



# 儲存組態

## Enterprise applications

NetApp  
May 03, 2024

# 目錄

儲存組態 .....	1
具有 NFS 檔案系統的 PostgreSQL 資料庫 .....	1
具有 SAN 檔案系統的 PostgreSQL .....	2

# 儲存組態

## 具有 NFS 檔案系統的 PostgreSQL 資料庫

PostgreSQL 資料庫可以裝載在 NFSv3 或 NFSv4 檔案系統上。最佳選項取決於資料庫以外的因素。

例如、在某些叢集式環境中、NFSv4 鎖定行為可能較為理想。（請參閱 ["請按這裡"](#) 如需其他詳細資料）

否則、資料庫功能應接近相同、包括效能。唯一的要求是使用 `hard` 掛載選項。這是為了確保軟性逾時不會產生無法恢復的 IO 錯誤。

如果選擇 NFSv4 作為傳輸協定、NetApp 建議使用 NFSv4.1。NFSv4.1 中的 NFSv4 傳輸協定有一些功能性增強功能、可改善 NFSv4.0 的恢復能力。

針對一般資料庫工作負載、請使用下列掛載選項：

```
rw,hard,nointr,bg,vers=[3|4],proto=tcp,rsize=65536,wsize=65536
```

如果需要大量連續 IO、則可依照下節所述增加 NFS 傳輸大小。

### NFS 傳輸大小

根據預設、ONTAP 將 NFS I/O 大小限制為 64K。

大多數應用程式和資料庫的隨機 I/O 使用的區塊大小要小得多、遠低於 64K 的最大值。大型區塊 I/O 通常是平行處理的、因此 64K 的最大值也不是取得最大頻寬的限制。

有些工作負載的上限為 64K、因此會造成限制。特別是、如果資料庫執行的 I/O 數量較少、但容量較大、則備份或還原作業或資料庫完整表格掃描等單執行緒作業、會更快、更有效率地執行。ONTAP 的最佳 I/O 處理大小為 256k。

指定 ONTAP SVM 的最大傳輸大小可變更如下：

```
Cluster01::> set advanced
Warning: These advanced commands are potentially dangerous; use them only
when directed to do so by NetApp personnel.
Do you want to continue? {y|n}: y
Cluster01::*> nfs server modify -vserver vserver1 -tcp-max-xfer-size
262144
Cluster01::*>
```

## 注意

切勿將 ONTAP 上允許的傳輸大小上限降至低於目前掛載之 NFS 檔案系統的 rsize/wsize 值。這可能會在某些作業系統中造成當機或甚至資料毀損。例如、如果 NFS 用戶端目前設定為 rsize/wsize 65536、則 ONTAP 最大傳輸大小可在 65536 到 1048576 之間調整、因為用戶端本身受到限制、因此沒有任何影響。將傳輸大小上限降至 65536 以下可能會損害可用度或資料。

在 ONTAP 層級增加傳輸大小後、將會使用下列掛載選項：

```
rw,hard,nointr,bg,vers=[3|4],proto=tcp,rsize=262144,wsiz=262144
```

## NFSv3 TCP 插槽表

如果 NFSv3 與 Linux 搭配使用、則正確設定 TCP 插槽表至關重要。

TCP 插槽表是與主機匯流排介面卡（HBA）佇列深度相當的 NFSv3。這些表格可控制任何時間都可以處理的 NFS 作業數量。預設值通常為 16、這對於最佳效能而言太低。相反的問題發生在較新的 Linux 核心上、這會自動將 TCP 插槽表格限制增加到要求使 NFS 伺服器飽和的層級。

為了達到最佳效能並避免效能問題、請調整控制 TCP 插槽表的核心參數。

執行 `sysctl -a | grep tcp.*.slot_table` 並觀察下列參數：

```
# sysctl -a | grep tcp.*.slot_table
sunrpc.tcp_max_slot_table_entries = 128
sunrpc.tcp_slot_table_entries = 128
```

所有 Linux 系統都應該包括在內 `sunrpc.tcp_slot_table_entries`、但只有部分包含在內 `sunrpc.tcp_max_slot_table_entries`。兩者都應設為 128。

## 注意

若未設定這些參數、可能會對效能造成重大影響。在某些情況下、效能會受到限制、因為 Linux 作業系統沒有發出足夠的 I/O 在其他情況下、隨著 Linux 作業系統嘗試發出的 I/O 數量超過可服務的數量、I/O 延遲也會增加。

## 具有 SAN 檔案系統的 PostgreSQL

具有 SAN 的 PostgreSQL 資料庫通常裝載在 xfs 檔案系統上、但如果作業系統廠商支援、也可以使用其他資料庫

雖然單一 LUN 通常可支援高達十萬 IOPS、但 IO 密集的資料庫通常需要使用含分段的 LVM。

## LVM 分拆

在快閃磁碟機時代之前、使用區塊延展來協助克服旋轉磁碟機的效能限制。例如、如果作業系統需要執行 1MB 讀取作業、則從單一磁碟機讀取 1MB 的資料時、需要大量的磁碟機磁頭搜尋和讀取、因為 1MB 會緩慢傳輸。

如果將 1MB 的資料分散在 8 個 LUN 上、則作業系統可能會同時執行 8 個 128K 讀取作業、並縮短完成 1MB 傳輸所需的時間。

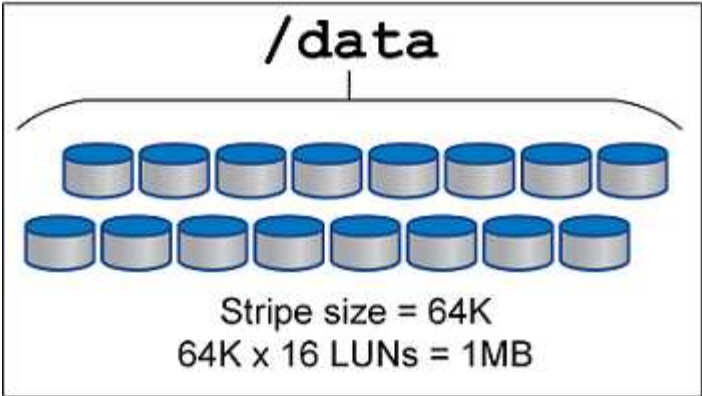
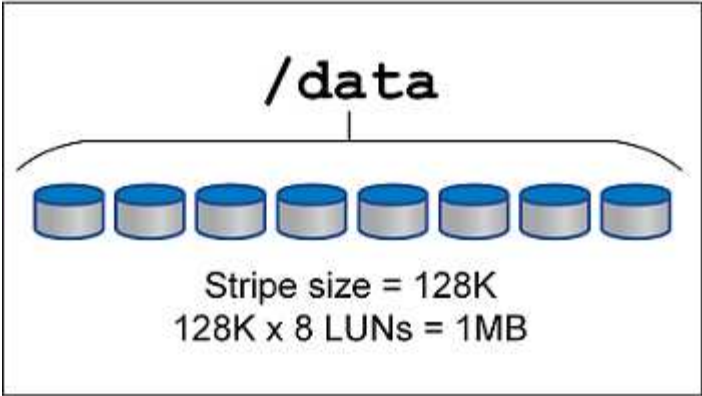
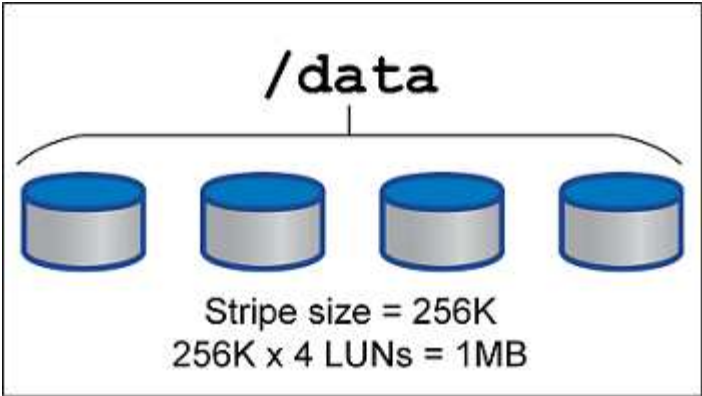
由於必須事先知道 I/O 模式、因此使用旋轉磁碟機進行分拆會更困難。如果串列區塊延展未針對真正的 I/O 模式正確調整、則等量區塊配置可能會損害效能。使用 Oracle 資料庫、特別是搭配 All Flash 組態、分拆作業更容易設定、並經證實可大幅提升效能。

依預設、邏輯磁碟區管理程式（例如 Oracle ASM 等量磁碟區）不屬於原生 OS LVM。其中有些 LUN 會將多個 LUN 連結在一起、成為串連的裝置、導致資料檔案存在於一台 LUN 裝置上、而只存在於一台 LUN 裝置上。這會造成熱點。其他 LVM 實作預設為分散式擴充。這與分拆類似、但卻是比較粗糙的。磁碟區群組中的 LUN 會切成大型片段、稱為區段、通常以百萬位元組為單位測量、然後邏輯磁碟區會分佈在這些區段中。結果是對檔案進行隨機 I/O、應該能在 LUN 之間妥善分配、但連續 I/O 作業的效率卻不如以前那麼高。

效能密集的應用程式 I/O 幾乎總是（a）以基本區塊大小為單位、或（b）1 MB。

等量分配組態的主要目標是確保單一檔案 I/O 可作為單一單元執行、而多區塊 I/O 的大小應為 1MB、可在等量磁碟區中的所有 LUN 之間平均平行處理。這表示等量磁碟區大小不得小於資料庫區塊大小、且等量磁碟區大小乘以 LUN 數量應為 1MB。

下圖顯示等量磁碟區大小和寬度調校的三個可能選項。選擇 LUN 數量以滿足上述效能需求、但在所有情況下、單一等量磁碟區內的總資料為 1MB。



## 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。