



Configurazioni Oracle

Enterprise applications

NetApp

December 17, 2024

Sommario

- Configurazioni Oracle 1
 - Panoramica 1
 - Istanza singola Oracle 1
 - Oracle Extended RAC 2
 - RAC tiebreaker 4

Configurazioni Oracle

Panoramica

L'utilizzo della sincronizzazione attiva di SnapMirror non aggiunge o modifica necessariamente le procedure consigliate per il funzionamento di un database.

La migliore architettura dipende dai requisiti di business. Ad esempio, se l'obiettivo è ottenere una protezione RPO=0:1 contro la perdita di dati, ma l'RTO non è più così semplice, utilizzare i database Oracle Single Instance e replicare le LUN con SM, potrebbe essere sufficiente così come meno costoso da un approccio basato su licenze Oracle. Un guasto del sito remoto non interromperebbe le operazioni e, di conseguenza, la perdita del sito primario porterebbe all'utilizzo di LUN online e pronte per l'uso nel sito rimasto.

Se l'RTO fosse più rigoroso, l'automazione Active-passive di base tramite script o cluster come Pacemaker o Ansible migliorerebbe i tempi di failover. Ad esempio, VMware ha può essere configurato in modo da rilevare un guasto della VM nel sito primario e rendere attiva la VM nel sito remoto.

Infine, per un failover estremamente rapido, Oracle RAC può essere implementato su più siti. L'RTO sarebbe praticamente pari a zero perché il database sarebbe sempre online e disponibile su entrambi i siti.

Istanza singola Oracle

Gli esempi illustrati di seguito mostrano alcune delle numerose opzioni per la distribuzione dei database Oracle Single Instance con la replica di sincronizzazione attiva SnapMirror.

[Oracle si con accesso non uniforme]

Failover con un sistema operativo preconfigurato

La sincronizzazione attiva SnapMirror fornisce una copia sincrona dei dati nel sito di disaster recovery, ma la loro disponibilità richiede un sistema operativo e le applicazioni associate. L'automazione di base può migliorare notevolmente il tempo di failover dell'ambiente complessivo. I prodotti Clusterware come Pacemaker vengono spesso utilizzati per creare un cluster in tutti i siti, e in molti casi il processo di failover può essere guidato con semplici script.

In caso di perdita dei nodi primari, il clusterware (o gli script) porterà i database online nel sito alternativo. Un'opzione è creare server di standby preconfigurati per le risorse SAN che compongono il database. Se il sito primario non funziona, il clusterware o l'alternativa con script esegue una sequenza di azioni simile alle seguenti:

1. Rileva i guasti del sito primario
2. Rilevamento di LUN FC o iSCSI
3. Montaggio di file system e/o montaggio di gruppi di dischi ASM
4. Avvio del database

Il requisito principale di questo approccio è rappresentato da un sistema operativo in esecuzione sul sito remoto. Deve essere preconfigurato con i file binari di Oracle, il che significa anche che attività come l'applicazione delle patch Oracle devono essere eseguite sul sito primario e di standby. In alternativa, è possibile eseguire il mirroring dei file binari di Oracle nel sito remoto e montarli se viene dichiarato un disastro.

La procedura di attivazione effettiva è semplice. Comandi come il rilevamento delle LUN richiedono solo pochi comandi per ogni porta FC. Il montaggio del file system non è altro che un `mount` comando e sia i database che ASM possono essere avviati e arrestati dalla CLI con un unico comando.

Failover con un sistema operativo virtualizzato

Il failover degli ambienti di database può essere esteso per includere il sistema operativo stesso. In teoria, questo failover può essere eseguito con le LUN di avvio, ma nella maggior parte dei casi con un sistema operativo virtualizzato. La procedura è simile ai seguenti passaggi:

1. Rileva i guasti del sito primario
2. Montaggio dei datastore che ospitano le macchine virtuali del server di database
3. Avvio delle macchine virtuali
4. Avviare i database manualmente o configurare le macchine virtuali per avviare automaticamente i database.

Ad esempio, un cluster ESX può estendersi su diversi siti. In caso di disastro, dopo lo switchover, è possibile portare online le macchine virtuali nel sito di disaster recovery.

Protezione dai guasti dello storage

Il diagramma precedente mostra l'utilizzo di "accesso non uniforme", in cui la SAN non è estesa tra i siti. Questa configurazione potrebbe essere più semplice e, in alcuni casi, potrebbe essere l'unica opzione data le attuali funzionalità SAN, ma significa anche che un guasto del sistema di storage primario causerebbe un'interruzione del database fino a quando l'applicazione non è stata sottoposta a failover.

Per una maggiore resilienza, la soluzione potrebbe essere implementata con "accesso uniforme". Ciò consentirebbe alle applicazioni di continuare a funzionare utilizzando i percorsi pubblicizzati dal sito opposto.

Oracle Extended RAC

Molti clienti ottimizzano il proprio RTO estendendo un cluster Oracle RAC tra i vari siti, ottenendo una configurazione completamente Active-Active. La progettazione complessiva diventa più complicata perché deve includere la gestione del quorum di Oracle RAC.

Il cluster RAC esteso tradizionale si basava sul mirroring ASM per garantire la protezione dei dati. Questo approccio funziona, ma richiede anche molti passaggi di configurazione manuale e impone un overhead all'infrastruttura di rete. Al contrario, consentendo alla sincronizzazione attiva di SnapMirror di assumersi la responsabilità della replica dei dati, la soluzione risulta notevolmente semplificata. Operazioni quali sincronizzazione, risincronizzazione dopo interruzioni, failover e gestione del quorum sono più semplici, inoltre la SAN non deve essere distribuita tra i siti, semplificando la progettazione e la gestione della SAN.

Replica

La chiave per comprendere le funzionalità RAC nella sincronizzazione attiva di SnapMirror è la visualizzazione dello storage come singolo set di LUN che si trovano sullo storage con mirroring. Ad esempio:

[Accesso logico Oracle]

Nessuna copia primaria o copia speculare. A livello logico, esiste una sola copia di ogni LUN e tale LUN è

disponibile nei percorsi SAN che si trovano su due sistemi di storage differenti. Dal punto di vista dell'host, non ci sono failover dello storage e modifiche del percorso. Diversi eventi di errore potrebbero causare la perdita di determinati percorsi verso la LUN, mentre gli altri percorsi rimangono online. La sincronizzazione attiva di SnapMirror garantisce la disponibilità degli stessi dati su tutti i percorsi operativi.

Configurazione dello storage

In questa configurazione di esempio, i dischi ASM sono configurati come in qualsiasi configurazione RAC a sito singolo sullo storage Enterprise. Poiché il sistema di storage garantisce la protezione dei dati, viene utilizzata la ridondanza esterna ASM.

Accesso uniforme e non privo di informazioni

La considerazione più importante con Oracle RAC sulla sincronizzazione attiva di SnapMirror è se utilizzare un accesso uniforme o non uniforme.

Un accesso uniforme significa che ogni host può vedere i percorsi su entrambi i cluster. Accesso non uniforme significa che gli host possono vedere solo i percorsi verso il cluster locale.

Nessuna delle due opzioni è specificamente consigliata o scoraggiata. Alcuni clienti dispongono prontamente di fibra ottica spenta per la connessione ai siti, altri non dispongono di tale connettività oppure la loro infrastruttura SAN non supporta un ISL a lunga distanza.

Accesso non uniforme

L'accesso non uniforme è più semplice da configurare dal punto di vista della SAN.

[Accesso non uniforme a Oracle RAC]

L'aspetto negativo principale "accesso non uniforme" dell'approccio è rappresentato dalla perdita della connettività ONTAP sito-sito o dalla perdita di un sistema di storage che causa la perdita delle istanze di database in un singolo sito. Questo ovviamente non è desiderabile, ma potrebbe essere un rischio accettabile in cambio di una configurazione SAN più semplice.

Accesso uniforme

Un accesso uniforme richiede l'estensione della SAN tra i siti. Il vantaggio principale consiste nel fatto che la perdita di un sistema di storage non provocherà la perdita di un'istanza del database. Al contrario, si otterrebbe una modifica multipathing in cui i percorsi sono attualmente in uso.

Esistono diversi modi per configurare l'accesso non uniforme.



Nei diagrammi di seguito sono presenti anche percorsi attivi ma non ottimizzati che sarebbero utilizzati durante semplici guasti del controller, ma tali percorsi non sono mostrati nell'interesse di semplificare i diagrammi.

AFF con impostazioni di prossimità

In presenza di una latenza significativa tra i siti, è possibile configurare i sistemi AFF con le impostazioni di prossimità dell'host. In questo modo, ciascun sistema storage può conoscere gli host locali e remoti e assegnare in maniera appropriata le priorità del percorso.

[RAC con accesso uniforme]

Durante il normale funzionamento, ogni istanza Oracle utilizzerebbe preferenzialmente i percorsi locali attivi/ottimizzati. Il risultato è che tutte le letture saranno gestite dalla copia locale dei blocchi. In questo modo, si ottiene la minore latenza possibile. L'io in scrittura viene analogamente inviato ai percorsi al controller locale. L'io deve ancora essere replicato prima di essere riconosciuto e quindi sarebbe comunque soggetto alla latenza aggiuntiva che attraversa la rete site-to-site, ma ciò non può essere evitato in una soluzione di replica sincrona.

ASA / AFF senza impostazioni di prossimità

In assenza di una latenza significativa tra i siti, è possibile configurare i sistemi AFF senza le impostazioni di prossimità dell'host oppure utilizzare ASA.

[RAC con accesso uniforme]

Ciascun host sarà in grado di utilizzare tutti i percorsi operativi su entrambi i sistemi storage. Questo consente di migliorare potenzialmente le prestazioni consentendo a ciascun host di sfruttare il potenziale in termini di prestazioni di due cluster, non di uno solo.

Con ASA, non solo tutti i percorsi verso entrambi i cluster sono considerati attivi e ottimizzati, ma anche i percorsi sui partner controller sarebbero attivi. Ne risulterebbero percorsi SAN all-Active sull'intero cluster, in qualsiasi momento.



I sistemi ASA possono anche essere utilizzati in una configurazione di accesso non uniforme. Poiché non esistono percorsi tra siti, non vi sarebbe alcun impatto sulle performance risultanti dall'io che attraversa l'ISL.

RAC tiebreaker

Sebbene il RAC esteso che utilizza la sincronizzazione attiva di SnapMirror sia un'architettura simmetrica rispetto all'io, esiste un'eccezione connessa alla gestione split-brain.

Cosa succede se il collegamento di replica viene perso e nessuno dei due siti ha un quorum? Che cosa dovrebbe accadere? Questa domanda è valida sia per Oracle RAC che per il comportamento di ONTAP. Se non è possibile replicare le modifiche tra i siti e si desidera riprendere le operazioni, uno dei siti dovrà sopravvivere e l'altro sito dovrà diventare non disponibile.

La "[Mediatore ONTAP](#)" soddisfa questo requisito al livello ONTAP. Esistono diverse opzioni per RAC Tiebreaking.

Gli Oracle Tiebreaker

Il metodo migliore per gestire i rischi di Oracle RAC split-brain consiste nell'utilizzare un numero dispari di nodi RAC, preferibilmente utilizzando un Tiebreaker a 3rd siti. Se un sito 3rd non è disponibile, l'istanza di Tiebreaker potrebbe essere posizionata in un sito dei due siti, designando effettivamente un sito di sopravvivenza preferito.

Oracle e css_Critical

Con un numero pari di nodi, il comportamento predefinito di Oracle RAC è che uno dei nodi nel cluster sarà considerato più importante degli altri nodi. Il sito con tale nodo con priorità più alta sopravviverà all'isolamento del sito, mentre i nodi sull'altro sito verranno eliminati. La priorità si basa su più fattori, ma è anche possibile

controllare questo comportamento utilizzando l' `css_critical` impostazione.

Nell' "esempio" architettura, i nomi host per i nodi RAC sono jfs12 e jfs13. Di seguito sono riportate le impostazioni correnti di `css_critical`:

```
[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.
```

Se si desidera che il sito con jfs12 sia il sito preferito, impostare questo valore su sì su un nodo del sito A e riavviare i servizi.

```
[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.
```

Informazioni sul copyright

Copyright © 2024 NetApp, Inc. Tutti i diritti riservati. Stampato negli Stati Uniti d'America. Nessuna porzione di questo documento soggetta a copyright può essere riprodotta in qualsiasi formato o mezzo (grafico, elettronico o meccanico, inclusi fotocopie, registrazione, nastri o storage in un sistema elettronico) senza previo consenso scritto da parte del detentore del copyright.

Il software derivato dal materiale sottoposto a copyright di NetApp è soggetto alla seguente licenza e dichiarazione di non responsabilità:

IL PRESENTE SOFTWARE VIENE FORNITO DA NETAPP "COSÌ COM'È" E SENZA QUALSIVOGLIA TIPO DI GARANZIA IMPLICITA O ESPRESSA FRA CUI, A TITOLO ESEMPLIFICATIVO E NON ESAUSTIVO, GARANZIE IMPLICITE DI COMMERCIALIZZABILITÀ E IDONEITÀ PER UNO SCOPO SPECIFICO, CHE VENGONO DECLINATE DAL PRESENTE DOCUMENTO. NETAPP NON VERRÀ CONSIDERATA RESPONSABILE IN ALCUN CASO PER QUALSIVOGLIA DANNO DIRETTO, INDIRETTO, ACCIDENTALE, SPECIALE, ESEMPLARE E CONSEGUENZIALE (COMPRESI, A TITOLO ESEMPLIFICATIVO E NON ESAUSTIVO, PROCUREMENT O SOSTITUZIONE DI MERCI O SERVIZI, IMPOSSIBILITÀ DI UTILIZZO O PERDITA DI DATI O PROFITTI OPPURE INTERRUZIONE DELL'ATTIVITÀ AZIENDALE) CAUSATO IN QUALSIVOGLIA MODO O IN RELAZIONE A QUALUNQUE TEORIA DI RESPONSABILITÀ, SIA ESSA CONTRATTUALE, RIGOROSA O DOVUTA A INSOLVENZA (COMPRESA LA NEGLIGENZA O ALTRO) INSORTA IN QUALSIASI MODO ATTRAVERSO L'UTILIZZO DEL PRESENTE SOFTWARE ANCHE IN PRESENZA DI UN PREAVVISO CIRCA L'EVENTUALITÀ DI QUESTO TIPO DI DANNI.

NetApp si riserva il diritto di modificare in qualsiasi momento qualunque prodotto descritto nel presente documento senza fornire alcun preavviso. NetApp non si assume alcuna responsabilità circa l'utilizzo dei prodotti o materiali descritti nel presente documento, con l'eccezione di quanto concordato espressamente e per iscritto da NetApp. L'utilizzo o l'acquisto del presente prodotto non comporta il rilascio di una licenza nell'ambito di un qualche diritto di brevetto, marchio commerciale o altro diritto di proprietà intellettuale di NetApp.

Il prodotto descritto in questa guida può essere protetto da uno o più brevetti degli Stati Uniti, esteri o in attesa di approvazione.

LEGENDA PER I DIRITTI SOTTOPOSTI A LIMITAZIONE: l'utilizzo, la duplicazione o la divulgazione da parte degli enti governativi sono soggetti alle limitazioni indicate nel sottoparagrafo (b)(3) della clausola Rights in Technical Data and Computer Software del DFARS 252.227-7013 (FEB 2014) e FAR 52.227-19 (DIC 2007).

I dati contenuti nel presente documento riguardano un articolo commerciale (secondo la definizione data in FAR 2.101) e sono di proprietà di NetApp, Inc. Tutti i dati tecnici e il software NetApp forniti secondo i termini del presente Contratto sono articoli aventi natura commerciale, sviluppati con finanziamenti esclusivamente privati. Il governo statunitense ha una licenza irrevocabile limitata, non esclusiva, non trasferibile, non cedibile, mondiale, per l'utilizzo dei Dati esclusivamente in connessione con e a supporto di un contratto governativo statunitense in base al quale i Dati sono distribuiti. Con la sola esclusione di quanto indicato nel presente documento, i Dati non possono essere utilizzati, divulgati, riprodotti, modificati, visualizzati o mostrati senza la previa approvazione scritta di NetApp, Inc. I diritti di licenza del governo degli Stati Uniti per il Dipartimento della Difesa sono limitati ai diritti identificati nella clausola DFARS 252.227-7015(b) (FEB 2014).

Informazioni sul marchio commerciale

NETAPP, il logo NETAPP e i marchi elencati alla pagina <http://www.netapp.com/TM> sono marchi di NetApp, Inc. Gli altri nomi di aziende e prodotti potrebbero essere marchi dei rispettivi proprietari.