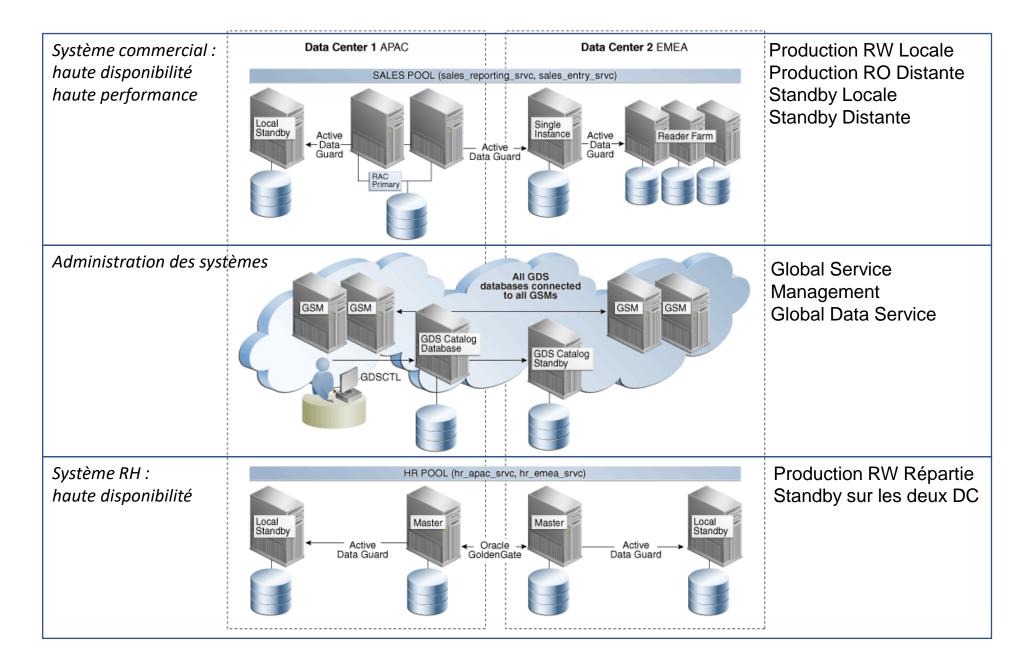
## Sécurisation par clustering

- L'idée est d'avoir :
  - Une sécurisation de la puissance de calcul
  - Une sécurisation des données
  - Les deux
- Les solutions dépendent des DBMS mais se regroupent sous le terme de Cluster
  - Un élément dirige les requêtes vers un ensemble de ressources les plus disponibles

# Cluster actif/actif

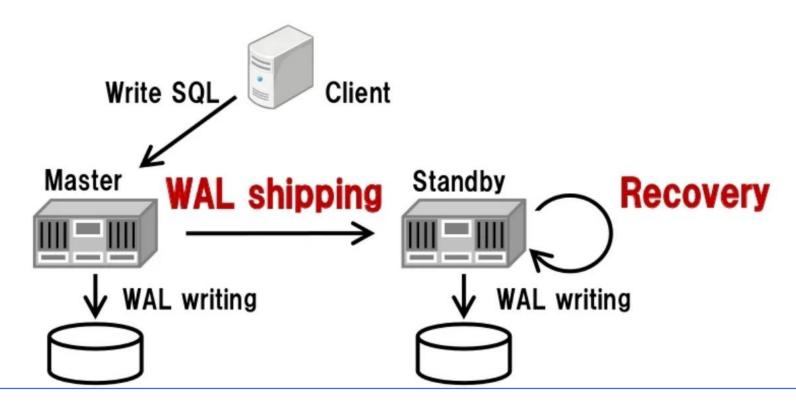
 Les données ne sont stockées qu'à un seul endroit. L'accès au disque (partagé) est réalisé via des FS spécialisés (ASM, VxSF, ...) permettant l'accès à un même fichier à n systèmes actifs. Les modifications physiques sont réalisées par un seul nœud et communiqué aux autres nœuds.





# Réplication PostgreSQL

- La réplication (master → standby) peut adresser toute l'instance ou une/des table(s)
- Et se réaliser en mode synchrone / asynchrone.

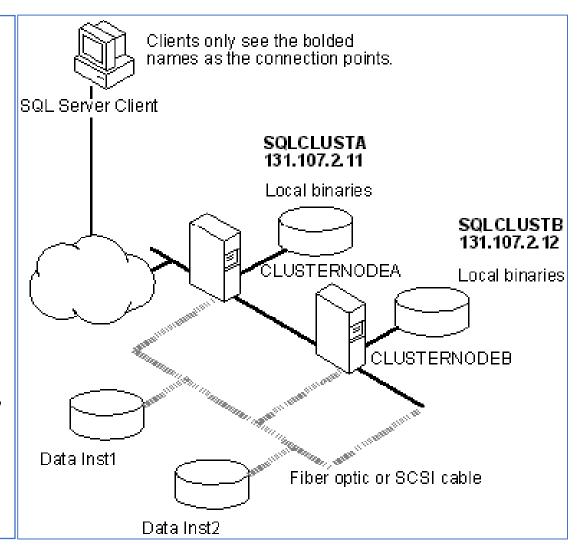


# Cluster actif/actif

- Fonctionne via une VIP (Virtual IP) gérée par le gestionnaire de clusteurisation (RAC / répartiteur de charge) pour adresser les nœuds via une seule adresse.
- Travaille en mode « mémoire globale partagée »
  - Induit un temps de communication important dans le cas de modification des données,
  - Pour RAC : Induit une utilisation des compteurs « par plage » qui ne garanti plus la continuité « historique » des numéros fournis (sauf à réduire la plage à 1) : chaque nœuds a « une plage » qu'il peut utiliser et se synchronise avec les autres lorsqu'il a tout utilisé.
- Permet la « spécialisation » des nœuds en fonction de leur architecture sous jacente (châssis optimisé pour le calcul, pour les transations, ...).

#### Cluster actif/passif: Failover Cluster SQL Server

- Un seul nœud est actif à un moment donné. En cas de défaillance du nœud, un autre nœud devient actif sur les mêmes données.
- Demande une couche OS de clusterisation gérant le nœud qui « voit » et peut « modifier » les données.
- Existe aussi sous Unix avec des outils comme (par exemple) Veritas, PowerHA, ... qui gèrent au niveau OS les « bascules » des stockages et des logiciels.



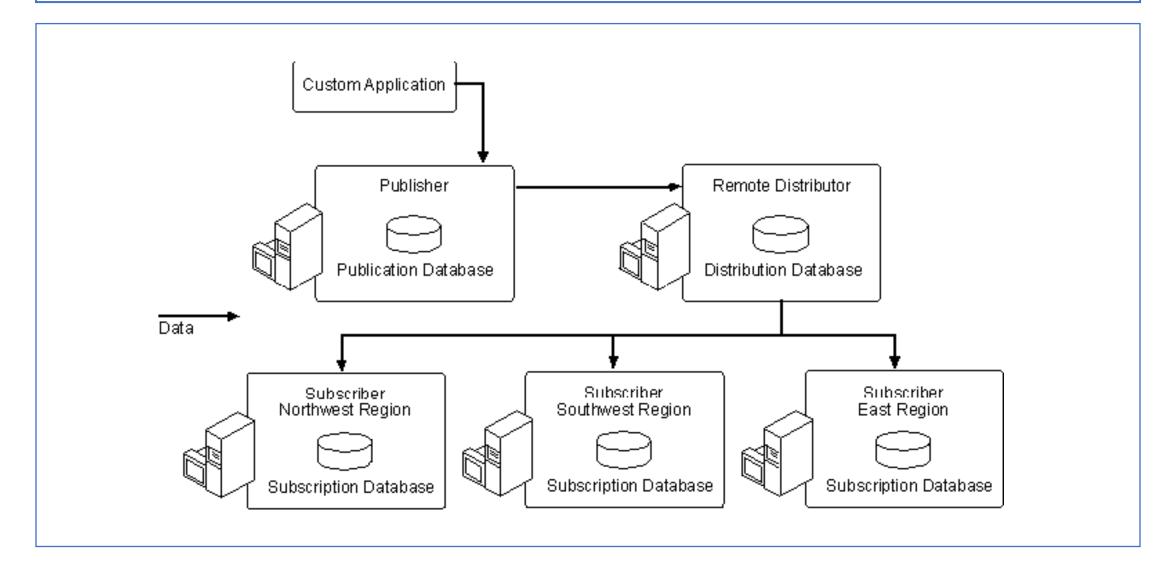
#### Architecture de réplication des données

- Gère la distribution en temps réel / différé des modifications sur les bases.
- Se base sur le principe de « nœud maître » (dans lequel sera effectué les modification) et de « nœuds esclaves » qui sont informés des modifications
- Permet la mise en place de nœuds en lecture seule mais aussi la bascule de responsabilité maître / esclave entre les nœuds
  - Pour gérer des incidents (perte de base)
    - Attention, le retour au nominal est coûteux car il demande l'arrêt de la base maître le temps de reconstruire l'esclave!
  - Pour gérer la proximité (service devant être rendu « au plus prêt » des utilisateurs/administrateurs)

#### Dataguard Oracle

- Réplication des données d'un serveur principal vers un (ou plusieurs) serveur secondaires via un processus LGWR (écriture / transmission des log). Chaque serveur ayant un stockage en propre et pouvant être situé dans des datacentres différents.
  - **Protection garantie** : la transaction n'est validée que lorsqu'une base secondaire à inscrit l'opération dans son log (PROTECTION),
  - **Protection instantanée** : la transaction est validée sans attente et transmise immédiatement (AVAILABILITY),
  - Protection différée : la transaction est transmise plus tard (PERFORMANCE).

# Réplication SQL Server (1/2)



# Réplication SQL Serveur (2/2)

- Mode snapshot :
  - Une image des données est distribuée
- Mode transaction :
  - Une image initiale est distribuée,
  - Puis uniquement les modifications.
- Mode merge :
  - Une image est préparée,
  - Le transfert et l'application peut être différée.

#### Scripting des opérations DBA

- Source par Oracle de scripts : <a href="https://oracle-base.com/dba/scripts">https://oracle-base.com/dba/scripts</a>
- Exemple : table\_defs.sql →
- Résultat : liste les colonnes d'une table
- Téléchargement du script sur le serveur Exécution dans sqlplus :

```
$ sqlplus dploix/dploix@dploix
SOL*Plus: Release 19.0.0.0.0 - Production on Sun Nov
1 16:50:09 2020
Version 19.3.0.0.0
Copyright (c) 1982, 2019, Oracle. All rights
reserved.
Heure de la derniere connexion reussie : Dim. Nov.
01 2020 16:35:58 +01:00
Connecte a :
Oracle Database 19c Enterprise Edition Release
19.0.0.0.0 - Production
Version 19.3.0.0.0
SQL> @table defs.sql jeux
JEUX
1 ID JEU
                           NUMBER
NOT NULL
2 NOM
                                       VARCHAR2
                                                  20
3 AGE MIN
                           NUMBER
                                       3
4 AGE MAX
                           NUMBER
5 ID TYPE JEU
                NUMBER
SOL>
```

```
-- File Name : https://oracle-base.com/dba/miscellaneous/table defs.sql
-- Description : Lists the column definitions for the specified table.
-- Call Syntax : @table defs (tablee-name or all)
-- Last Modified: 24/09/2003
COLUMN column id FORMAT 99
COLUMN data type FORMAT A10
COLUMN nullable FORMAT A8
COLUMN size FORMAT A6
BREAK ON table name SKIP 2
SET PAGESIZE 0
SET LINESIZE 200
SET TRIMOUT ON
SET TRIMSPOOL ON
SET VERIFY OFF
SELECT table name,
      column id,
      column name,
      data type,
        WHEN data type IN ('VARCHAR2', 'CHAR') THEN TO CHAR(data length)
        WHEN data scale IS NULL OR data scale = 0 THEN TO CHAR(data precision)
        END) "SIZE",
      DECODE(nullable, 'Y', '', 'NOT NULL') nullable
FROM user tab columns
WHERE table name = DECODE(UPPER('&1'), 'ALL', table name, UPPER('&1'))
ORDER BY table name, column id;
SET PAGESIZE 14
SET LINESIZE 80
```