



Configurations Oracle

Enterprise applications

NetApp
January 16, 2025

Sommaire

- Configurations Oracle 1
 - Présentation 1
 - Instance unique Oracle 1
 - RAC étendu Oracle 2
 - Disjoncteur d'attache RAC 4

Configurations Oracle

Présentation

L'utilisation de la synchronisation active SnapMirror n'ajoute pas nécessairement aux meilleures pratiques d'exploitation d'une base de données ou ne modifie pas nécessairement ces pratiques.

La meilleure architecture dépend des besoins de l'entreprise. Par exemple, si l'objectif est de bénéficier d'une protection RPO=0 contre la perte de données, mais que l'objectif RTO est assoupli, l'utilisation de bases de données Oracle Single instance et la réplication des LUN avec SM-AS peuvent suffire et être moins coûteuses d'un standard de licences Oracle. Toute panne du site distant n'interrompt pas les opérations, et la perte du site principal entraînerait la présence de LUN en ligne et prêtes à être utilisées sur le site survivant.

Si le RTO était plus strict, l'automatisation actif-passif de base via des scripts ou des clusters comme Pacemaker ou Ansible améliorerait le délai de basculement. Par exemple, VMware HA peut être configuré pour détecter une panne de VM sur le site principal et activer cette dernière sur le site distant.

Enfin, pour un basculement extrêmement rapide, Oracle RAC peut être déployé sur plusieurs sites. L'objectif de délai de restauration serait essentiellement égal à zéro, car la base de données serait en ligne et disponible à tout moment sur les deux sites.

Instance unique Oracle

Les exemples décrits ci-dessous illustrent certaines des nombreuses options de déploiement des bases de données Oracle Single instance avec la réplication SnapMirror Active Sync.

[Oracle si avec accès non uniforme]

Basculement avec un système d'exploitation préconfiguré

La synchronisation active SnapMirror fournit une copie synchrone des données au niveau du site de reprise d'activité. Toutefois, la mise à disposition des données requiert un système d'exploitation et les applications associées. L'automatisation de base peut considérablement améliorer le délai de basculement de l'environnement global. Les produits Clusterware tels que Pacemaker sont souvent utilisés pour créer un cluster sur les sites et, dans la plupart des cas, le processus de basculement peut être piloté par des scripts simples.

En cas de perte des nœuds principaux, le cluster (ou les scripts) mettra les bases de données en ligne sur le site secondaire. Une option consiste à créer des serveurs de secours préconfigurés pour les ressources SAN qui constituent la base de données. En cas de défaillance du site principal, le logiciel de mise en cluster ou l'alternative scriptée effectue une séquence d'actions similaires à celles décrites ci-dessous :

1. Détection d'une défaillance du site principal
2. Effectuez la détection des LUN FC ou iSCSI
3. Montage de systèmes de fichiers et/ou montage de groupes de disques ASM
4. Démarrage de la base de données

Cette approche doit avant tout se passer d'un système d'exploitation en cours d'exécution sur le site distant.

Elles doivent être préconfigurées avec des binaires Oracle, ce qui signifie également que des tâches telles que l'application de correctifs Oracle doivent être effectuées sur les sites principal et de secours. Les binaires Oracle peuvent également être mis en miroir vers le site distant et montés en cas d'incident.

La procédure d'activation réelle est simple. Les commandes telles que la découverte de LUN ne nécessitent que quelques commandes par port FC. Le montage du système de fichiers n'est rien de plus qu'une `mount` commande et les bases de données et ASM peuvent être démarrés et arrêtés sur l'interface de ligne de commande à l'aide d'une seule commande.

Basculement avec un système d'exploitation virtualisé

Le basculement des environnements de base de données peut être étendu pour inclure le système d'exploitation lui-même. En théorie, ce basculement peut être effectué avec des LUN de démarrage, mais le plus souvent avec un système d'exploitation virtualisé. La procédure est similaire aux étapes suivantes :

1. Détection d'une défaillance du site principal
2. Montage des datastores hébergeant les machines virtuelles du serveur de base de données
3. Démarrage des machines virtuelles
4. Démarrage manuel des bases de données ou configuration des machines virtuelles pour démarrer automatiquement les bases de données.

Par exemple, un cluster ESX peut couvrir des sites. En cas d'incident, les machines virtuelles peuvent être mises en ligne sur le site de reprise après incident après le basculement.

Protection contre les défaillances du stockage

Le diagramme ci-dessus montre l'utilisation de "[accès non uniforme](#)", où le SAN n'est pas étendu entre les sites. Cela peut être plus simple à configurer et, dans certains cas, peut être la seule option étant donné les fonctionnalités SAN actuelles, mais cela signifie également que la défaillance du système de stockage principal entraînerait une panne de la base de données jusqu'à ce que l'application ait été ratée.

Pour une résilience supplémentaire, la solution pourrait être déployée avec "[accès uniforme](#)". Cela permettrait aux applications de continuer à fonctionner en utilisant les chemins annoncés à partir du site opposé.

RAC étendu Oracle

De nombreux clients optimisent leur RTO en étendant un cluster Oracle RAC sur plusieurs sites, offrant une configuration entièrement active/active. La conception globale devient plus complexe car elle doit inclure la gestion du quorum d'Oracle RAC.

La mise en cluster RAC étendue traditionnelle s'est appuyée sur la mise en miroir ASM pour assurer la protection des données. Cette approche fonctionne, mais elle implique également de nombreuses étapes manuelles de configuration et entraîne une surcharge de l'infrastructure réseau. À l'inverse, la réplication des données peut être prise en charge par la synchronisation active SnapMirror, ce qui simplifie considérablement la solution. Les opérations telles que la synchronisation, la resynchronisation après les interruptions, les basculements et la gestion du quorum sont plus simples. En outre, le SAN n'a pas besoin d'être distribué entre les sites, ce qui simplifie la conception et la gestion du SAN.

La réplication

Pour comprendre la fonctionnalité RAC sur SnapMirror Active Sync, il est essentiel de considérer le stockage

comme un ensemble unique de LUN hébergés sur un stockage en miroir. Par exemple :

[Accès logique à Oracle]

Il n'y a pas de copie principale ou miroir. Pour schématiser, il n'y a qu'une seule copie de chaque LUN et cette LUN est disponible sur les chemins SAN situés sur deux systèmes de stockage différents. Du point de vue de l'hôte, il n'y a pas de basculement de stockage ; il y a des changements de chemin. Plusieurs défaillances peuvent entraîner la perte de certains chemins vers la LUN, tandis que les autres chemins restent en ligne. La synchronisation active SnapMirror garantit la disponibilité des mêmes données sur tous les chemins opérationnels.

Configuration de stockage sous-jacente

Dans cet exemple de configuration, les disques ASM sont configurés de la même manière que dans n'importe quelle configuration RAC à site unique sur le stockage d'entreprise. Étant donné que le système de stockage assure la protection des données, la redondance ASM externe est utilisée.

Accès uniforme ou non informé

L'élément le plus important à prendre en compte avec Oracle RAC sur SnapMirror Active Sync est de savoir s'il faut utiliser un accès uniforme ou non.

Un accès uniforme signifie que chaque hôte peut voir les chemins sur les deux clusters. L'accès non uniforme signifie que les hôtes peuvent uniquement voir les chemins vers le cluster local.

Aucune de ces options n'est spécifiquement recommandée ou déconseillée. Certains clients ont facilement accès à la fibre noire pour connecter les sites, d'autres ne disposent pas d'une telle connectivité ou leur infrastructure SAN ne prend pas en charge l'ISL longue distance.

Accès non uniforme

L'accès non uniforme est plus simple à configurer du point de vue du SAN.

[Accès Oracle RAC non uniforme]

L'inconvénient principal de cette "accès non uniforme" approche est que la perte de la connectivité ONTAP site à site ou la perte d'un système de stockage entraînera la perte des instances de base de données sur un site. Cela n'est évidemment pas souhaitable, mais cela peut constituer un risque acceptable en échange d'une configuration SAN plus simple.

Accès uniforme

L'accès uniforme requiert l'extension du SAN sur les sites. Le principal avantage est que la perte d'un système de stockage n'entraîne pas la perte d'une instance de base de données. Au lieu de cela, cela entraînerait une modification des chemins d'accès multiples dans lesquels les chemins sont actuellement utilisés.

Il existe plusieurs façons de configurer l'accès non uniforme.



Dans les schémas ci-dessous, il existe également des chemins actifs mais non optimisés qui seraient utilisés en cas de défaillances simples du contrôleur, mais ces chemins ne sont pas affichés dans l'intérêt de simplifier les diagrammes.

AFF avec paramètres de proximité

En cas de latence importante entre les sites, les systèmes AFF peuvent être configurés avec des paramètres de proximité des hôtes. Cela permet à chaque système de stockage d'identifier les hôtes locaux et distants, et d'attribuer les priorités de chemin en conséquence.

[RAC avec accès uniforme]

En fonctionnement normal, chaque instance Oracle utilisera de préférence les chemins locaux actifs/optimisés. Par conséquent, toutes les lectures seront traitées par la copie locale des blocs. La latence est ainsi la plus faible possible. Les E/S d'écriture sont envoyées de la même manière vers le contrôleur local. L'E/S doit toujours être répliquée avant d'être reconnue, ce qui entraîne toujours une latence supplémentaire en traversant le réseau site à site, mais cela ne peut pas être évité dans une solution de réplication synchrone.

ASA / AFF sans paramètres de proximité

S'il n'y a pas de latence significative entre les sites, les systèmes AFF peuvent être configurés sans paramètres de proximité des hôtes, ou ASA peut être utilisé.

[RAC avec accès uniforme]

Chaque hôte pourra utiliser tous les chemins opérationnels sur les deux systèmes de stockage. Cela améliore considérablement les performances en permettant à chaque hôte d'exploiter le potentiel de performance de deux clusters, et non d'un seul.

Avec ASA, non seulement tous les chemins vers les deux clusters sont considérés comme actifs et optimisés, mais les chemins sur les contrôleurs partenaires sont également actifs. Il en résulte des chemins SAN entièrement actifs sur l'ensemble du cluster, à tout moment.



Les systèmes ASA peuvent également être utilisés dans une configuration d'accès non uniforme. Étant donné qu'il n'existe aucun chemin entre les sites, les performances ne seraient pas améliorées par le franchissement de l'ISL par les E/S.

Disjoncteur d'attache RAC

Bien que le RAC étendu utilisant la synchronisation active SnapMirror soit une architecture symétrique par rapport aux E/S, il existe une exception qui est connectée à la gestion du split-brain.

Que se passe-t-il si le lien de réplication est perdu et qu'aucun des sites n'a le quorum ? Que doit-on faire ? Cette question s'applique à la fois au comportement d'Oracle RAC et de ONTAP. Si les modifications ne peuvent pas être répliquées sur tous les sites et que vous souhaitez reprendre les opérations, l'un des sites devra survivre et l'autre site devra être indisponible.

Le système "[Médiateur de ONTAP](#)" répond à cette exigence au niveau de la couche ONTAP. Il existe plusieurs options pour le tricover RAC.

Disjoncteurs Oracle

La meilleure méthode pour gérer les risques Oracle RAC split-brain consiste à utiliser un nombre impair de nœuds RAC, de préférence à l'aide d'un Tiebreaker 3rd site. Si un troisième site n'est pas disponible, l'instance Tiebreaker pourrait être placée sur un site des deux sites, ce qui la désignerait en fait un site de survivant préféré.

Oracle et CSS_Critical

Avec un nombre pair de nœuds, le comportement par défaut d'Oracle RAC est que l'un des nœuds du cluster sera considéré plus important que les autres nœuds. Le site avec ce nœud de priorité supérieure survivra à l'isolation du site tandis que les nœuds de l'autre site seront supprimés. La hiérarchisation est basée sur plusieurs facteurs, mais vous pouvez également contrôler ce comportement à l'aide du `css_critical` paramètre.

Dans l'"[exemple](#)" architecture, les noms d'hôte des nœuds RAC sont `jfs12` et `jfs13`. Les paramètres actuels de `css_critical` sont les suivants :

```
[root@jfs12 ~]# /grid/bin/crsctl get server css_critical
CRS-5092: Current value of the server attribute CSS_CRITICAL is no.

[root@jfs13 trace]# /grid/bin/crsctl get server css_critical
CRS-5092: Current value of the server attribute CSS_CRITICAL is no.
```

Si vous voulez que le site avec `jfs12` soit le site préféré, définissez cette valeur sur oui sur un site Un noeud et redémarrez les services.

```
[root@jfs12 ~]# /grid/bin/crsctl set server css_critical yes
CRS-4416: Server attribute 'CSS_CRITICAL' successfully changed. Restart
Oracle High Availability Services for new value to take effect.

[root@jfs12 ~]# /grid/bin/crsctl stop crs
CRS-2791: Starting shutdown of Oracle High Availability Services-managed
resources on 'jfs12'
CRS-2673: Attempting to stop 'ora.crsd' on 'jfs12'
CRS-2790: Starting shutdown of Cluster Ready Services-managed resources on
server 'jfs12'
CRS-2673: Attempting to stop 'ora.ntap.ntappdb1.pdb' on 'jfs12'
...
CRS-2673: Attempting to stop 'ora.gipcd' on 'jfs12'
CRS-2677: Stop of 'ora.gipcd' on 'jfs12' succeeded
CRS-2793: Shutdown of Oracle High Availability Services-managed resources
on 'jfs12' has completed
CRS-4133: Oracle High Availability Services has been stopped.

[root@jfs12 ~]# /grid/bin/crsctl start crs
CRS-4123: Oracle High Availability Services has been started.
```

Informations sur le copyright

Copyright © 2025 NetApp, Inc. Tous droits réservés. Imprimé aux États-Unis. Aucune partie de ce document protégé par copyright ne peut être reproduite sous quelque forme que ce soit ou selon quelque méthode que ce soit (graphique, électronique ou mécanique, notamment par photocopie, enregistrement ou stockage dans un système de récupération électronique) sans l'autorisation écrite préalable du détenteur du droit de copyright.

Les logiciels dérivés des éléments NetApp protégés par copyright sont soumis à la licence et à l'avis de non-responsabilité suivants :

CE LOGICIEL EST FOURNI PAR NETAPP « EN L'ÉTAT » ET SANS GARANTIES EXPRESSES OU TACITES, Y COMPRIS LES GARANTIES TACITES DE QUALITÉ MARCHANDE ET D'ADÉQUATION À UN USAGE PARTICULIER, QUI SONT EXCLUES PAR LES PRÉSENTES. EN AUCUN CAS NETAPP NE SERA TENU POUR RESPONSABLE DE DOMMAGES DIRECTS, INDIRECTS, ACCESSOIRES, PARTICULIERS OU EXEMPLAIRES (Y COMPRIS L'ACHAT DE BIENS ET DE SERVICES DE SUBSTITUTION, LA PERTE DE JOUISSANCE, DE DONNÉES OU DE PROFITS, OU L'INTERRUPTION D'ACTIVITÉ), QUELLES QU'EN SOIENT LA CAUSE ET LA DOCTRINE DE RESPONSABILITÉ, QU'IL S'AGISSE DE RESPONSABILITÉ CONTRACTUELLE, STRICTE OU DÉLICTELLE (Y COMPRIS LA NÉGLIGENCE OU AUTRE) DÉCOULANT DE L'UTILISATION DE CE LOGICIEL, MÊME SI LA SOCIÉTÉ A ÉTÉ INFORMÉE DE LA POSSIBILITÉ DE TELS DOMMAGES.

NetApp se réserve le droit de modifier les produits décrits dans le présent document à tout moment et sans préavis. NetApp décline toute responsabilité découlant de l'utilisation des produits décrits dans le présent document, sauf accord explicite écrit de NetApp. L'utilisation ou l'achat de ce produit ne concède pas de licence dans le cadre de droits de brevet, de droits de marque commerciale ou de tout autre droit de propriété intellectuelle de NetApp.

Le produit décrit dans ce manuel peut être protégé par un ou plusieurs brevets américains, étrangers ou par une demande en attente.

LÉGENDE DE RESTRICTION DES DROITS : L'utilisation, la duplication ou la divulgation par le gouvernement sont sujettes aux restrictions énoncées dans le sous-paragraphe (b)(3) de la clause Rights in Technical Data-Noncommercial Items du DFARS 252.227-7013 (février 2014) et du FAR 52.227-19 (décembre 2007).

Les données contenues dans les présentes se rapportent à un produit et/ou service commercial (tel que défini par la clause FAR 2.101). Il s'agit de données propriétaires de NetApp, Inc. Toutes les données techniques et tous les logiciels fournis par NetApp en vertu du présent Accord sont à caractère commercial et ont été exclusivement développés à l'aide de fonds privés. Le gouvernement des États-Unis dispose d'une licence limitée irrévocable, non exclusive, non cessible, non transférable et mondiale. Cette licence lui permet d'utiliser uniquement les données relatives au contrat du gouvernement des États-Unis d'après lequel les données lui ont été fournies ou celles qui sont nécessaires à son exécution. Sauf dispositions contraires énoncées dans les présentes, l'utilisation, la divulgation, la reproduction, la modification, l'exécution, l'affichage des données sont interdits sans avoir obtenu le consentement écrit préalable de NetApp, Inc. Les droits de licences du Département de la Défense du gouvernement des États-Unis se limitent aux droits identifiés par la clause 252.227-7015(b) du DFARS (février 2014).

Informations sur les marques commerciales

NETAPP, le logo NETAPP et les marques citées sur le site <http://www.netapp.com/TM> sont des marques déposées ou des marques commerciales de NetApp, Inc. Les autres noms de marques et de produits sont des marques commerciales de leurs propriétaires respectifs.