



資料庫組態

Enterprise applications

NetApp
May 09, 2024

目錄

資料庫組態	1
Oracle 資料庫區塊大小	1
Oracle 資料庫參數：DB_FILE_Multifblock_read_count	1
Oracle 資料庫參數：filesystemio_options	2
Oracle Real Application Clusters (RAC) 逾時	3

資料庫組態

Oracle 資料庫區塊大小

ONTAP 內部使用可變的區塊大小、這表示 Oracle 資料庫可以設定任何所需的區塊大小。不過、檔案系統區塊大小可能會影響效能、在某些情況下、較大的重做區塊大小可能會改善效能。

資料檔案區塊大小

部分作業系統提供多種檔案系統區塊大小選擇。對於支援 Oracle 資料檔案檔案的檔案系統、使用壓縮時區塊大小應為 8KB。不需要壓縮時、可以使用 8KB 或 4KB 的區塊大小。

如果資料檔案放在具有 512 位元組區塊的檔案系統上、則可能會有未對齊的檔案。根據 NetApp 建議、LUN 和檔案系統可能已正確對齊、但檔案 I/O 可能未對齊。這種錯誤的調整會導致嚴重的效能問題。

支援重做記錄的檔案系統必須使用重做區塊大小的倍數。這通常需要重做記錄檔系統和重做記錄本身都使用 512 位元組的區塊大小。

重做區塊大小

重做率極高時、4KB 區塊大小可能會執行得更好、因為重做率較高、可在較少且更有效率的作業中執行 I/O。如果重做速率大於 50Mbps、請考慮測試 4KB 區塊大小。

在具有 4KB 區塊大小和許多極小型交易的檔案系統上、使用 512 位元組區塊大小的重做記錄檔來識別資料庫中的一些客戶問題。將多個 512 位元組的變更套用到單一 4KB 檔案系統區塊所涉及的額外負荷、導致效能問題、而這些問題已透過將檔案系統變更為使用 512 位元組的區塊大小來解決。



* NetApp 建議 * 除非相關客戶支援或專業服務組織建議您變更重做區塊大小、否則請勿變更重做區塊大小、否則變更將以正式產品文件為基礎。

Oracle 資料庫參數：DB_FILE_Multifblock_read_count

。db_file_multiblock_read_count 參數控制 Oracle 在連續 I/O 期間讀取為單一作業的 Oracle 資料庫區塊數量上限

不過、此參數不會影響 Oracle 在任何及所有讀取作業期間讀取的區塊數、也不會影響隨機 I/O 只有連續 I/O 的區塊大小會受到影響。

Oracle 建議使用者不要設定此參數。如此可讓資料庫軟體自動設定最佳值。這通常表示此參數設為可產生 1MB I/O 大小的值。例如、1MB 讀取 8KB 區塊需要 128 個區塊才能讀取、因此此參數的預設值為 128。

NetApp 在客戶站台上觀察到的大多數資料庫效能問題、都涉及此參數的設定不正確。使用 Oracle 版本 8 和 9 變更此值的理由是正確的。因此、參數可能會在不知情的情況下出現在中 init.ora 檔案、因為資料庫已就地升級至 Oracle 10 及更新版本。傳統設定為 8 或 16、而預設值為 128、會大幅損害連續 I/O 效能。



* NetApp 建議 * 設定 `db_file_multiblock_read_count` 參數不應出現在 `init.ora` 檔案：NetApp 從未遇到過變更此參數可改善效能的情況、但在許多情況下、它會對連續 I/O 處理量造成明顯損害。

Oracle 資料庫參數： `filesystemio_options`

Oracle 初始化參數 `filesystemio_options` 控制非同步和直接 I/O 的使用

與一般的看法相反、非同步和直接 I/O 並不相互排斥。NetApp 發現、在客戶環境中、此參數經常設定錯誤、而這種錯誤設定直接導致許多效能問題。

非同步 I/O 表示 Oracle I/O 作業可以平行化。在各種作業系統上均可使用非同步 I/O 之前、使用者已設定數個 `dbwriter` 程序、並變更伺服器程序組態。透過非同步 I/O、作業系統本身就能以高效率且平行的方式代表資料庫軟體執行 I/O。此程序不會讓資料面臨風險、而且關鍵作業（例如 Oracle 重做記錄）仍會同步執行。

直接 I/O 會略過作業系統緩衝區快取。UNIX 系統上的 I/O 通常會流經作業系統緩衝區快取。這對不維護內部快取的應用程式很有用、但 Oracle 在 SGA 中擁有自己的緩衝區快取。在幾乎所有情況下、最好是啟用直接 I/O 並將伺服器 RAM 分配給 SGA、而非仰賴作業系統緩衝區快取。Oracle SGA 更有效率地使用記憶體。此外、當 I/O 流經作業系統緩衝區時、它會受到額外處理、因此會增加延遲。當低延遲是關鍵需求時、在大量寫入 I/O 時、延遲特別明顯。

的選項 `filesystemio_options` 是：

- * 非同步 *。* Oracle 將 I/O 要求提交給作業系統以進行處理。此程序可讓 Oracle 執行其他工作、而非等待 I/O 完成、進而增加 I/O 平行化。
- **directio**。Oracle 直接針對實體檔案執行 I/O、而非透過主機作業系統快取來路由 I/O。
- * 無 *。* Oracle 使用同步和緩衝 I/O 在此組態中、選擇共享與專用伺服器程序與 `dbWriters` 數量更為重要。
- **setall**。Oracle 同時使用非同步和直接 I/O 在幾乎所有情況下、都是使用 `setall` 是最佳的。



◦ `filesystemio_options` 參數在 DNFS 和 ASM 環境中無效。使用 DNFS 或 ASM 時、會自動同時使用非同步和直接 I/O

有些客戶過去曾遇到非同步 I/O 問題、尤其是先前的 Red Hat Enterprise Linux 4 (RHEL4) 版本。網際網路上有些過時的建議仍建議避免非同步 IO、因為資訊過時。在所有目前的作業系統上、非同步 I/O 都是穩定的。沒有理由停用它、作業系統沒有已知的錯誤。

如果資料庫已使用緩衝 I/O、則直接 I/O 的交換器也可能需要變更 SGA 大小。停用緩衝 I/O 可消除主機作業系統快取為資料庫提供的效能優勢。將 RAM 新增回 SGA 可修復此問題。最終結果應該是 I/O 效能的改善。

雖然 Oracle SGA 使用 RAM 幾乎比使用 OS 緩衝區快取更好、但可能無法判斷最佳值。例如、最好在資料庫伺服器上使用具有極小型 SGA 大小的緩衝 I/O、其中有許多間歇性作用中的 Oracle 執行個體。這種配置可讓所有執行中的資料庫執行個體靈活使用作業系統上的剩餘可用 RAM。這是非常不尋常的情況、但在某些客戶據點已發現這種情況。



* NetApp 建議 * 設定 `filesystemio_options` 至 `'setall'` 但請注意、在某些情況下、遺失主機緩衝區快取可能需要增加 Oracle SGA。

Oracle Real Application Clusters (RAC) 逾時

Oracle RAC 是一款叢集軟體產品、內含多種類型的內部活動訊號處理程序、可監控叢集的健全狀況。



中的資訊 "遺失計數" 本節包含使用網路儲存設備的 Oracle RAC 環境的重要資訊、在許多情況下、需要變更預設 Oracle RAC 設定、以確保 RAC 叢集在網路路徑變更和儲存設備容錯移轉 / 切換作業之後仍能順利運作。

磁碟逾時

主要儲存相關 RAC 參數為 `disktimeout`。此參數控制投票檔案 I/O 必須完成的臨界值。如果是 `disktimeout` 超過參數、RAC 節點就會從叢集中移出。此參數的預設值為 200。此值應足以用於標準儲存設備接管和恢復程序。

NetApp 強烈建議您在將 RAC 組態投入生產之前、先徹底測試這些組態、因為許多因素會影響接管或恢復作業。除了完成儲存容錯移轉所需的時間之外、連結集合化控制傳輸協定 (LACP) 變更也需要額外的時間才能傳播。此外、SAN 多重路徑軟體必須偵測 I/O 逾時、然後在替代路徑上重試。如果資料庫處於極活躍狀態、則必須在處理投票磁碟 I/O 之前、先佇列並重新嘗試大量 I/O。

如果無法執行實際的儲存接管或恢復、則可以在資料庫伺服器上進行纜線拉出測試來模擬影響。



- NetApp 建議 * 下列事項：
- 離開 `disktimeout` 預設值為 200 的參數。
- 務必徹底測試 RAC 組態。

遺失計數

◦ `misscount` 參數通常只會影響 RAC 節點之間的網路心跳。預設值為 30 秒。如果網格二進位檔位於儲存陣列上、或作業系統開機磁碟機不是本機磁碟機、此參數可能會變得很重要。這包括 FC SAN 上具有開機磁碟機的主機、NFS 開機作業系統、以及位於虛擬化資料存放區 (例如 VMDK 檔案) 上的開機磁碟機。

如果因儲存接管或恢復而中斷開機磁碟機的存取、網格二進位位置或整個作業系統可能會暫時停止運作。ONTAP 完成儲存作業所需的時間、以及作業系統變更路徑和恢復 I/O 所需的時間、可能會超過 `misscount` 臨界值。因此、節點會在連線到開機 LUN 或網格二進位檔恢復後立即停止。在大多數情況下、會發生遷離和後續重新開機、而不會出現記錄訊息來指出重新開機的原因。並非所有組態都會受到影響、因此請在 RAC 環境中測試任何 SAN 開機、NFS 開機或資料存放區型主機、以便在與開機磁碟機的通訊中斷時、RAC 保持穩定。

非本機開機磁碟機或非本機檔案系統代管的情況 `grid` 二進位檔案 `misscount` 需要變更以符合 `disktimeout`。如果變更此參數、請進行進一步測試、以識別對 RAC 行為的任何影響、例如節點容錯移轉時間。



- NetApp 建議 * 下列事項：
- 離開 `misscount` 參數的預設值為 30、除非符合下列其中一項條件：
 - `grid` 二進位檔位於網路附加磁碟機上、包括 NFS、iSCSI、FC 和資料存放區型磁碟機。
 - 作業系統是 SAN 開機。
- 在這種情況下、請評估網路中斷對 OS 或的存取造成的影響 `GRID_HOME` 檔案系統。在某些情況下、這類中斷會導致 Oracle RAC 精靈停止運作、進而導致 `misscount` 根據的逾時和遷離。逾時預設為 27 秒、即的值 `misscount` 減號 `reboottime`。在這種情況下、請增加 `misscount` 200 比對 `disktimeout`。

版權資訊

Copyright © 2024 NetApp, Inc. 版權所有。台灣印製。非經版權所有人事先書面同意，不得將本受版權保護文件的任何部分以任何形式或任何方法（圖形、電子或機械）重製，包括影印、錄影、錄音或儲存至電子檢索系統中。

由 NetApp 版權資料衍伸之軟體必須遵守下列授權和免責聲明：

此軟體以 NETAPP「原樣」提供，不含任何明示或暗示的擔保，包括但不限於有關適售性或特定目的適用性之擔保，特此聲明。於任何情況下，就任何已造成或基於任何理論上責任之直接性、間接性、附隨性、特殊性、懲罰性或衍生性損害（包括但不限於替代商品或服務之採購；使用、資料或利潤上的損失；或企業營運中斷），無論是在使用此軟體時以任何方式所產生的契約、嚴格責任或侵權行為（包括疏忽或其他）等方面，NetApp 概不負責，即使已被告知有前述損害存在之可能性亦然。

NetApp 保留隨時變更本文所述之任何產品的權利，恕不另行通知。NetApp 不承擔因使用本文所述之產品而產生的責任或義務，除非明確經過 NetApp 書面同意。使用或購買此產品並不會在依據任何專利權、商標權或任何其他 NetApp 智慧財產權的情況下轉讓授權。

本手冊所述之產品受到一項（含）以上的美國專利、國外專利或申請中專利所保障。

有限權利說明：政府機關的使用、複製或公開揭露須受 DFARS 252.227-7013（2014 年 2 月）和 FAR 52.227-19（2007 年 12 月）中的「技術資料權利 - 非商業項目」條款 (b)(3) 小段所述之限制。

此處所含屬於商業產品和 / 或商業服務（如 FAR 2.101 所定義）的資料均為 NetApp, Inc. 所有。根據本協議提供的所有 NetApp 技術資料和電腦軟體皆屬於商業性質，並且完全由私人出資開發。美國政府對於該資料具有非專屬、非轉讓、非轉授權、全球性、有限且不可撤銷的使用權限，僅限於美國政府為傳輸此資料所訂合約所允許之範圍，並基於履行該合約之目的方可使用。除非本文另有規定，否則未經 NetApp Inc. 事前書面許可，不得逕行使用、揭露、重製、修改、履行或展示該資料。美國政府授予國防部之許可權利，僅適用於 DFARS 條款 252.227-7015(b)（2014 年 2 月）所述權利。

商標資訊

NETAPP、NETAPP 標誌及 <http://www.netapp.com/TM> 所列之標章均為 NetApp, Inc. 的商標。文中所涉及的所有其他公司或產品名稱，均為其各自所有者的商標，不得侵犯。