XCubeSAN

RAIDEE

QSAN

www.QSAN.com

版權

©版權所有 2019 QSAN廣盛科技保留所有權利。未經 QSAN廣盛科技書面許可，不得複製或傳播本文件的任何部分。

2019 年 9 月

此版本適用於 QSAN XCubeSAN系列。QSAN認為本出版物在發布之日內容準確無誤。資訊如有更改，恕不另行通知。

商標

QSAN、QSAN標誌、XCubeSAN和 QSAN.com是 QSAN廣盛科技的商標或註冊商標。

Microsoft 、Windows、Windows Server和 Hyper-V 是 Microsoft Corporation 在美國和/或其他國家/地區的商標或註冊商標。

Linux 是 Linus Torvalds 在美國和/或其他國家/地區的商標。

UNIX是 The Open Group在美國和其他國家/[illegible]

Mac OS X Apple Inc. /

Java Java Oracle /

VMware ESXi vSphere是 VMware, Inc / /

Citrix Xen Citrix Systems, Inc / /

注意

此 XCubeSAN系列技術白皮書適用於以下 XCubeSAN型號：

XCubeSAN儲存系統 4U 19吋 機架式機箱型號

型號名稱	控制器類型	外形、槽位計數和機架單元
XS5224D	雙控制器	LFF 24磁碟 4U 機箱
XS3224D	雙控制器	LFF 24磁碟 4U 機箱
XS3224S	單控制器	LFF 24磁碟 4U 機箱
XS1224D	雙控制器	LFF 24磁碟 4U 機箱
XS1224S	單控制器	LFF 24磁碟 4U 機箱

XCubeSAN儲存系統 3U 19吋 機架式機箱型號

型號名稱	控制器類型	外形、槽位計數和機架單元
XS5216D	雙控制器	LFF16 磁碟 3U 機箱
XS3216D	雙控制器	LFF16 磁碟 3U 機箱
XS3216S	單控制器	LFF16 磁碟 3U 機箱
XS1216D	雙控制器	LFF16 磁碟 3U 機箱
XS1216S	單控制器	LFF16 磁碟 3U 機箱

XCubeSAN儲存系統 2U 19吋 機架式機箱型號

型號名稱	控制器類型	外形、槽位計數和機架單元
XS5212D	雙控制器	LFF12 磁碟 2U 機箱
XS5212S	單控制器	LFF 12磁碟 2U 機箱
XS3212D	雙控制器	LFF 12磁碟 2U 機箱

XS3212S	單控制器	LFF 12磁碟 2U 機箱
XS1212D	雙控制器	LFF 12磁碟 2U 機箱
XS1212S	單控制器	LFF 12磁碟 2U 機箱
XS5226D	雙控制器	SFF 26 磁碟 2U 機箱
XS5226S	單控制器	SFF 26磁碟 2U 機箱
XS3226D	雙控制器	SFF 26磁碟 2U 機箱
XS3226S	單控制器	SFF 26磁碟 2U 機箱
XS1226D	雙控制器	SFF 26磁碟 2U 機箱
XS1226S	單控制器	SFF 26磁碟 2U 機箱

文件中所包含資訊的準確性已被審查。但它可能包括印刷錯誤或技術不準確，這將定期對文件進行更改，而這些更改將納入該出版物的新版本。QSAN可能會對產品進行改進或更改，所有功能和產品規格如有更改，恕不另行通知或承擔義務。本文件中的所有陳述、資訊和建議均不構成任何明示或暗示的擔保。

此處包含的任何效能資料都是在受控環境中確定的。因此，在其他作業環境中獲得的結果可能會有很大差異。在開發級系統上進行

，　　，

子包括個人、公司、品牌和產品的名稱。所有這些名稱都是虛構的，與實際商業企業使用的名稱和地址如有任何相似之處完全是巧合。

 © 版權所有 2019 QSAN廣盛科技保留所有權利。

目錄

注意 i

RAID EE 技術 1

執行摘要 1

讀者 1

概觀 1

運作原理 3

配置 RAID EE儲存池 8

建立 RAID EE儲存池 8

RAID EE儲存池 14

RAID EE 17

測試結果 20

測試案例 1：RAID 5與 RAID 5EE 20

測試案例 2：RAID 60與 RAID 60EE 23

結論 26

適用於 26

參考 26

附錄 28

相關文件 28

技術支援 28

RAID EE 技術

執行摘要

已經存在超過 30 年的 RAID架構正在經歷一波轉型。對於 TB 級大容量硬碟，原始 RAID技術無法解決重建時間過長的問題。基於傳統區塊技術的新一代 RAID技術，我們稱之為 RAID EE，這是被視為解決傳統 RAID缺陷的途徑。

資訊：

RAID EE技術在 SANOS韌體版本 1.3.0 中可用，並且在 SANOS韌體版本 1.4.1 中效能有大幅度的提升。

讀者

本文件適用於有興趣了解 RAID EE以解決重建時間過長問題的 QSAN客戶和合作夥伴。我們假定讀者熟悉 QSAN產品並具有一般 IT 經驗，包括作為系統或網路管理員的知識。如有任何疑問，請參閱產品的使用手冊，或聯絡 QSAN技術支援以獲得進一步的幫助。

概觀

RAID（獨立磁碟冗餘陣列）將基於特定演算法組合多個獨立的物理磁碟

碟 效能 RAID

碟對

及高效能應用的出現，傳統 RAID逐漸暴露出其缺陷。

隨著硬碟容量的增加，重建 RAID資料所需的時間也大大增加。這是當今企業儲存管理中最麻煩的問題之一。在過去幾天硬碟容量只有 10GB 到 100GB 的情況下，RAID重建的工作可以在 10 分鐘甚至 10 分鐘以上完成，這在沒有特別關注的情況下還不是問題。 但是，隨著磁碟容量增長到數百 GB 甚至 TB，RAID重建時間增加到數小時甚至數天，這成為儲存管理中的主要問題。

例如，一個傳統 RAID 5內含有 8 加 1 顆同位檢查的 6TB NL-SAS磁碟上需要2.5
資料 式系統的整體效能。如果使用者限制重建
優先權，則重建時間將更長。重要的是，在耗時的重建過程中，大量的存取操作可能導致儲存池中其他磁碟的故障，從而大大增加了磁碟故障的可能性和 丟失的風險。

傳統 RAID 架構的局限性

傳統的 RAID架構由一定數量的磁碟組成磁碟組 (也稱為 RAID組)
LUN

RAID

- y
  加載的資料集中在備援磁碟上形成瓶頸。
- y 其次，儲存卷資料存取僅限於屬於磁碟組的成員磁碟；這限制了主機的效能，因為儲存裝置正在執行存取和重建 I/O。

為什麼 RAID 重建耗費時間

隨著磁碟容量的增長，RAID重建時間呈線性增長，當使用具有 4TB 以上硬碟容量的 RAID磁碟時，將傳統 RAID架構所需的重建時間提高到數十小時。

有幾個因素會影響 RAID重建時間：

- y **硬碟容量：**硬碟容量構成磁碟組，硬碟容量越大，重建時間越長。
- y **磁碟數量：** 備援磁碟所需的時間。磁碟越多，重建時間越長。
- y **重建作業優先權：**在 RAID重建期間，系統仍需要假設對前端主機的 I/O 存取。分配給RAID重建作業的優先權越高，重建速度越快，但前端主機獲得 I/O 效能的能力越低。
- y **快速重建：**啟用快速重建功能只需要重建儲存卷的實際容量，未使用的磁碟組空間不用重建。如果儲存卷僅使用磁碟組中的部分空間，則將縮短重建時間。
- y RAID **等級：**具有直接區塊到區塊複製的 RAID 1和 RAID 10將比具有同位檢查計算的 RAID 5和 RAID 6重建地更快。

鑑於每個磁碟可能出現故障，磁碟組中包含的磁碟越多，累積故障的可能性就越大，因此磁碟組中的磁碟數量有上限的限制。與以前的因素相比，磁碟容量對重建速度的影響越來越大，這已成為主要因素。如此長的重建時間顯然是任何使用者都不能接受的。為了解決傳統 RAID的問題，我們實現了 RAID EE技術。

 © 版權所有 2019 QSAN廣盛科技保留所有權利。

運作原理

RAID EE在磁碟組中新增了更多備援磁碟，我們將其稱為 RAID EE **備援磁碟**，以分離原始的全域、本地和專用備援磁碟。[redacted] 磁碟
磁碟 [redacted]
的所有 都是重建資料的目標，因此傳統 RAID重建的瓶頸消失了，重建效能得到了顯著提升。如果新增了新 ，則備援區中的 將複製回新加入的 。

為 RAID EE提供了四個新的 RAID等級，有：

- RAID 5EE (E 代表增強型)，至少需要 4 顆磁碟和一顆 RAID EE備援 ，可以容忍 2 顆磁碟故障。新增更多 RAID EE備援磁碟將容忍更多磁碟故障。
- RAID 6EE 至少需要 5 顆磁碟。
- RAID 50EE 至少需要 7 顆磁碟。
- RAID 60EE 至少需要 9 顆磁碟。

資訊：

磁碟組中的 RAID EE備援磁碟數量為 1 到 8 顆磁碟。

帶有 1 顆 RAID EE 備援磁碟的 RAID 5EE 範例

舉一個例子來描述它是如何工作的。以下範例是具有 5 顆磁碟的 RAID 5EE，4 顆磁碟用於 RAID ，另外一顆磁碟用於 RAID EE備援磁碟。初始化後， 區塊分配如下。P 代表同位檢查，S 代表 RAID EE備援磁碟，現在它的內容是空的。

圖表 1　　RAID 5EE的資料區塊分佈

假設磁碟 2 發生故障。RAID 5EE處於降級模式。

圖表 2　　磁碟 2 已故障

故障　　中的資料將被重建回備援區域。此操作稱為 EE 重建。重建後，分佈的　　就像 RAID 5，它可以容忍另一個故障的磁碟。我們可以想像，RAID EE備援磁碟越多，重建速度

© 版權所有 2019 QSAN廣盛科技保留所有權利。

圖表 5　　RAID 60EE的資料區　分佈

RAID 60EE的重建和回拷與上述類似[redacted]

RAID EE

RAID EE

 © 版權所有 2019 QSAN廣盛科技保留所有權利。

配置 RAID EE 儲存池

本節將介紹配置 RAID EE儲存池的操作。

圖表 6　　儲存池功能子菜單

建立 RAID EE 儲存池

以下是建立配置 4 顆磁碟的 RAID 5EE儲存池範例。在首次建立儲存池時，包含一個磁碟組，且磁碟組中的最大　　數為 64。

1. 選擇**儲存池**功能子菜單，點擊**建立儲存池**按鈕。首先掃描可用磁碟。

提示：

200　　20 ~ 30

© 版權所有 2019 QSAN廣盛科技保留所有權利。

閒置

Access Patterns			
% of Access Specification	*Transfer Size Request*	*% Reads*	*% Random*
File Server Access Pattern (as defined by Intel)			
10%	*0.5 KB*	*80%*	*100%*
5%	*1 KB*	*80%*	*100%*
5%	*2 KB*	*80%*	*100%*
60%	*4 KB*	*80%*	*100%*
2%	*8 KB*	*80%*	*100%*
4%	*16 KB*	*80%*	*100%*
4%	*32 KB*	*80%*	*100%*
10%	*64 KB*	*80%*	*100%*
Workstation Access Pattern (as defined by StorageReview.com)			
100%	*8 KB*	*80%*	*80%*
Database Access Pattern (as defined by Intel/StorageReview.com)			
100%	*8 KB*	*67%*	*100%*

圖表 16　　由 StorageReview.com提供的存取

y

- 立　　RAID 5儲存　　顆磁碟

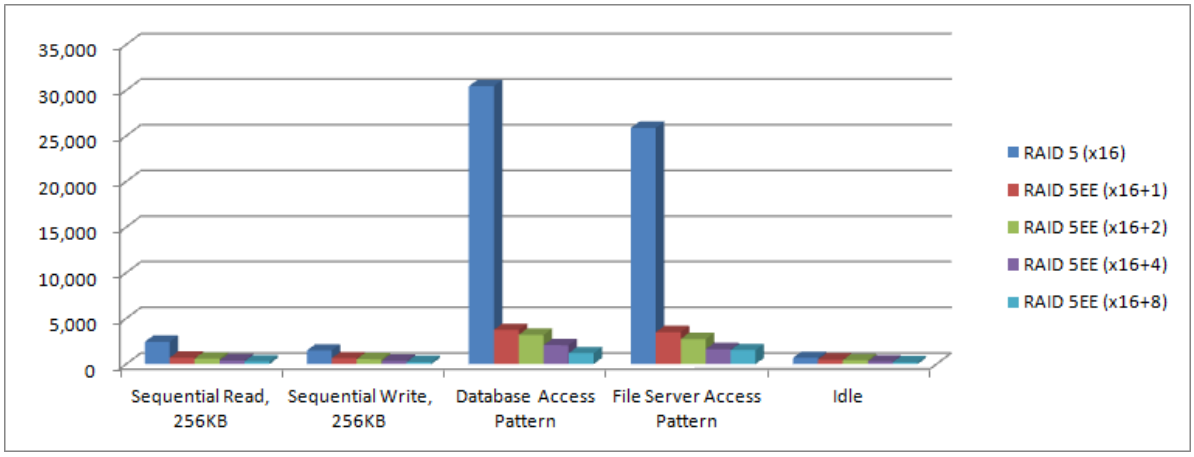
I/O

- 按順序建立帶有 1/2/4/8 x RAID EE 備援磁碟的 RAID 5EE。初始化後，拔出一顆磁碟，RAID EE開始重建。使用不同的 I/O 存取模式計算重建時間。然後插入一顆磁碟並將其設定為專用備援磁碟，它開始回拷。最後，計算回拷時間。

表格 5　　RAID 5和 RAID 5EE的測試結果

	RAID 5 (x16)	RAID 5EE (x16+1)	RAID 5EE (x16+2)	RAID 5EE (x16+4)	RAID 5EE (x16+8)
循序讀取， 256KB	40'43"	11'39"	10'09"	6'46"	4'58"
改善百分比		71%	75%	83%	88%
回拷時間		34'08"	30'57"	28'30"	30'56"
循序寫入， 256KB	24'54"	10'37"	9'08"	6'12"	4'02"
改善百分比		57%	63%	75%	84%
回拷時間		22'29"	30'23"	28'50"	31'54"

© 版權所有 2019 QSAN廣盛科技保留所有權利。



圖表 17　　RAID 5和 RAID 5EE的圖表

摘要

- RAIDEE可將重建時間提高多達 96%。
- 使用 RAID EE的備 磁碟越多，重建時間就越少。
- 當存取為
- 存取模式　　回拷　　。
- 當　　回拷　　I/O。

測試案例 2：RAID 60 與 RAID 60EE

此測試提供 RAID 60和 RAID 60EE　　RAID EE備 磁碟　　RAID 60EE

-
  - 　　RS700 X7/PS4(　　：Intel Xeon E5-2600 v2 /　　8GB)
    iSCSI　　：Intel 82574L Gigabit 網路連接
    作業系統： Windows Server 2012 R2
- 儲存設備
  - 型號：XCubeSAN XS5224D
    記憶體：16GB (2 x 8GB在插槽 1 和 3) 每個控制器
    韌體版本：1.4.1
    磁碟：24 x Seagate Constellation ES,ST500NM0001, 500GB,SAS 6Gb/s

- 硬碟儲存池：
  RAID 60 儲存池，控制器 1 中具有 16 x NL-SAS 硬碟
  RAID 60EE 儲存池，控制器 1 中具有 17 (16+1 x RAIDEE 備援磁碟) x NL-SAS 硬碟
  RAID 60EE 儲存池，控制器 1 中具有 18 (16+2 x RAIDEE 備援磁碟) x NL-SAS 硬碟
  RAID 60EE 儲存池，控制器 1 中具有 20 (16+4 x RAIDEE 備援磁碟) x NL-SAS 硬碟
  RAID 60EE 儲存池，控制器 1 中具有 24 (16+8 x RAIDEE 備援磁碟) x NL-SAS 硬碟
- 硬碟儲存卷：1TB 在儲存池中

y I/O 模式

- 工具：IOmeter V1.1.0
- 工作：1
- 突出 (列隊深度)：128
- 存取規則：
  備份模式 (循序讀取 / 寫入，256KB (MB/s))
  資料庫存取模式 (由 Intel/StorageReview.com 定義，8KB，67%讀取，100%隨機)
  檔案伺服器存取模式 (由 Intel 定義，參見圖表 16，
  http://www.storagereview.com/articles/200003/20000313OSandBM_5.html)
  閒置

y

- 立 RAID60 儲存池。初始化後，拔出然後插入一顆磁碟。使用不同的 I/O 存取模式
- 按順序建立帶有 1/2/4/8 x RAID EE 備 磁碟的 RAID 60EE。初始化後，拔出一顆磁碟，RAID EE開始重建。使用不同的 I/O 存取模式計算重建時間。然後插入一顆磁碟並將其設定為專用備援磁碟，它開始回拷。最後，計算回拷時間。

 © 版權所有 2019 QSAN 廣盛科技保留所有權利。

表格 6　　RAID 60和 RAID 60EE的測試結果

	RAID 60 (x16)	RAID 60EE (x16+1)	RAID 60EE (x16+2)	RAID 60EE (x16+4)	RAID 60EE (x16+8)
循序讀取，256KB	24'58"	13'04"	11'37"	9'23"	4'02"
改善百分比		48%	53%	62%	84%
回拷時間		46'41"	41'38"	34'01"	35'43"
循序寫入，256KB	20'16"	13'59"	12'28"	9'40"	4'26"
改善百分比		31%	38%	52%	78%
回拷時間		33'32"	44'33"	31'54"	35'04"
資料庫存取模式	623'22"	93'24"	58'17"	35'34"	16'52"
改善百分比		85%	91%	94%	97%
回拷時間		843'11"	876'59"	492'25"	383'16"
檔案伺服器存取模式	458'03"	63'36"	55'23"	34'24"	15'51"
改善百分比		86%	88%	92%	97%
回拷時間		1215'34"	1087'47"	673'21"	478'57"
閒置	13'47"	10'49"	10'50"	7'24"	3'30"
改善百分比		22%	21%	46%	75%
回拷時間		26'46"	27'32"	29'30"	33'22"

舉個例子，用 256KB 循序讀取 RAID60 的重建時間為 24 分 58 秒。與具有 1 顆 RAID EE備 ᆟ磁碟的 RAID60EE相比，重建時間為 13 分 4 秒。它改善 ((24 x 60 +58) - (13 x 60 + 4)) / (24 x 60 +58) = (1498 ᄀ 784) / 1498 = 0.4766 == 48%。

圖表 18　　The Chart of RAID 60 and RAID 6EE

摘要

- RAID EE可將重建時間提高多達 97%。
- 使用 RAID EE的備援磁碟越多，重建時間就越少。
- 當存取為讀取時，重建時間更短。
- 存取模式　　回拷　　。
- 當　　回拷　　I/O。

結論

隨著　　容量的增長，RAID重建時間呈線性增長。磁碟組中包含的　　越多，累積故障增加的可能性就越大，　　容量對重建速度的影響也越大。使用 RAID EE技術將大大降低這些風險。

適用於

- XCubeSAN XS5200 / XS3200 / XS1200 韌體版本 1.4.1

XCubeSAN SANOS 4.0

 © 版權所有 2019 QSAN廣盛科技保留所有權利。