#### Administration des bases de données

Mehdi Haddad mehdi.haddad@u-pec.fr

2017 - 2018

#### Plan du cours

- 1. Archivage des fichier de journalisation
- 2. Recovery Manager
- 3. Sauvegarde
- 4. Récupération

#### Sauvegarde et récupération

- Assurer la sécurité des données est une des tâches principales de l'administrateur.
- Cette sécurité est assurée par :
  - la mise en œuvre d'une protection des fichiers sensibles de la base :
    - fichiers de contrôle;
    - fichiers de journalisation.
  - ▶ la mise en place d'une stratégie de sauvegarde/restauration :
    - adaptée aux contraintes de l'entreprise;
    - testée et documentée.
- ▶ La protection des fichiers de contrôle et des fichiers de journalisation s'effectue par multiplexage (voir TP2).

#### Sauvegarde et récupération

Les questions à se poser pour définir la stratégie sont les suivantes :

- ▶ Est-il acceptable de perdre des données?
- ▶ Est-il possible d'arrêter périodiquement la base?
- ► Est-il possible de réaliser une sauvegarde complète de la base pendant l'arrêt?

La réponse à la question "est-il acceptable de perdre des données?" est rarement "oui". Si exceptionnellement, la réponse est "oui", il faut déterminer jusqu'à quelle limite : 1 jour, 2 jours, 1 semaine?

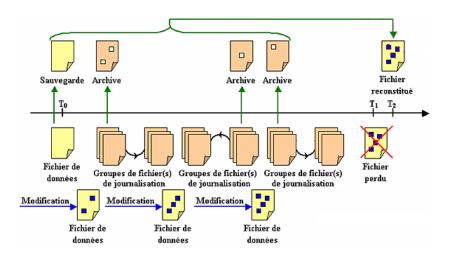
#### Sauvegarde et récupération

Il est également nécessaire de déterminer la nature de l'activité sur la base :

- ▶ Les données sont-elles mises à jour quotidiennement par les utilisateurs? C'est typiquement le cas dans une application transactionnelle.
- Les données sont-elles mises à jour périodiquement (toutes les nuits, toutes les semaines) et simplement consultées dans la journée? C'est typiquement le cas avec une application décisionnelle.

- ▶ Les fichiers de journalisation constituent un journal des modifications apportées à la base.
- ▶ Ils sont organisés en groupes écrits de manière circulaire : les informations sauvegardées sont donc par défaut périodiquement écrasées.
- ▶ Ces fichiers de journalisation peuvent être réappliqués à une sauvegarde de fichier de données

- ▶ Il faut rejouer toutes les modifications survenues entre la sauvegarde et un incident ayant endommagé le fichier (restauration de média), à condition d'avoir conservé tous les fichiers de journalisation.
- Ceci est possible en faisant fonctionner la base de données en mode ARCHIVELOG.
- ► Ce mode de fonctionnement permet de garantir zéro perte de données en cas d'incident sur un fichier de données.



- À un instant T0, une sauvegarde d'un fichier de données est réalisée.
- ▶ Après T0, l'activité de mise à jour se poursuit, générant des entrées dans les fichiers de journalisation.
- L'archivage étant activé, les fichiers de journalisation pleins sont archivés.

- ▶ À l'instant T1, un incident se produit et le fichier de données est perdu.
- ▶ La récupération du fichier de données consiste à prendre la dernière sauvegarde du fichier (qui ne contient évidemment pas les modifications effectuées depuis).
- ▶ Puis à appliquer sur cette sauvegarde les fichiers de journalisation archivés (qui eux, contiennent la trace des modifications apportées depuis la dernière sauvegarde).
- ▶ Afin de ramener le fichier de données dans l'état où il se trouvait juste avant l'incident (pour être plus précis, dans l'état de la dernière transaction validée).

- ▶ Activer l'archivage des fichiers de journalisation s'effectue en mettant la base de données dans le mode ARCHIVELOG : ce mode permet de garantir qu'un groupe de fichiers de journalisation ne sera pas réutilisé tant qu'il n'a pas été archivé.
- ▶ Placer la base de données en mode ARCHIVELOG démarre automatiquement deux processus d'archivage (ARC0 et ARC1) lors de l'ouverture de la base de données; dans les versions précédentes, il fallait le faire explicitement. Par contre, il est toujours opportun, de positionner certains paramètres d'initialisation qui concernent les processus d'archivage.
- ▶ La base de données peut être créée immédiatement en mode ARCHIVELOG. Généralement, la base de données est créée en mode NOARCHIVELOG puis passée en ARCHIVELOG.

#### Trouver des informations sur l'archivage

- ➤ Dans SQL\*Plus, vous pouvez utiliser la commande ARCHIVE LOG LIST (dans une connexion SYSDBA) pour obtenir des informations sur l'archivage;
- Plusieurs vues du dictionnaire de données permettent d'obtenir des informations sur l'archivage :
  - V\$DATABASE : mode de fonctionnement de la base de données (colonne LOG MODE);
  - ▶ V\$LOG : statut des groupes vis-à-vis de l'archivage (colonne ARCHIVED);
  - V\$ARCHIVED\_LOG: informations sur les fichiers de journalisation archivés;
  - V\$ARCHIVE\_DEST: informations sur les destinations d'archivage.

#### V\$ARCHIVED\_LOG

Les colonnes intéressantes de la vue V\$ARCHIVED\_LOG sont les suivantes :

- ▶ RECID : Identifiant de l'enregistrement.
- ▶ NAME : Chemin complet de l'archive.
- ➤ SEQUENCE#: Numéro de séquence du fichier de journalisation correspondant.
- ► FIRST\_CHANGE# : Plus petit numéro SCN écrit dans l'archive.
- ▶ FIRST\_TIME : Date et heure du plus petit numéro SCN.
- ► NEXT\_CHANGE# : Plus petit numéro SCN écrit dans l'archive suivante.
- ► NEXT\_TIME : Date et heure du plus petit numéro SCN écrit dans l'archive suivante.
- ► COMPLETION\_TIME : Date et heure de l'archivage.

#### V\$ARCHIVE\_DEST

Les colonnes intéressantes de la vue V\$ARCHIVE\_DEST sont les suivantes :

- ▶ DEST NAME : Nom de la destination.
- ▶ DESTINATION : Chemin de la destination.
- ► STATUS : Statut de la destination (VALID, ERROR, etc.).
- ► ERROR : Message d'erreur (en cas d'erreur).

#### Stratégies de sauvegarde disponibles

- ▶ Une sauvegarde peut être cohérente ou incohérente.
- ▶ Une sauvegarde peut être complète, partielle ou incrémentale.

#### Sauvegarde cohérente

- Une sauvegarde cohérente est une sauvegarde de la totalité de la base de données après un arrêt propre de la base de données (pas après un SHUTDOWN ABORT ou un arrêt anormal de l'instance)
- ▶ Ce type de sauvegarde est aussi souvent appelé "sauvegarde base fermée".
- Après un arrêt propre de la base de données, toutes les modifications ont été écrites dans les fichiers de données qui sont bien synchrones.
- ▶ Une base de données restaurée à partir d'une sauvegarde cohérente peut être ouverte immédiatement : il est inutile d'appliquer les fichiers de journalisation.
- ▶ C'est le seul mode de sauvegarde disponible lorsque la base de données fonctionne en mode NOARCHIVELOG.

#### Sauvegarde incohérente

- Une sauvegarde incohérente est une sauvegarde effectuée alors que la base de données est ouverte et que l'activité de mise à jour se poursuit pendant la sauvegarde.
- ▶ Ce type de sauvegarde est aussi souvent appelé "sauvegarde base ouverte".
- ▶ Les fichiers sauvegardés ne sont pas synchrones du point de vue des modifications enregistrées.
- ▶ Lorsqu'une base de données est restaurée à partir d'une sauvegarde incohérente, il faut appliquer les fichiers de journalisation pour rendre les fichiers cohérents.
- ▶ Les sauvegardes incohérentes ne sont possibles que lorsque la base de données fonctionne en mode ARCHIVELOG.

#### Sauvegarde complète, partielle ou incrémentale

- ▶ Une sauvegarde complète est une sauvegarde de la totalité de la base de données.
- ▶ Une sauvegarde partielle est une sauvegarde incluant uniquement une partie de la base de données. Les sauvegardes partielles sont forcément incohérentes entre elles. Pour qu'elles soient exploitables en restauration (ce qui est normalement l'objectif), il faut que la base de données fonctionne en mode ARCHIVELOG.
- ▶ Une sauvegarde incrémentale est une sauvegarde qui ne contient que les blocs modifiés depuis la dernière sauvegarde; une sauvegarde incrémentale peut être complète ou partielle.

### Quelle stratégie pour le mode de fonctionnement de la base?

		Pertes de données acceptables	
		Oui	Non
Sauvegarde base fermée possible	Oui	ARCHIVELOG NOARCHIVELOG	ARCHIVELOG
	Non	ARCHIVELOG	ARCHIVELOG

### Quelle stratégie pour le mode de fonctionnement de la base?

- ▶ Le mode ARCHIVELOG est obligatoire si au moins une des contraintes suivantes existe :
  - ▶ Aucune perte de donnée n'est autorisée.
  - La base de données ne peut pas être fermée pour être sauvegardée.
- ▶ Le mode NOARCHIVELOG est possible si :
  - ▶ Des pertes de données sont acceptables.
  - La base de données peut être fermée pour être sauvegardée en totalité.
- ▶ Un autre avantage du mode ARCHIVELOG est que la base de données peut rester ouverte lorsqu'un incident survient sur un fichier de données qui n'appartient ni au tablespace SYSTEM, ni au tablespace UNDO actif.

#### Quelle stratégie pour la sauvegarde?

- ▶ La première règle est de réaliser des sauvegardes fréquentes (au minimum tous les jours) et de conserver plusieurs cycles de sauvegarde (en cas de problème avec une sauvegarde).
- ➤ Si la base de données fonctionne en mode ARCHIVELOG, vous pouvez réaliser des sauvegardes bases ouvertes; il n'y a pas de raison de s'en priver.
- ▶ Si la durée de sauvegarde et la taille des sauvegardes ne posent pas de problème (même en conservant plusieurs sauvegardes), vous pouvez réaliser systématiquement (tous les jours) des sauvegardes complètes.

#### Quelle stratégie pour la sauvegarde?

- Si la durée de sauvegarde et/ou la taille des sauvegardes posent un problème, vous pouvez réaliser des sauvegardes incrémentales et/ou des sauvegardes partielles.
- Dans le cas de sauvegardes partielles, vous devez simplement être très rigoureux dans le suivi et veiller à tout sauvegarder sur un cycle complet de sauvegardes partielles.
- ▶ Il est important de réaliser des sauvegardes très fréquemment pour pouvoir procéder à une restauration en un temps raisonnable : partir d'une sauvegarde datant d'un mois et réappliquer tous les fichiers de journalisation archivés depuis un mois risque de se révéler très long si la base de données est activement mise à jour.

#### Présentation du Recovery Manager (RMAN)

- ▶ RMAN est un outil ligne de commande qui permet de réaliser des sauvegardes et des récupérations d'une base de données appelée base de données cible (target database).
- RMAN utilise un référentiel (repository) pour stocker des informations sur sa configuration, les sauvegardes réalisées, la structure de la base cible, les fichiers de journalisation archivés, etc.
- ➤ Ce référentiel est toujours stocké dans le fichier de contrôle de la base cible.
- La durée de conservation des informations dans le fichier de contrôle est déterminée par le paramètre d'initialisation.

### Présentation du Recovery Manager (RMAN)

- ▶ Une sauvegarde RMAN peut se faire sous la forme d'une copie image (image copy) ou d'un jeu de sauvegarde (backup set).
- ▶ Une copie image est une copie à l'identique du fichier (analogue à une copie par une commande du système d'exploitation).
- ▶ Un jeu de sauvegarde contient un ou plusieurs fichiers sauvegardés.
- ▶ Chaque fichier d'un jeu est appelé élément de sauvegarde (backup piece). Par défaut, un jeu de sauvegarde comprend un seul élément de sauvegarde, mais il est possible de limiter la taille de ces éléments.
- ▶ Dans ce cas, un jeu de sauvegarde peut contenir plusieurs éléments de sauvegarde si la taille totale de la sauvegarde est supérieure à la limite. Le jeu de sauvegarde a un format propriétaire RMAN.

#### Sauvegarde

- ► La commande BACKUP permet d'effectuer une sauvegarde.
- ▶ Pour que cette commande fonctionne, il faut que la base de données soit montée ou ouverte car RMAN a besoin d'accéder au fichier de contrôle de la base cible, notamment pour y enregistrer l'existence de la sauvegarde.
- ▶ Les sauvegardes base ouverte de la totalité de la base de données, de fichiers de données ou de tablespace ne sont possibles que si la base de données fonctionne en mode ARCHIVELOG.
- ▶ Si la base de données fonctionne en mode NOARCHIVELOG, pour effectuer une sauvegarde de ce type, il faut au préalable arrêter la base de données (proprement) puis la monter.

#### Sauvegarde

- ▶ RMAN peut sauvegarder des fichiers de données, des fichiers de contrôle, des fichiers de journalisation archivés, le fichier de paramètres serveur ou des éléments de sauvegarde (d'une sauvegarde précédente).
- Une sauvegarde RMAN peut être réalisée sous la forme d'une copie image (image copy) ou d'un jeu de sauvegarde (backup set). Par défaut, la sauvegarde s'effectue dans un jeu de sauvegarde.
- ▶ Lorsque RMAN effectue une sauvegarde de fichiers de données dans un jeu de sauvegarde, il ne sauvegarde pas les blocs jamais utilisés des fichiers, ce qui permet de gagner de la place.

#### Sauvegarde

- ► En complément, il est possible de compresser le jeu de sauvegarde.
- ▶ Cela ralentit légèrement la sauvegarde, consomme un peu de CPU, mais diminue la taille de la sauvegarde de manière importante (typiquement, division par 5).
- ▶ Ces deux fonctionnalités ne sont pas disponibles dans le cas d'une copie image (copie bit à bit du fichier d'origine).

#### Syntaxe BACKUP

#### BACKUP [comment] quoi [option]

Dans la commande BACKUP, la seule clause obligatoire est la clause quoi qui indique ce qu'il faut sauvegarder. Toutes les autres clauses sont optionnelles et ont des valeurs par défaut. La clause comment peut prendre une ou plusieurs des valeurs suivantes :

- ► INCREMENTAL LEVEL n [CUMULATIVE] Indique que la sauvegarde est une sauvegarde incrémentale.
- ► VALIDATE
  Indique simplement de vérifier que la sauvegarde peut être réalisée (teste la présence des fichiers et leur non-corruption). Cette option est équivalente à l'utilisation de la commande VALIDATE.
- ▶ AS COPY ou AS [COMPRESSED] BACKUPSET Indique s'il faut faire une sauvegarde sous la forme d'une copie image ou d'un jeu de sauvegarde, éventuellement compressé.

#### Syntaxe BACKUP

La clause quoi peut prendre une ou plusieurs des valeurs suivantes :

- ► DATABASE Sauvegarde de la totalité de la base de données.
- ► TABLESPACE cible Sauvegarde d'un ou plusieurs tablespaces.
- DATAFILE cible
   Sauvegarde d'un ou plusieurs fichiers de données.
- ► CURRENT CONTROLFILE Sauvegarde du fichier de contrôle courant.
- SPFILE
   Sauvegarde du fichier de paramètres serveur.
- ➤ ARCHIVELOG cible Sauvegarde des fichiers de journalisation archivés.

#### Syntaxe BACKUP

La clause option peut prendre une des valeurs suivantes :

- ► INCLUDE CURRENT CONTROLFILE
  Inclure le fichier de contrôle courant dans la sauvegarde.
- ► PLUS ARCHIVELOG Inclure les fichiers de journalisation archivés dans la sauvegarde.
- ▶ DELETE [ALL] INPUT Supprimer les éléments sauvegardés (valable uniquement pour une sauvegarde de fichiers de journalisation archivés ou une sauvegarde de jeu de sauvegarde).
- ► FORMAT [=] 'format'
  Spécifier un format pour la sauvegarde (chemin et format de nom de fichier).
- ► TAG [=] 'nom' Associer un nom à la sauvegarde.
- ▶ NOT BACKED UP clause\_depuis Indiquer de ne sauvegarder que les éléments qui n'ont pas été sauvegardés depuis un certain temps.

#### Exemple Backup

▶ Pour sauvegarder la totalité de la base, il suffit d'utiliser l'option DATABASE dans la commande BACKUP :

#### BACKUP DATABASE :

- ▶ RMAN utilise les informations du fichier de contrôle de la base cible pour définir la liste des fichiers de données à sauvegarder. En complément, il sauvegarde le fichier de contrôle et le fichier de paramètres serveur (voir ci-après).
- ▶ La commande BACKUP VALIDATE DATABASE peut être utilisée pour vérifier que la base de données est en "bon état" (aucun fichier inaccessible, aucun fichier corrompu).

#### Sauvegarde incrémentale

- Avec RMAN, il est possible de réaliser des sauvegardes incrémentales, de la totalité de la base de données, de tablespaces individuels ou de fichiers de données individuels.
- ▶ L'objectif est de ne sauvegarder que les blocs qui ont été modifiés depuis la dernière sauvegarde.
- Les sauvegardes incrémentales présentent comme principal intérêt de réduire la taille des sauvegardes, notamment lorsque l'activité de mise à jour est relativement faible sur la base de données.
- ▶ Pour réaliser une sauvegarde incrémentale, il suffit d'inclure l'option INCREMENTAL LEVEL n [CUMULATIVE] dans la commande BACKUP.

#### Sauvegarde incrémentale

Une sauvegarde incrémentale peut être de niveau 0 ou de niveau 1, différentielle ou cumulative :

- ▶ Une sauvegarde incrémentale de niveau 0 sauvegarde toujours tous les blocs utilisés des fichiers de données. Elle est équivalente à une sauvegarde complète (mais une sauvegarde complète n'est pas considérée par RMAN comme une sauvegarde incrémentale de niveau 0).
- ▶ Une sauvegarde incrémentale différentielle de niveau 1 sauvegarde tous les blocs modifiés depuis la dernière sauvegarde incrémentale de niveau 0 ou 1. C'est le comportement par défaut.
- ▶ Une sauvegarde incrémentale cumulative de niveau 1 sauvegarde tous les blocs modifiés depuis la dernière sauvegarde incrémentale de niveau 0.

#### Sauvegarde incrémentale

- Les sauvegardes incrémentales cumulatives sont plus intéressantes pour la rapidité de récupération (moins de sauvegardes intermédiaires à appliquer) mais nécessitent plus d'espace disque.
- ▶ Lors d'une sauvegarde incrémentale de niveau 1, RMAN est obligé de lire tous les blocs utilisés pour trouver ceux qui ont été modifiés et doivent donc être sauvegardés.
- ► En conséquence, la durée de la sauvegarde n'est pas sensiblement réduite par rapport à une sauvegarde de niveau 0 (pas de gain sur la lecture, simple gain sur l'écriture).

#### Récupération

La stratégie de récupération dépend de plusieurs facteurs :

- ▶ De la nature du(des) fichier(s) endommagé(s) ou perdu(s)
  - ▶ fichier de données;
  - ▶ fichier de contrôle;
  - fichier de paramètres serveur;
  - fichier de journalisation.
- ▶ Du mode de fonctionnement de la base :
  - ► ARCHIVELOG
  - NOARCHIVELOG
- Des sauvegardes disponibles.

#### Que faire en cas de problème?

- 1. identifier la nature du problème;
- 2. définir le mode opératoire en tenant compte du mode de fonctionnement de la base et des sauvegardes disponibles.

#### Restauration ou récupération

Dans une opération de "restauration" ou de "récupération", il existe en fait deux étapes bien précises et bien distinctes :

- ▶ L'étape de restauration (restore) consiste à extraire d'une sauvegarde les fichiers nécessaires.
- ▶ L'étape de récupération (recover) consiste à appliquer les fichiers de journalisation aux fichiers récupérés de la sauvegarde.

- ► En mode NOARCHIVELOG, le mode opératoire est très simple :
  - restaurer la dernière sauvegarde complète de la base;
  - redémarrer la base.
- ➤ Toutes les modifications apportées depuis la dernière sauvegarde sont perdues.
- ▶ A priori, la restauration en mode NOARCHIVELOG ne permet pas de ramener la base de données à l'état où elle se trouvait juste avant l'incident
- ▶ Elle permet juste de ramener la base de données à l'état où elle se trouvait au moment de la sauvegarde.

#### Les situations sont les suivantes :

- ▶ Un cycle complet de basculement des fichiers de journalisation n'a pas eu lieu depuis la sauvegarde.
- ▶ Le fichier de données perdu n'est pas critique pour la base de données (n'appartient pas au tablespace SYSTEM, ni au tablespace d'annulation actif), ni pour l'application (ce n'est pas le tablespace principal de l'application).
- ➤ Tous les fichiers de contrôle sont perdus mais les autres fichiers (données et journalisation) sont intacts.

- ▶ Si les fichiers de journalisation n'ont pas subi un cycle complet de basculements depuis la sauvegarde utilisée, toutes les mises à jour effectuées depuis la sauvegarde en question sont encore "disponibles" dans les fichiers de journalisation.
- ▶ Dans ce cas, il faut réaliser une récupération comme si la base de données était en mode ARCHIVELOG

- ➤ Si le fichier de données perdu n'est pas critique pour la base de données ni pour l'application et que le problème est survenu alors que la base de données était arrêtée.
- La situation est plutôt favorable car les fichiers qui restent sont cohérents entre eux : si ce problème de fichier n'existait pas, le prochain démarrage ne nécessiterait pas de récupération de l'instance.

- ► En mode ARCHIVELOG, le mode opératoire de base pour une perte de fichier(s) de données est le suivant :
  - restaurer la dernière sauvegarde de chaque fichier perdu;
  - ▶ appliquer les fichiers de journalisation (archives puis ceux en ligne);
  - ▶ redémarrer la base (si la récupération n'a pas été faite base ouverte).
- ► Toutes les modifications apportées depuis les sauvegardes utilisées sont récupérées. La récupération est dite complète.

- Ce type de récupération est simple et ne pose pas de problème s'il reste au moins un fichier de contrôle, un membre par groupe de fichier de journalisation et que toutes les archives de fichiers de journalisation sont disponibles.
- ➤ Sur la base de ce scénario, différentes situations peuvent conduire à une récupération incomplète :
  - volontairement, pour s'arrêter avant un ordre SQL malencontreux;
  - ▶ involontairement, si des fichiers de journalisation sont perdus (une archive ou tout un groupe de fichiers de journalisation en ligne).

## Les incidents sur les fichiers de contrôle et de journalisation

- ▶ Les incidents sur les fichiers de contrôle et les fichiers de journalisation peuvent être classés en deux catégories : "peu graves" et "très graves".
- ▶ Incidents peu graves :
  - perte d'un ou plusieurs fichiers de contrôle, du moment qu'il en reste au moins un;
  - perte d'un ou plusieurs fichiers de journalisation, du moment qu'il en reste au moins un par groupe.
- ▶ Incidents plus graves et plus complexes à traiter :
  - perte de tous les fichiers de contrôle : moyennement grave si les autres fichiers sont intacts;
  - ▶ perte de tous les membres d'un groupe de fichiers de journalisation : la gravité dépend du statut du groupe perdu (CURRENT, ACTIVE, INACTIVE).

#### Les incidents les fichiers de contrôle

- ➤ Ces situations sont évitées si les fichiers de contrôle et les fichiers de journalisation sont correctement multiplexés.
- ▶ La perte de tous les fichiers de contrôle n'est pas la situation la plus complexe à traiter, s'il existe des sauvegardes récentes du fichier de contrôle et si les autres fichiers (particulièrement, les fichiers de journalisation) sont intacts.
- ▶ Dans ce cas, une récupération complète est possible.

#### Les incidents les fichiers de journalisation

- ▶ La perte de tous les membres d'un groupe de fichiers de journalisation est bien plus complexe à traiter.
- ▶ La situation de départ doit être analysée avec soin (statut du groupe perdu, état des autres fichiers, etc.), afin de choisir le bon mode opératoire.
- ▶ Pour les situations complexes, il est vivement conseillé de se faire aider par le support Oracle.