



儲存組態

Enterprise applications

NetApp
May 09, 2024

目錄

儲存組態	1
使用 NFS 的 MySQL	1
MySQL 搭配 SAN	1

儲存組態

使用 NFS 的 MySQL

MySQL 文件建議您將 NFSv4 用於 NAS 部署。

ONTAP NFS 傳輸大小

根據預設、ONTAP 將 NFS IO 大小限制為 64K。使用 MySQL 資料庫的隨機 IO 使用的區塊大小要小得多、遠低於 64K 上限。大型區塊 IO 通常是平行處理的、因此 64K 最大值也不是限制。

有些工作負載的上限為 64K、因此會造成限制。尤其是、如果資料庫執行的 IO 數量較少、但容量較大、例如完整表格掃描備份作業等單執行緒作業、將會更快、更有效率地執行。ONTAP 搭配資料庫工作負載的最佳 IO 處理大小為 256k。針對下列特定作業系統所列出的 NFS 裝載選項已相應從 64K 更新至 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
```



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。

• NetApp 推薦 *



設定下列 NFSv4 Fstab (/etc/fstab) 設定：

```
nfs4 rw,  
hard,nointr,bg,vers=4,proto=tcp,noatime,rsize=262144,wsizer=262144
```



NFSv3 的常見問題是停電後鎖定的 InnoDB 記錄檔。使用時間或切換記錄檔可解決此問題。不過、NFSv4 有鎖定作業、並追蹤開啟的檔案和委派。

MySQL 搭配 SAN

有兩個選項可以使用常見的雙磁碟區模型、來設定 MySQL 與 SAN。

只要 I/O 和容量需求在單一 LUN 檔案系統的限制範圍內、就可以將較小的資料庫放在一對標準 LUN 上。例如、需要大約 2K 隨機 IOPS 的資料庫、可以裝載在單一 LUN 上的單一檔案系統上。同樣地、只有 100GB 大小的資料庫也能放在單一 LUN 上、而不會造成管理問題。

大型資料庫需要多個 LUN。例如、需要 10 萬 IOPS 的資料庫最可能需要至少八個 LUN。由於磁碟機的 SCSI 通道數量不足、單一 LUN 將成為瓶頸。10TB 資料庫同樣難以在單一 10TB LUN 上進行管理。邏輯磁碟區管理程式的設計旨在將多個 LUN 的效能和容量功能結合在一起、以改善效能和管理能力。

在這兩種情況下、一對 ONTAP 磁碟區應該足夠。透過簡單的組態、資料檔案 LUN 會與記錄 LUN 一樣置於專用磁碟區中。在邏輯 Volume Manager 組態下、資料檔案 Volume 群組中的所有 LUN 都將位於專用磁碟區、而記錄 Volume 群組的 LUN 則位於第二個專用磁碟區。

- NetApp 建議 * 在 SAN 上使用兩個檔案系統進行 MySQL 部署：
- 第一個檔案系統會儲存所有 MySQL 資料、包括資料表空間、資料和索引。
- 第二個檔案系統會儲存所有記錄（二進位記錄、慢速記錄和交易記錄）。



以這種方式分隔資料有多種原因、包括：

- 資料檔案和記錄檔的 I/O 模式各不相同。將它們區隔、就能透過 QoS 控制提供更多選項。
- 最佳化使用 Snapshot 技術需要能夠個別還原資料檔案。將資料檔案與記錄檔混合會干擾資料檔案還原。
- NetApp SnapMirror 技術可用於為資料庫提供簡單、低 RPO 的災難恢復功能、但是資料檔案和記錄檔需要不同的複寫排程。



使用這種基本的兩個磁碟區配置、以符合未來需求的解決方案、以便在需要時使用所有 ONTAP 功能。

- NetApp 建議 * 使用 ext4 檔案系統格式化磁碟機、因為有下列功能：
- 延伸的區塊管理方法、用於日誌檔案系統（JFS）中的區塊管理功能、以及延伸檔案系統（XFS）的延遲分配功能。
- ext4 允許檔案系統最多 1 個 exbibyte（ 2^{60} 位元組）、檔案最多 16 個 tebibytes（ $16 * 2^{40}$ 位元組）。相反地、ext3 檔案系統僅支援 16TB 的最大檔案系統大小、最大檔案大小為 2TB。
- 在 ext4 檔案系統中、多區塊分配（mballoc）會在單一作業中為檔案分配多個區塊、而非如 ext3 中逐一分配。此組態可減少多次呼叫區塊分配器的成本、並最佳化記憶體配置。
- 雖然 XFS 是許多 Linux 套裝作業系統的預設值、但它以不同方式管理中繼資料、不適合某些 MySQL 組態。



- NetApp 建議 * 將 4K 區塊大小選項與 mkfs 公用程式搭配使用、以符合現有區塊 LUN 大小。



```
mkfs.ext4 -b 4096
```

NetApp LUN 以 4KB 實體區塊儲存資料、產生八個 512 位元組的邏輯區塊。

如果您未設定相同的區塊大小、I/O 將無法正確對齊實體區塊、而且可能會在 RAID 群組中的兩個不同磁碟機中寫入、導致延遲。



請務必調整 I/O、以順利進行讀寫作業。但是、當 I/O 從不在實體區塊開頭的邏輯區塊開始時、I/O 會未對齊。I/O 作業只有在邏輯區塊（實體區塊中的第一個邏輯區塊）開始時才會對齊。

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。