Audit activité base Oracle / SAP

Version : 1.0
Date : 04/02/2009
Auteur : xxxxxxxxxxx

SOMMAIRE

1 OBJECTIFS DE CE DOCUMENT....3

2 Résultats de l'étude....4

2.1. Bref exposé des constatations....4

2.1.1 Le CPU....4
2.1.2 L'efficacité des entrées sorties....4
2.1.3 La gestion de la mémoire....4
2.1.4 Les QUERY perturbateurs....4

2.2. Recommandations d'amélioration....5

2.2.1 La piste mémoire....5
2.2.2 La piste entrées/sorties....5
2.2.3 La piste « statistiques »....6
2.2.4 Pistes diverses....6
2.2.4.1 Éléments « parasites »....6
2.2.4.2 Réseau....6
2.2.4.3 Tuning supplémentaire de l'allocation mémoire....6

3 Constatations détaillées....7

3.1. Efficacité du cache de données....7

3.2. Perte d'efficacité due aux balayages de tables....7

3.3. Efficacité des entrées/sorties....7

3.3.1 Limitation de la bande passante....7
3.3.2 Saturation....7

4 Annexes....8

4.1. Efficacité du cache....8

4.2. Quelques mesures sur les I/O....9

4.3. Importance du « balayage »....9

4.4. Exemples de saturation....10

4.5. Origine des attentes de la base de données....12

4.6. Requêtes non performantes....12

4.7. Analyse de l'activité des différents « FILE SYSTEMS »....15

1 OBJECTIFS DE CE DOCUMENT.

Ce document est destiné à présenter les résultats de l'audit de performances effectué les 2 et 3 février 2009 sur la base Oracle du serveur xxxx.

Ce document est organisé comme suit :

- Résultats de l'étude
  - Bref exposé des constatations
  - Recommandations d'amélioration
- Constatations détaillées
- Annexes : copies d'écrans

2 RÉSULTATS DE L'ÉTUDE.

2.1. BREF EXPOSÉ DES CONSTATATIONS.

Les constatations tournent principalement autour de quatre éléments :

- Le CPU
- L'efficacité des Entrées/sorties
- La gestion de la mémoire
- Les « QUERY » perturbateurs.

2.1.1 LE CPU.

Pendant toute la durée des tests la charge CPU, de 20 à 40%, a été modérée et n'est certainement pas la source des ralentissements constatés.

2.1.2 L'EFFICACITÉ DES ENTRÉES SORTIES.

Au vu de la taille de la base (# 300 G Octets) l'utilisation d'une baie partagée est justifiée au moins du point de vue économique.

Toutefois certaines limitations liées au débit global apparaissent nettement lors des mesures, il sera nécessaire d'effectuer des recherches plus ciblées permettant d'identifier de manière plus précise le goulot d'étranglement qui « bride » les entrées / sorties de cette machine.

Le volume d'I/O est augmenté par deux phénomènes :

- insuffisance de la taille du « cache » de la base de données, les données absentes du cache doivent être lues depuis les disques.
- Insuffisance de la lecture prévisionnelle (MULTIBLOCK_READ), l valeur de 8 est faible.
- Présence de « grosses » requêtes de style batch qui accèdent à de très gros volumes de données et « épuisent » le cache de la base qui doit donc ensuite recharger des données depuis le disque augmentant encore le volume I/O.

2.1.3 LA GESTION DE LA MÉMOIRE.

Mémoire totale # 10Go dont 8 accordées à l'instance centrale de SAP et dont l'utilisation constatée ne dépasse pas 2,6 G Octets.

Il est donc possible de réduire cette réservation et d'utiliser la mémoire devenue disponible pour augmenter la taille des caches de la base de données.

2.1.4 LES QUERY PERTURBATEURS.

Nous avons constaté la présence de « grosses » requêtes n'utilisant pas ou peu les index et provoquant des balayages de tables de plusieurs millions (dizaine de millions) de rangs pour finalement n'en récupérer que quelques dizaines de milliers.

Lors des tests nous avons constaté qu'un certain nombre des requêtes lourdes utilisaient des balayages de tables plutôt que des accès indexés. Les origines de ce phénomène sont multiples :

- Certains cas semblent être dus à l'âge des statistiques qui ne permettent pas à l'optimiseur de requêtes de faire un « bon » choix puisqu'il se fonde sur des informations erronées ou incomplètes.
- D'autres cas sont visiblement dus à une inadéquation des index aux requêtes que nous avons analysées.

2.2. RECOMMANDATIONS D'AMÉLIORATION.

Plusieurs pistes ont été mises au jour lors de cette analyse et les recommandations sont de quatre ordres :

- Gestion mémoire
- Entrées / sorties
- Statistiques de la base de données
- Autres éléments.

2.2.1 LA PISTE MÉMOIRE.

Nous avons vu que l'on peut « récupérer » de la mémoire réservée par SAP et inutilisée.

Diminuer la mémoire réservée SAP de 8 à 4 G Octets récupérant ainsi 4 000 M Octets dont 1536 M Octets seraient transférés à Oracle, la meilleure utilisation de cette mémoire serait de :

- Augmenter la taille de « SGA_MAX_SIZE » de 2600 M Octets à 4096 M Octets
- Augmenter le CACHE_SIZE de 880 M Octets à 2304 Méga Octets.

Effets attendus :

- Meilleure efficacité du cache → transactions plus rapides
- Meilleure « résistance » du cache aux grosses requêtes → transactions plus régulières et temps de réponse utilisateur plus constant. Au lieu de varier de 1 à 4 les temps de réponse ne varieront plus que de 1 à 2 (ces chiffres ne sont qu'un exemple). L'irrégularité de temps de réponse est ce qui agace la plus les utilisateurs.
- Meilleure efficacité du cache → un peu moins d'I/O

Remarque :

Le total de la mémoire allouée actuellement dépasse la mémoire physique, si SAP voulait utiliser les 8 G Octets réservés la limite de la mémoire physique (10 G Octets) serait largement dépassée provoquant un swapping désastreux pour les performances. La réduction de 8 à 4 G Octets pallie à ce risque.

2.2.2 LA PISTE ENTRÉES/SORTIES.

Première manière de baisser le nombre d'I/O :

augmenter la valeur du paramètre « DB_FILE_MULTIBOLCK_READ » de 8 à 16 afin de diminuer le nombre d'I/O.

- On pourra ensuite après expérimentation de passer éventuellement à 24 ou 32 mais cela risque alors de pénaliser quelque peu le TP utilisateur.

Il a été constaté que certains fichiers (« data files » Oracle) sont beaucoup plus sollicités que les autres :

- une répartition plus optimisée dans la baie de disques pourrait participer à l'amélioration des performances. Il faudra pour cela disposer d'informations et de statistiques sur le fonctionnement de la baie de disques.

Voir les détails d'implantation et de statistiques I/O en 4.7

Effets attendus :

- Meilleur « rendement » global des entrées / sorties
- Diminution des temps de lecture / écriture entraînant une amélioration des performances ressenties pas les utilisateurs.

Du fait de l'implantation des FILE SYSTEM dans une baie mutualisée il peut être assez difficile d'obtenir des informations pertinentes permettant de prendre les meilleures décisions d'optimisation.

2.2.3 LA PISTE « STATISTIQUES ».

l'optimiseur de requêtes d'Oracle utilise des statistiques sur la répartition de la valeur des données au sein des tables et index pour « choisir » un « bon » plan d'exécution de la requête demandée.

Si ces statistiques ne sont pas à jour par rapport aux données la décision peut être franchement mauvaise et augmenter le nombre d'I/O nécessaire à la réalisation de l'objectif fixé.

Effets attendus :

- Meilleurs plan d'exécution des demandes entraînant :
  - Diminution des temps de réponse
  - Légère diminution des I/O

Il est conseillé d'utiliser l'outil SAP « BRCONNECT » pour automatiser ces opérations de mise à jour des statistiques.

2.2.4 PISTES DIVERSES.

Quelques autres éléments peuvent facilement être contrôlés et éventuellement modifiés.

2.2.4.1 Éléments « parasites »

Le paramétrage Oracle (INIT.ORA) montre la présence d'un certain nombre de « EVENTS » positionnés, ces éléments sont en général destinés au débogage et sont souvent consommateurs de ressources. Il faudrait vérifier leur pertinence (demandés par SAP) et, s'ils ne sont pas utiles les supprimer.

2.2.4.2 Réseau.

L'analyse des temps de latence de la base a montré une attente importante au niveau de la couche réseau, il ne semble pas que les serveurs, application SAP et base de données bénéficient d'un réseau privé ce qui peut être la cause de cette attente assez importante.

2.2.4.3 Tuning supplémentaire de l'allocation mémoire.

Oracle permet de diviser le cache de données et d'affecter certaines tables à certaines parties du cache qui seront gérées avec des règles différentes. On peut ainsi affecter une partie du cache à des tables de taille « raisonnable » très sollicitées et qui ne risqueront plus ainsi d'être évincées du cache par des balayages intempestifs (notion de BUFFER KEEP) et conserveront une performance optimale.

Toutefois au vu du nombre da tables des applications SAP cette action est assez délicate à définir.

Une action qui pourrait aussi être envisagée serait d'isoler dans une partie du cache les tables soufrant de balayages intempestifs, leur « nocivité » serait ainsi limitée aux requêtes qui utilisent ces tables (notion de BUFFER RECYCLE).

3 CONSTATATIONS DÉTAILLÉES

Aucune saturation ni même de charge élevée des CPU n'a été constatée lors de la période de tests.

3.1. EFFICACITÉ DU CACHE DE DONNÉES

95% et même 97% sont généralement admis comme limite d'un bon rendement de la base de données qui garantit un fonctionnement efficace.

Lors des tests nous avons constaté que les taux obtenus sont souvent bien en dessous de cette valeur et atteignent parfois des taux instantanés voisins de 40% qui déclenchent une augmentation catastrophique des temps de réponse vus du poste des utilisateurs.

Voir quelques illustrations en annexe 4.1 .

3.2. PERTE D'EFFICACITÉ DUE AUX BALAYAGES DE TABLES.

Lors de ces tests nous avons pu constater très nombreux cas de surcharge d'I/O dues à des balayages de tables. A certains moments la base fournit plus de lignes obtenues par balayage que de lignes obtenues pas accès direct.

Voir illustration de ce phénomène en 4.3.

3.3. EFFICACITÉ DES ENTRÉES/SORTIES.

Les entrées sorties marquent deux catégories de problèmes :

- Limitation de la bande passante
- Saturation conduisant à des délais anormaux.

3.3.1 LIMITATION DE LA BANDE PASSANTE.

Nous avons constaté que le débit d'I/O total (tous disques confondus) variait assez peu en fonction de la charge comme si un élément dans le circuit de lecture limitait la bande passante. Si un FILE SYSTEM avait un débit élevé alors les autres avaient un débit plus faible.

Lors de quelques mesures (après 16h00 alors que l'activité de la société est moins élevée) nous avons constaté des débits instantanés plus de trois fois plus élevés que ceux constatés aux heures de pointe d'activité.

Voir quelques exemples en annexe 4.2

3.3.2 SATURATION.

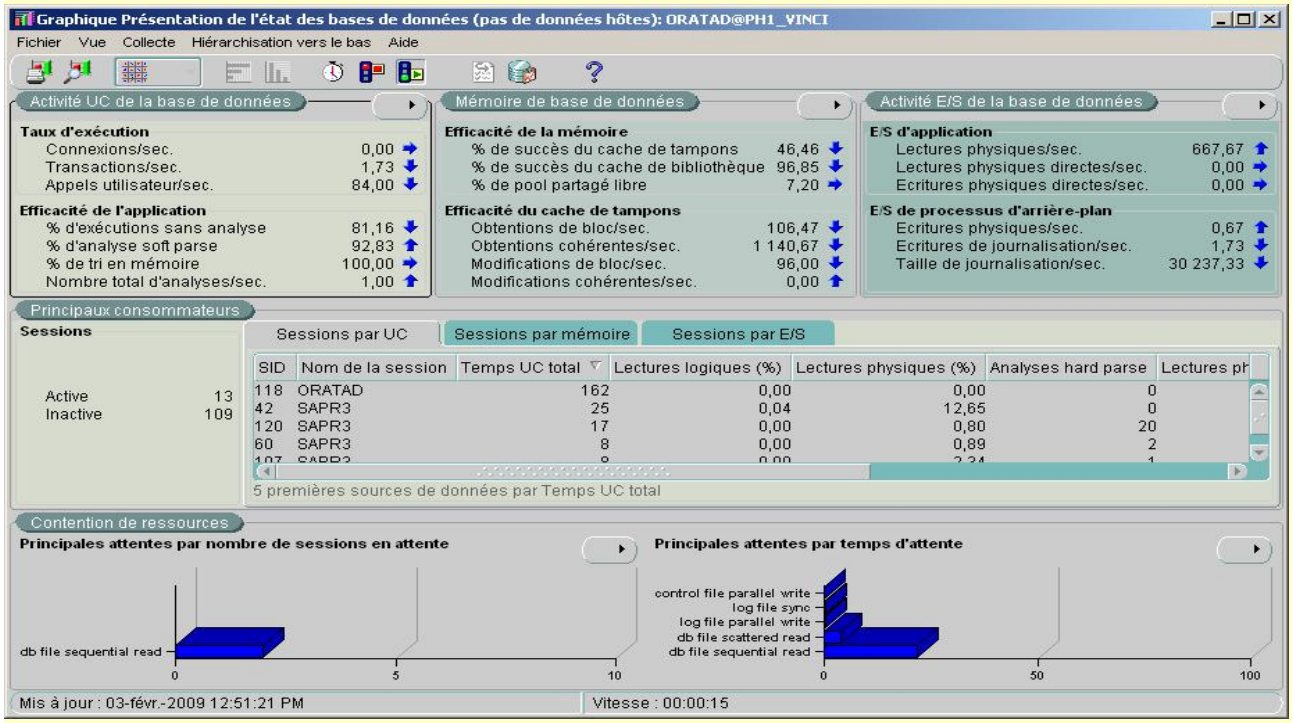
Lors de quelques occasions nous avons constaté un quasi blocage des I/O avec un débit fortement réduit et des temps d'attente 10 à 20 fois plus élevés que la moyenne et atteignant plus de 50 millisecondes sur quelques devices alors qu'en général le temps constaté est de 2 à 4 millisecondes.

Voir une illustration en annexe 4.4

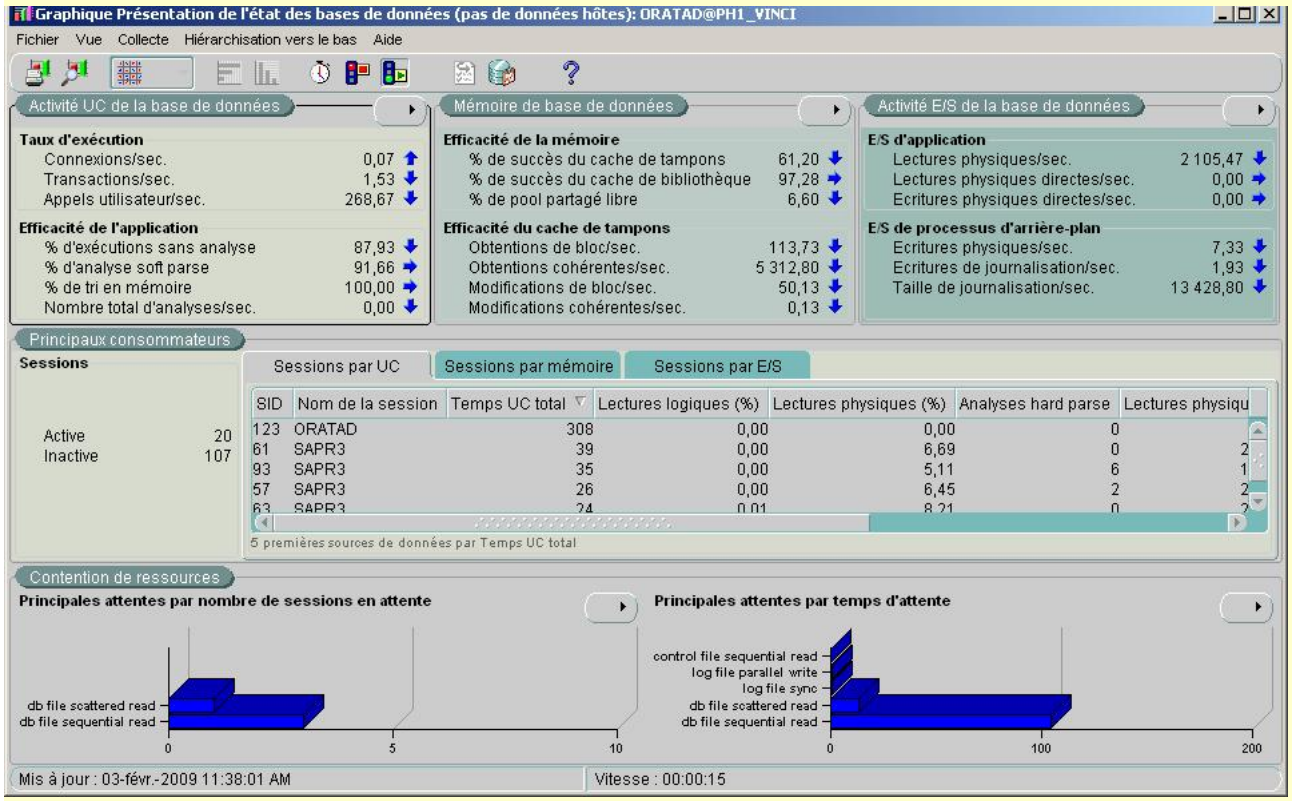
4 ANNEXES.

4.1. EFFICACITÉ DU CACHE.

Quelques captures d'écran illustrant la baisse importante du rendement de la base de données.



Cache à 47% avec un taux de lecture faible < 700 lectures par seconde.



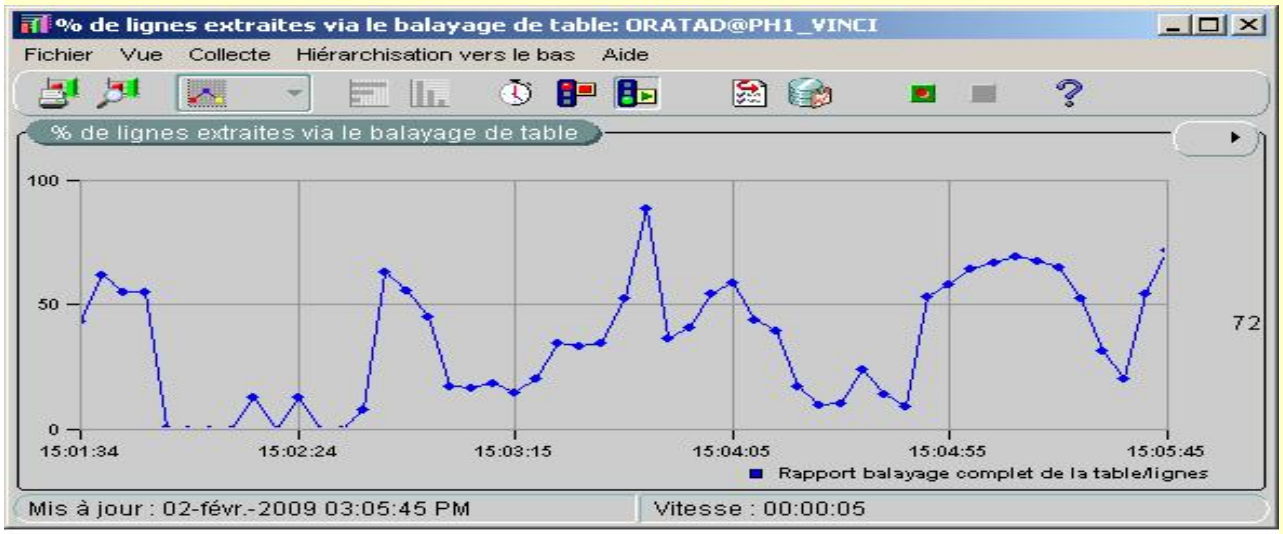
Cache à 61% avec un taux de lecture élevé à 2100 lectures par seconde.

4.2. QUELQUES MESURES SUR LES I/O.

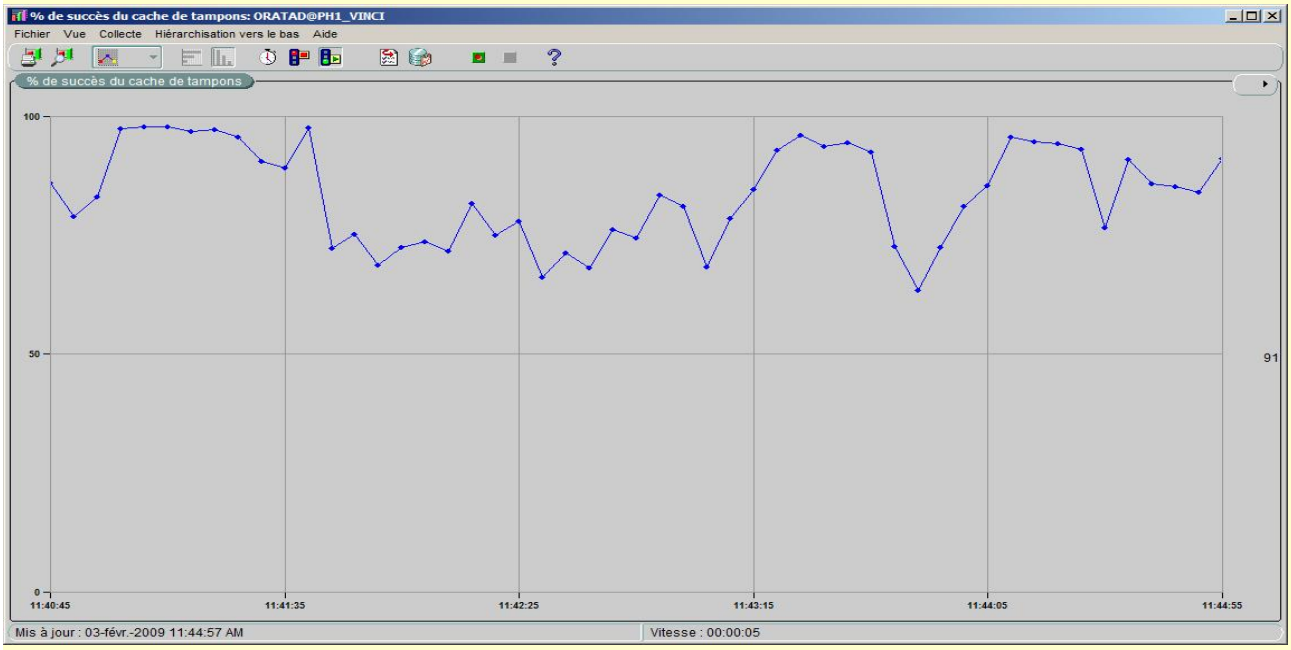
Les chiffres présentés ici sont des mesures instantanées résultant de l'utilisation de l'outil SAR, les mesures reflètent l'état des deux ou trois dernières secondes.

Les deux premiers tableaux représentent bien l'activité disque telle que nous l'avons constatée lors des deux journées de tests. Le troisième tableau est une exception qui semble montrer que le sous-ensemble disques est capable de performances bien meilleures que ce que nous avons couramment constaté. Pendant cette activité le CPU est monté à environ 45% montrant que les unités de calcul ne sont pas le goulot d'étranglement.

4.3. IMPORTANCE DU « BALAYAGE »

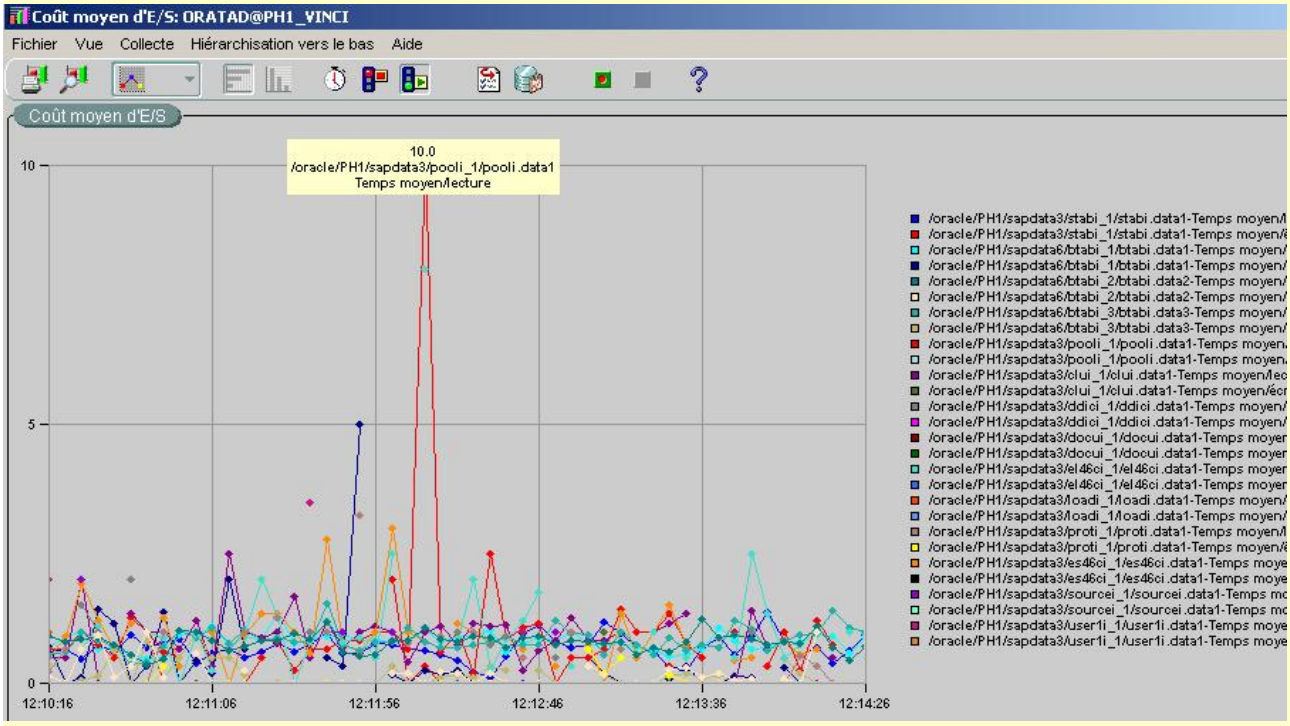


Le pourcentage de lignes extraites par balayage devrait rester faible (<= 15%) or il est souvent beaucoup plus fort. Déclenchant en général un pompage du cache donnant des graphes d'efficacité du cache tels que celui qui suit, pour rappel le taux de succès du cache doit être supérieur à 95% pour un rendement correct :

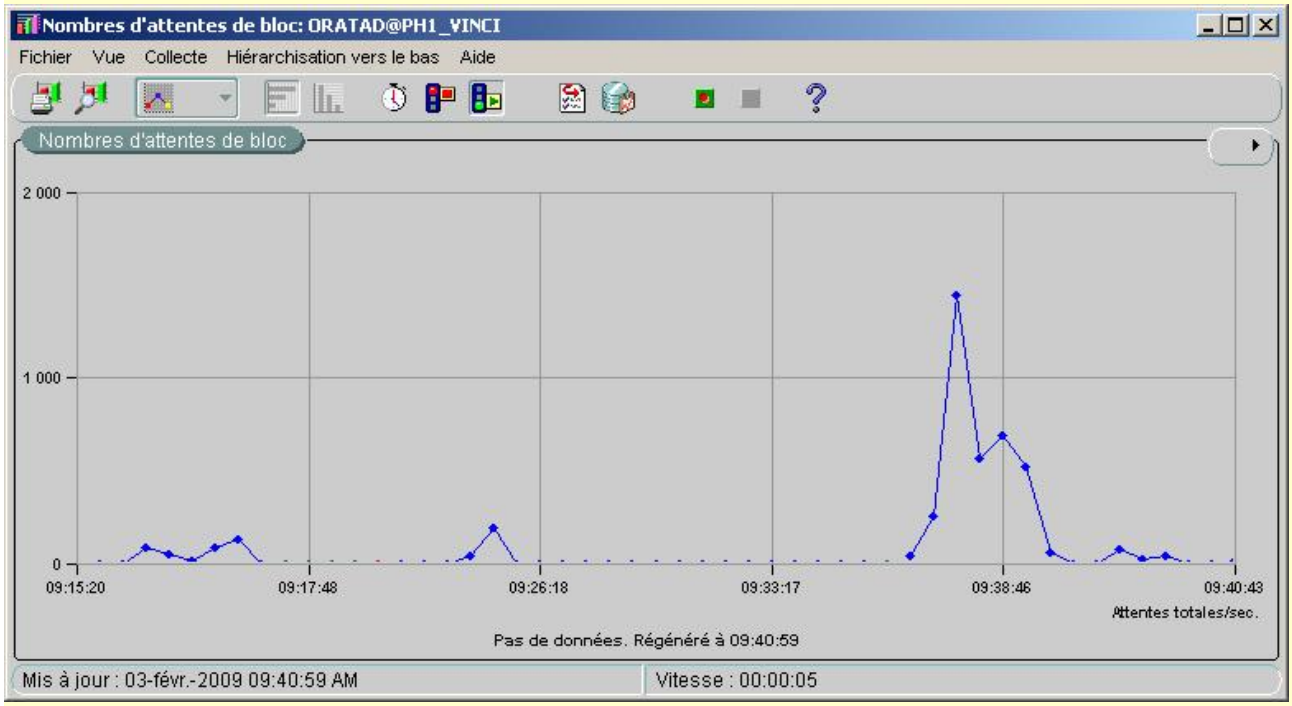


4.4. EXEMPLES DE SATURATION.

Suivi des temps d'I/O.



Alors que la moyenne est < 2,5 millisecondes quelques unités montrent des délais beaucoup plus importants.



Alors que l'activité est assez régulière des pics importants apparaissent.

Un autre exemple sur les mesures brutes des disques par « SAR » :

03/02/2009 15:37							
	device	%busy	Avg queue	R+W /sec	Blocks sec	Avg wait	Avg service
Average	c2t0d0	2,50	0,50	4	40	0	7,28
Average	c3t2d0	1,00	0,50	2	32	0	4,45
Average	c11t0d0	28,50	0,50	173	3 308	0	2,00
Average	c9t0d0	26,50	0,50	172	3 259	0	1,83
Average	c11t0d2	55,00	0,50	237	3 792	0	3,06
Average	c9t0d2	63,00	0,50	244	3 896	0	3,31
Average	c10t0d1	63,00	0,50	192	3 072	0	4,08
Average	c8t0d1	63,00	0,50	191	3 056	0	4,29
Average	c11t0d4	87,50	0,50	217	3 472	0	6,11
Average	c9t0d4	88,50	0,50	209	3 344	0	6,19
Average	c10t0d5	0,50	0,50	1	16	0	5,99
				Total	**27 287**	**Moyenne**	**4,42**

03/02/2009 15:43							
	device	%busy	Avg queue	R+W /sec	Blocks sec	Avg wait	Avg service
Average	c2t0d0	2,46	0,50	4	45	0	8,13
Average	c3t2d0	1,48	0,50	3	41	0	6,68
Average	c11t0d0	1,97	0,50	17	323	0	1,05
Average	c9t0d0	1,48	0,50	17	292	0	1,08
Average	c11t0d2	38,92	0,50	349	5 580	0	1,25
Average	c9t0d2	41,87	0,50	439	7 031	0	1,03
Average	c10t0d1	52,22	0,50	363	5 809	0	1,56
Average	c8t0d1	54,68	0,50	370	5 919	0	1,59
Average	c11t0d4	47,29	0,50	144	2 301	0	3,30
Average	c9t0d4	42,36	0,50	146	2 341	0	2,97
				Total	**29 682**	**Moyenne**	**2,86**

A une heure moins chargée les performances IO maximales (débit en blocs/seconde) sont très supérieures :

03/02/2009 16:11							
	device	%busy	Avg queue	R+W /sec	Blocks sec	Avg wait	Avg service
Average	c2t0d0	2,48	0,50	5	50	0	8,65
Average	c3t2d0	1,49	0,50	3	43	0	5,97
Average	c11t0d0	2,48	0,50	8	230	0	2,89
Average	c9t0d0	1,98	0,50	6	166	0	3,36
Average	c11t0d2	74,26	0,50	925	14 804	0	1,17
Average	c9t0d2	69,31	0,50	1 035	16 554	0	0,93
Average	c10t0d1	93,56	0,50	936	14 978	0	2,26
Average	c8t0d1	86,63	0,50	1 070	17 117	0	1,64
Average	c11t0d4	28,71	0,50	139	2 226	0	2,12
Average	c9t0d4	33,17	0,50	168	2 685	0	2,11
Average	c8t0d5	1,98	0,50	4	71	0	3,91
Average	c10t0d5	0,50	0,50	2	40	0	0,74
				Total	**68 964**	**Moyenne**	**2,98**

4.5. ORIGINE DES ATTENTES DE LA BASE DE DONNÉES.

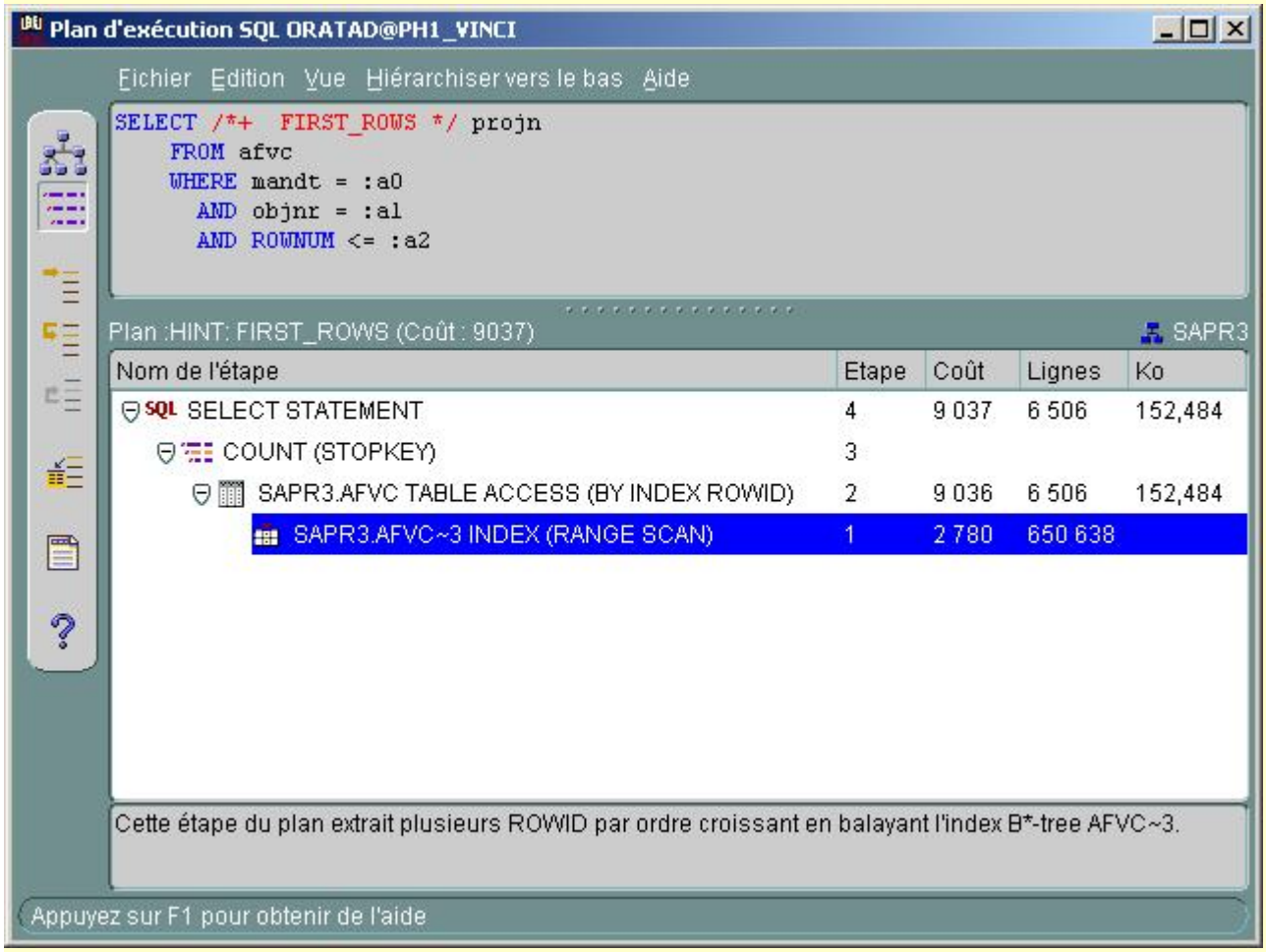
```
TYPE  TOTAL_WAITS TOTAL_TIMEOUTS TIME_WAITED AVERAGE_WAIT
----- ----------- -------------- ----------- ------------
BUF       3446838          31881     4481955   ,769047882
FILE    177470344              3   113103854   1,56909193
LIBR         1932             29       14181   ,136238629
Log       1916781              2     2112825   ,907212382
Net     425293191              0   684840629   ,621010456
PX         754555          13932     2992781    ,25212503
ZDIVR     2075784         312708   164937118   ,012585305
```

Les attentes liées à la couche réseau représentent un élément important des attentes de la base de données. Il y a peut-être là matière à amélioration. Ces attentes sont, en principe, le signe d'une concurrence sur le brin réseau reliant les serveurs d'application à la base de données.

Cela peut aussi être du partiellement au traitement de gros états tels que ceux dont nous avons vu l'influence sur les temps de réponse.

4.6. REQUÊTES NON PERFORMANTES.

Nous avons aussi repéré un certain nombre de requêtes non performantes déclenchant une baisse significative des performances.



L'accès est fait ici en utilisant l'index AFVC~3 qui contient les champs :

```
AFVC~3          AFVC                          1 MANDT
AFVC~3          AFVC                          2 PROJN
```

Or le champ « MANDT » ne possède qu'une seule valeur et la deuxième clef d'accès « OBJNR » ne figure pas dans cet index, cet ordre a donc balayé 650 000 rangs pour en sélectionner 6500.

Il faudrait étudier ces requêtes et créer éventuellement les index nécessaires à une meilleure performance.

Quelques tables particulièrement remarquées lors des tests :

BKPF

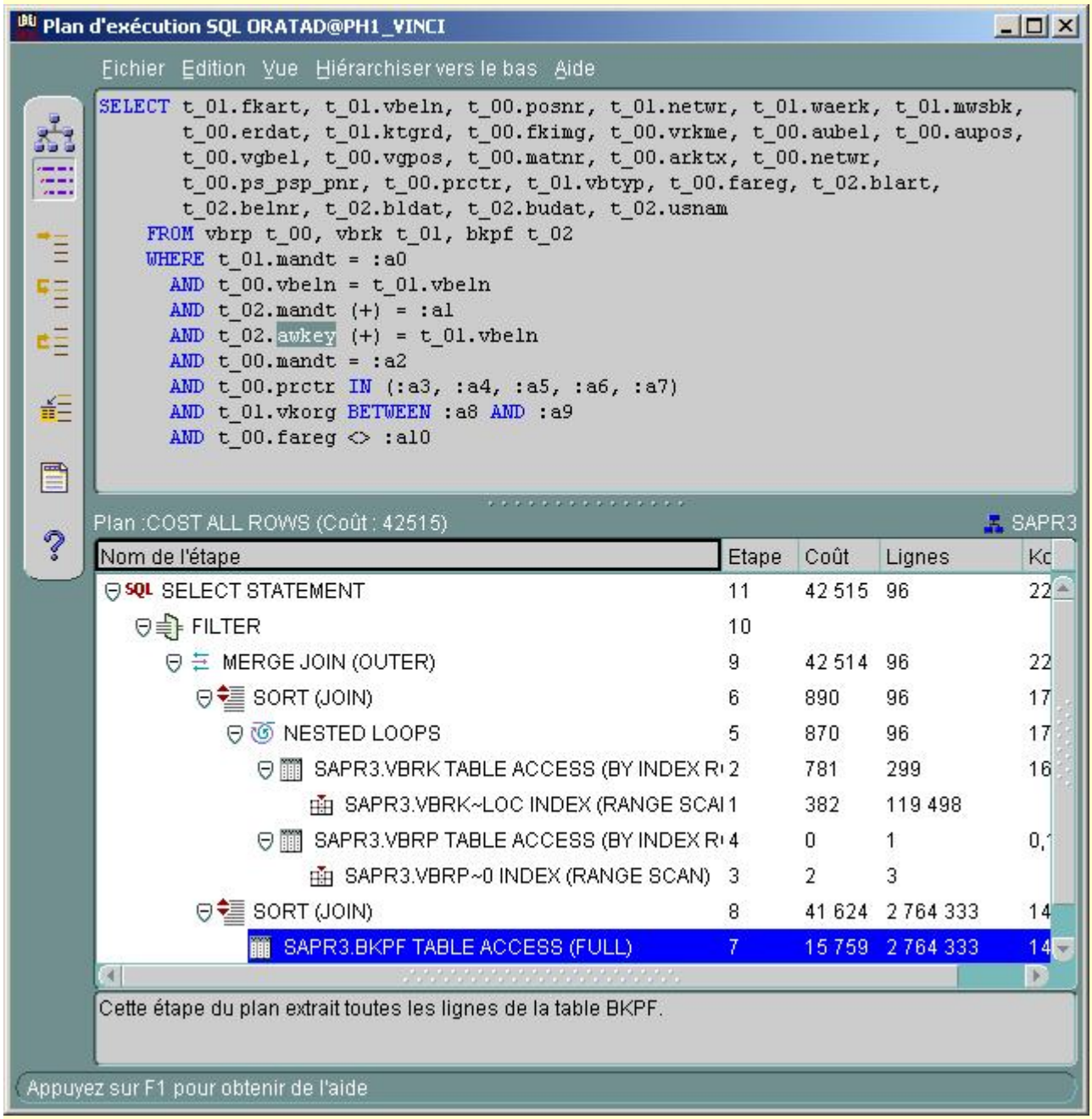
CMFP

GLPCA

MAKT

MARA

MARC



Ici une table de 2 760 000 rangs a été balayée pour en récupérer 119 500 déclenchant la lecture de plus de 140 millions d'octets depuis les disques.

Détails de l'étape 1

E/S - Coût :	5 731
Type d'index :	NON-UNIQUE
Statut de l'optimiseur :	ANALYZED
Nom de l'index :	STPO~K
Dernière analyse :	27-nov.-2006
Profondeur de l'index B*tree :	2
Nombre de clés distinctes :	1
Facteur de cluster :	127470

Ici trace d'un index non analysé depuis plus de deux ans, les décisions de l'optimiseur de requêtes peuvent être gravement perturbées.

Nous avons même trouvé un index jamais analysé (heureusement sur une petite table) :

Plan d'exécution SQL ORATAD@PH1_VINCI

Fichier Edition Vue Hiérarchiser vers le bas Aide

```
SELECT t_00.lgnum, t_00.tanum, t_00.tapos, t_00.vltyp, t_00.nltyp, t_01.bwlvs,
       t_01.bdatu, t_01.bzeit
    FROM ltap t_00, ltak t_01
    WHERE t_01.mandt = :a0
      AND t_01.lgnum = t_00.lgnum
      AND t_01.tanum = t_00.tanum
      AND t_00.mandt = :a1
      AND t_00.lgnum = :a2
      AND t_00.pquit = ' '
```

Détails de l'étape 1

E/S - Coût :	2
Type d'index :	UNIQUE
Nom de l'index :	LTAP~0
Dernière analyse :	Jamais

Plan :COST ALL ROWS (Coût : 4) SAPR3

Nom de l'étape	Etape	Coût	Lignes	Ko
SELECT STATEMENT	6	4	14	0,752
NESTED LOOPS	5	3	14	0,752
SAPR3.LTAP TABLE ACCESS (BY INDEX ROWID)	2	0	14	0,342
SAPR3.LTAP~0 INDEX (RANGE SCAN)	1	2	227	
SAPR3.LTAK TABLE ACCESS (BY INDEX ROWID)	4	0	1	0,029
SAPR3.LTAK~0 INDEX (UNIQUE SCAN)	3	1	1	

Cette étape du plan extrait plusieurs ROWID par ordre croissant en balayant l'index B*-tree LTAP~0.

Appuyez sur F1 pour obtenir de l'aide

4.7. ANALYSE DE L'ACTIVITÉ DES DIFFÉRENTS « FILE SYSTEMS ».

L'espace « UNDO » se trouve dans un file system assez chargé il serait indiqué de le déplacer vers un endroit moins sollicité.

L'espace « 5 » supporte 36% des I/O.