



XCubeSAN 系列技术白皮书

RAID EE 技术



QSAN 广盛科技
www.QSAN.com



版权

©版权所有 2019 QSAN 广盛科技保留所有权利。未经 QSAN 广盛科技书面许可，不得复制或传播本文档的任何部分。

2019 年 9 月

此版本适用于 QSAN XCubeSAN 系列。QSAN 认为本出版物在发布之日内容准确无误。信息如有更改，恕不另行通知。

商标

QSAN、QSAN 标志、XCubeSAN 和 QSAN.com 是 QSAN 广盛科技的商标或注册商标。

Microsoft、Windows、Windows Server 和 Hyper-V 是 Microsoft Corporation 在美国和/或其他国家/地区的商标或注册商标。

Linux 是 Linus Torvalds 在美国和/或其他国家/地区的商标。

UNIX 是 The Open Group 在美国和其他国家/地区的注册商标。

Mac 和 OS X 是 Apple Inc.在美国和其他国家/地区的注册商标。

Java 和所有基于 Java 的商标和标志是 Oracle 和/或其附属公司的商标或注册商标。

VMware、ESXi 和 vSphere 是 VMware, Inc. 在美国和/或其他国家/地区的注册商标或商标。

Citrix 和 Xen 是 Citrix Systems, Inc. 在美国和/或其他国家/地区的注册商标或商标。

本文件中用于宣称拥有商标和名称的实体或其产品的其他商标和商品名称均为其各自所有者的财产。

注意

此 XCubeSAN 系列技术白皮书适用于以下 XCubeSAN 型号：

XCubeSAN 存储系统 4U 19" 机架式机箱型号

型号名称	控制器类型	外形、盘位计数和机架单元
XS5224D	双控制器	LFF 24 磁盘 4U 机箱
XS3224D	双控制器	LFF 24 磁盘 4U 机箱
XS3224S	单控制器	LFF 24 磁盘 4U 机箱
XS1224D	双控制器	LFF 24 磁盘 4U 机箱
XS1224S	单控制器	LFF 24 磁盘 4U 机箱

XCubeSAN 存储系统 3U 19" 机架式机箱型号

型号名称	控制器类型	外形、盘位计数和机架单元
XS5216D	双控制器	LFF 16 磁盘 3U 机箱
XS3216D	双控制器	LFF 16 磁盘 3U 机箱
XS3216S	单控制器	LFF 16 磁盘 3U 机箱
XS1216D	双控制器	LFF 16 磁盘 3U 机箱
XS1216S	单控制器	LFF 16 磁盘 3U 机箱

XCubeSAN 存储系统 2U 19" 机架式机箱型号

型号名称	控制器类型	外形、盘位计数和机架单元
XS5212D	双控制器	LFF 12 磁盘 2U 机箱
XS5212S	单控制器	LFF 12 磁盘 2U 机箱
XS3212D	双控制器	LFF 12 磁盘 2U 机箱

XS3212S	单控制器	LFF 12 磁盘 2U 机箱
XS1212D	双控制器	LFF 12 磁盘 2U 机箱
XS1212S	单控制器	LFF 12 磁盘 2U 机箱
XS5226D	双控制器	SFF 26 磁盘 2U 机箱
XS5226S	单控制器	SFF 26 磁盘 2U 机箱
XS3226D	双控制器	SFF 26 磁盘 2U 机箱
XS3226S	单控制器	SFF 26 磁盘 2U 机箱
XS1226D	双控制器	SFF 26 磁盘 2U 机箱
XS1226S	单控制器	SFF 26 磁盘 2U 机箱

文件中所包含信息的准确性已被审查。但它可能包括印刷错误或技术不准确，这将定期对文档进行更改，而这些更改将纳入该出版物的新版本。QSAN 可能会对产品进行改进或更改，所有功能和产品规格如有更改，恕不另行通知或承担义务。本文档中的所有陈述、信息和建议均不构成任何明示或暗示的担保。

此处包含的任何性能数据都是在受控环境中确定的。因此，在其他操作环境中获得的结果可能会有很大差异。在开发级系统上进行的一些测试，并无法保证这些测试在一般的系统上是相同的。此外，可以通过外推估计一些测量值，实际结果可能有所不同，本文档的用户应验证其特定环境的适用数据。

此信息包含日常商业操作中使用的数据和报告的示例。为了尽可能完整地说明它们，这些例子包括个人、公司、品牌和产品的名称。所有这些名称都是虚构的，与实际商业企业使用的名称和地址如有任何相似之处完全是巧合。

目录

注意	i
RAID EE 技术.....	1
执行摘要.....	1
读者	1
概观	1
运作原理.....	3
配置 RAID EE 存储池	8
创建 RAID EE 存储池.....	8
列出 RAID EE 存储池.....	14
RAID EE 存储池上的操作	17
测试结果.....	20
测试案例 1：RAID 5 与 RAID 5EE	20
测试案例 2：RAID 60 与 RAID 60EE	23
结论	26
适用于	26
参考	26
附录	28
相关文件.....	28
技术支持.....	28

RAID EE 技术

执行摘要

已经存在超过 30 年的 RAID 架构正在经历一波转型。对于 TB 级大容量硬盘，原始 RAID 技术无法解决重建时间过长的问题。基于传统块技术的新一代 RAID 技术，我们称之为 RAID EE，这是被视为解决传统 RAID 缺陷的途径。



信息：

RAID EE 技术在 SANOS 固件版本 1.3.0 中可用，并且在 SANOS 固件版本 1.4.1 中性能有大幅度的提升。

读者

本文档适用于有兴趣了解 RAID EE 以解决重建时间过长问题的 QSAN 客户和合作伙伴。我们假定读者熟悉 QSAN 产品并具有一般 IT 经验，包括作为系统或网络管理员的知识。如有任何疑问，请参阅产品的使用手册，或联系 QSAN 技术支持以获得进一步的帮助。

概观

RAID（独立磁盘冗余阵列）将基于特定算法组合多个独立的物理磁盘，以形成虚拟逻辑磁盘，从而提供更大容量、更高性能或更好的数据容错能力。RAID 作为一种成熟可靠的数据保护标准，已成为存储系统的基础技术。然而，随着近年来磁盘对数据存储的需求快速增长以及高性能应用的出现，传统 RAID 逐渐暴露出其缺陷。

随着硬盘容量的增加，重建 RAID 数据所需的时间也大大增加。这是当今企业存储管理中最麻烦的问题之一。在过去几天硬盘容量只有 10GB 到 100GB 的情况下，RAID 重建的工作可以在 10 分钟甚至 10 分钟以上完成，这在没有特别关注的情况下还不是问题。但是，随着磁盘容量增长到数百 GB 甚至 TB，RAID 重建时间增加到数小时甚至数天，这成为存储管理中的主要问题。

例如，一个传统 RAID 5 内含有 8 加 1 颗奇偶校验的 6TB NL-SAS 磁盘上需要 2.5 天才能重建数据。重建过程会消耗系统资源，从而降低应用程序系统的整体性能。如果使用者限制重建优先权，则重建时间将更长。重要的是，在耗时的重建过程中，大量的访问操作可能导致存储池中其他磁盘的故障，从而大大增加了磁盘故障的可能性和数据丢失的风险。

传统 RAID 架构的局限性

传统的 RAID 架构由一定数量的磁盘组成磁盘组 (也称为 RAID 组)。您还可以将某些磁盘指定为空闲的热备用磁盘。存储池被分组以提供存储卷的容量，然后最终将 LUN 映像到主机以成为主机上的存储空间。

这种 RAID 架构有几个局限性：

- 首先，当磁盘组的磁盘损坏并且需要重建时，只有磁盘组的成员磁盘参与重建作业，此时加载的数据集中在备用磁盘上形成瓶颈。
- 其次，存储卷数据访问仅限于属于磁盘组的成员磁盘；这限制了主机的性能，因为存储装置正在执行访问和重建 I/O。

为什么 RAID 重建耗费时间

随着磁盘容量的增长，RAID 重建时间呈线性增长，当使用具有 4TB 以上硬盘容量的 RAID 磁盘时，将传统 RAID 架构所需的重建时间提高到数十小时。

有几个因素会影响 RAID 重建时间：

- **硬盘容量：**硬盘容量构成磁盘组，硬盘容量越大，重建时间越长。
- **磁盘数量：**磁盘组中包含的磁盘数量会影响系统从其余健康磁盘读取数据并将其写入热备用磁盘所需的时间。磁盘越多，重建时间越长。
- **重建作业优先权：**在 RAID 重建期间，系统仍需要假设对前端主机的 I/O 访问。分配给 RAID 重建作业的优先权越高，重建速度越快，但前端主机获得 I/O 性能的能力越低。
- **快速重建：**启用快速重建功能只需要重建存储卷的实际容量，未使用的磁盘组空间不用重建。如果存储卷仅使用磁盘组中的部分空间，则将缩短重建时间。
- **RAID 级别：**具有直接块到块复制的 RAID 1 和 RAID 10 将比具有奇偶校验计算的 RAID 5 和 RAID 6 重建地更快。

鉴于每个磁盘可能出现故障，磁盘组中包含的磁盘越多，累积故障的可能性就越大，因此磁盘组中的磁盘数量有上限的限制。与以前的因素相比，磁盘容量对重建速度的影响越来越大，这已成为主要因素。如此长的重建时间显然是任何用户都不能接受的。为了解决传统 RAID 的问题，我们实现了 RAID EE 技术。

运作原理

RAID EE 在磁盘组中新增更多备用磁盘，我们将其称为 **RAID EE 备用磁盘**，以分离原始的全局、本地和专用备用磁盘。备用区域保留在磁盘组的每个条带中，并透过磁盘旋转分布在磁盘组中。磁盘组中的磁盘发生故障时，缺少的数据将重建到保留的备用区中。由于集合中的所有磁盘都是重建数据的目标，因此传统 RAID 重建的瓶颈消失了，重建性能得到了显著提升。如果新增了新磁盘，则备用区中的数据将复制回新加入的磁盘。

为 RAID EE 提供了四个新的 RAID 级别，有：

- **RAID 5EE** (E 代表增强型)，至少需要 4 颗磁盘和一颗 RAID EE 备用磁盘，可以容忍 2 颗磁盘故障。新增更多 RAID EE 备用磁盘将容忍更多磁盘故障。
- **RAID 6EE** 至少需要 5 颗磁盘。
- **RAID 50EE** 至少需要 7 颗磁盘。
- **RAID 60EE** 至少需要 9 颗磁盘。



信息：

磁盘组中的 RAID EE 备用磁盘数量为 1 到 8 颗磁盘。

带有 1 颗 RAID EE 备用磁盘的 RAID 5EE 示例

举一个例子来描述它是如何工作的。以下示例是具有 5 颗磁盘的 RAID 5EE，4 颗磁盘用于 RAID 磁盘，另外一颗磁盘用于 RAID EE 备用磁盘。初始化后，数据块分配如下。P 代表奇偶校验，S 代表 RAID EE 备用磁盘，现在它的内容是空的。

D1	D2	D3	D4	D5
1	2	3	P	S
S	4	5	6	P
P	S	7	8	9
10	P	S	11	12
13	14	P	S	15

图表 1 RAID 5EE 的数据块分布

假设磁盘 2 发生故障。RAID 5EE 处于降级模式。

D1	D2	D3	D4	D5
1	2	3	P	S
S	4	5	6	P
P	S	7	8	9
10	P	S	11	12
13	14	P	S	15

图表 2 磁盘 2 已故障

故障磁盘中的数据将被重建回备用区域。此操作称为 **EE 重建**。重建后，分布的数据就像 RAID 5，它可以容忍另一个故障的磁盘。我们可以想象，RAID EE 备用磁盘越多，重建速度就越快。

D1		D3	D4	D5
1		3	P	2
4		5	6	P
P		7	8	9
10		P	11	12
13		P	14	15

图表 3 故障磁盘中的数据重建回空的块

当新磁盘加入 RAID EE 磁盘组时，备用区中重建的数据将被复制回新的磁盘。此操作称为回拷。

D1	D2	D3	D4	D5
1	2	3	P	S
S	4	5	6	P
P	S	7	8	9
10	P	S	11	12
13	14	P	S	15

图表 4 数据被回拷

回拷后，它回到 RAID 5EE 的正常状态。

带有 2 颗 RAID EE 备用磁盘的 RAID 60EE 示例

再举一个带有 10 颗磁盘的 RAID 60EE 示例。8 颗磁盘用于 RAID 磁盘，2 颗磁盘用于 RAID EE 备用磁盘。初始化后，数据块分配如下。

Quantity of Disk Group = 10									
Subgroup 1				Subgroup 2				Spare = 2	
D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
1	2	P	Q	3	4	P	Q	S	S
S	5	6	P	Q	7	8	P	Q	S
S	S	9	10	P	Q	11	12	P	Q
Q	S	S	13	14	P	Q	15	16	P
P	Q	S	S	17	18	P	Q	19	20
24	P	Q	S	S	21	22	P	Q	23
27	28	P	Q	S	S	25	26	P	Q
Q	31	32	P	Q	S	S	29	30	P
P	Q	35	36	P	Q	S	S	33	34
38	P	Q	39	40	P	Q	S	S	37

图表 5 RAID 60EE 的数据块分布

RAID 60EE 的重建和回拷与上述类似；这里不再重复。

RAID EE 级别摘要

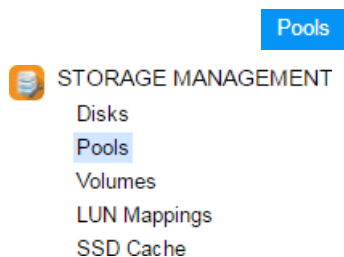
以下是 RAID EE 级别摘要。

表格 1 RAID EE 级别摘要

	RAID 5EE	RAID 6EE	RAID 50EE	RAID 60EE
最少磁盘数量	4	5	7	9
容错 (G = 子组, S = RAID EE 备用磁盘 = 1 ~ 8)	2 ~ 9 颗磁盘故障 (例如, 1 (RAID 5) + S 颗备用磁盘)	3 ~ 10 颗磁盘故障 (例如, 2 (RAID 6) + S 颗备用磁盘)	G+1 ~ G+8 颗 磁盘故障 (例如, 2 个子组 (RAID 50) + S 颗备 用磁盘)	2xG+1 ~ 2xG+8 颗磁盘 故障 (例如, 2x2 个子 组 (RAID 60) + S 颗备用磁盘)
读取性能	非常好	非常好	非常好	非常好
写入性能	好	持平到好	好	持平到好
容量 (N = 磁盘数量, M = 磁盘容量, G = 子组, S = RAID EE 颗备用磁 盘)	$(N-1-S) \times M$ (例如, (10 颗磁盘 - 1 - 2 颗备用磁 盘) $\times$ 1TB = 7TB)	$(N-2-S) \times M$ (例如, (10 颗磁盘 - 2 - 2 颗备用磁盘) $\times$ 1TB = 6TB)	$(N-G-S) \times M$ (例如, (10 颗磁盘 - 2 个子组 - 2 颗 备用磁盘) $\times$ 1TB = 6TB)	$(N-2 \times G-S) \times M$ (例如, (10 颗磁盘 - 2x2 个子组 - 2 颗 备用磁盘) $\times$ 1TB = 4TB)
容量利用率 (最少 ~ 26 颗磁盘)	18% ~ 92% (例如, 7/10 = 70%)	17% ~ 88% (例如, 6/10 = 60%)	29% ~ 88% (例如, 6/10 = 60%)	25% ~ 80% (例如, 4/10 = 40%)
典型应用	数据仓储、网 站服务、存盘	数据存盘、高 可用性解决方 案、具有大容 量要求的服务 器	大型数据库、 文件服务器、 应用服务器	数据存盘、高 可用性解决方 案、具有大容 量要求的服务 器

配置 RAID EE 存储池

本节将介绍配置 RAID EE 存储池的操作。



图表 6 存储池功能子菜单

创建 RAID EE 存储池

以下是创建配置 4 颗磁盘的 RAID 5EE 存储池示例。在首次创建存储池时，包含一个磁盘组，且磁盘组中的最大磁盘数为 64。

1. 选择存储池功能子菜单，单击创建存储池按钮。首先扫描可用磁盘。



提示：

如果您的系统有超过 200 颗磁盘，扫描磁盘可能需要 20 ~ 30 秒。请耐心等待。

Create Pool

General

Disk Selection

RAID Configuration

Disk Properties

Summary

Pool Type

Please select a pool type.

☒ Thick Provisioning

☐ Thin Provisioning

☐ Auto Tiering (Thin Provisioning Enabled)

Pool Properties

Please enter a pool name and select preferred controller setting.

Pool Name : ⓘ

Preferred Controller : ▼

The I/O resources will be managed by the preferred controller which you specified.

SED Pool

☐ Enable SED Pool

Enabling SED pool will use the secure SEDs to create a pool. Intermixing SEDs and non-SEDs are not supported in a pool.

Next Cancel

图表 7 创建 RAID EE 存储池步骤 1

2. 选择存储池类型。
3. 输入存储池名称，存储池名称的最大长度为 15 个字符，有效字符为 [A~Z | a~z | 0~9 | - _ < >]。
4. 从下拉列表中选择首选控制器，此存储池中的后端 I/O 资源将由您指定的首选控制器处理，此选项在安装双控制器时可以使用。
5. 点选启用 **SED 存储池** 复选框，启用 SED 存储池将使用安全的 SED 创建存储池，不支持混合 SED 和非 SED 磁盘在同一个存储池中。
6. 单击下一步按钮继续。

General

Disk Selection

RAID Configuration

Disk Properties

Summary

Select Disks

Please select disks to add a disk group in a thick provisioning pool. The maximum quantity of disk in a disk group is 64.

Enclosure ID: 0 (Head Unit: XS5216)

	Enclosure ID	Slot	Health	Capacity	Disk Type	Manufacturer	Model
☐	0	1	Good	372.36 GB	SAS SSD 12.0Gb/s	SEAGATE	ST400FM0053
☐	0	2	Good	372.36 GB	SAS SSD 12.0Gb/s	SEAGATE	ST400FM0053
☐	0	3	Good	372.36 GB	SAS SSD 12.0Gb/s	SEAGATE	ST400FM0053
☐	0	4	Good	372.36 GB	SAS SSD 12.0Gb/s	SEAGATE	ST400FM0053
☐	0	5	Good	744.96 GB	SAS SSD 12.0Gb/s	MICRON	S630DC-800
☐	0	6	Good	744.96 GB	SAS SSD 12.0Gb/s	MICRON	S630DC-800
☐	0	7	Good	744.96 GB	SAS SSD 12.0Gb/s	MICRON	S630DC-800
☐	0	8	Good	744.96 GB	SAS SSD 12.0Gb/s	MICRON	S630DC-800
☒	0	9	Good	1.09 TB	SAS HDD 12.0Gb/s	SEAGATE	ST1200MM0088
☒	0	10	Good	1.09 TB	SAS HDD 12.0Gb/s	SEAGATE	ST1200MM0088
☒	0	11	Good	1.09 TB	SAS HDD 12.0Gb/s	SEAGATE	ST1200MM0088
☒	0	12	Good	1.09 TB	SAS HDD 12.0Gb/s	SEAGATE	ST1200MM0088
☐	0	13	Good	5.46 TB	NL-SAS HDD 12.0Gb/s	SEAGATE	ST6000NM0014
☐	0	14	Good	5.46 TB	NL-SAS HDD 12.0Gb/s	SEAGATE	ST6000NM0014
☐	0	15	Good	5.46 TB	NL-SAS HDD 12.0Gb/s	SEAGATE	ST6000NM0014

Back

Next

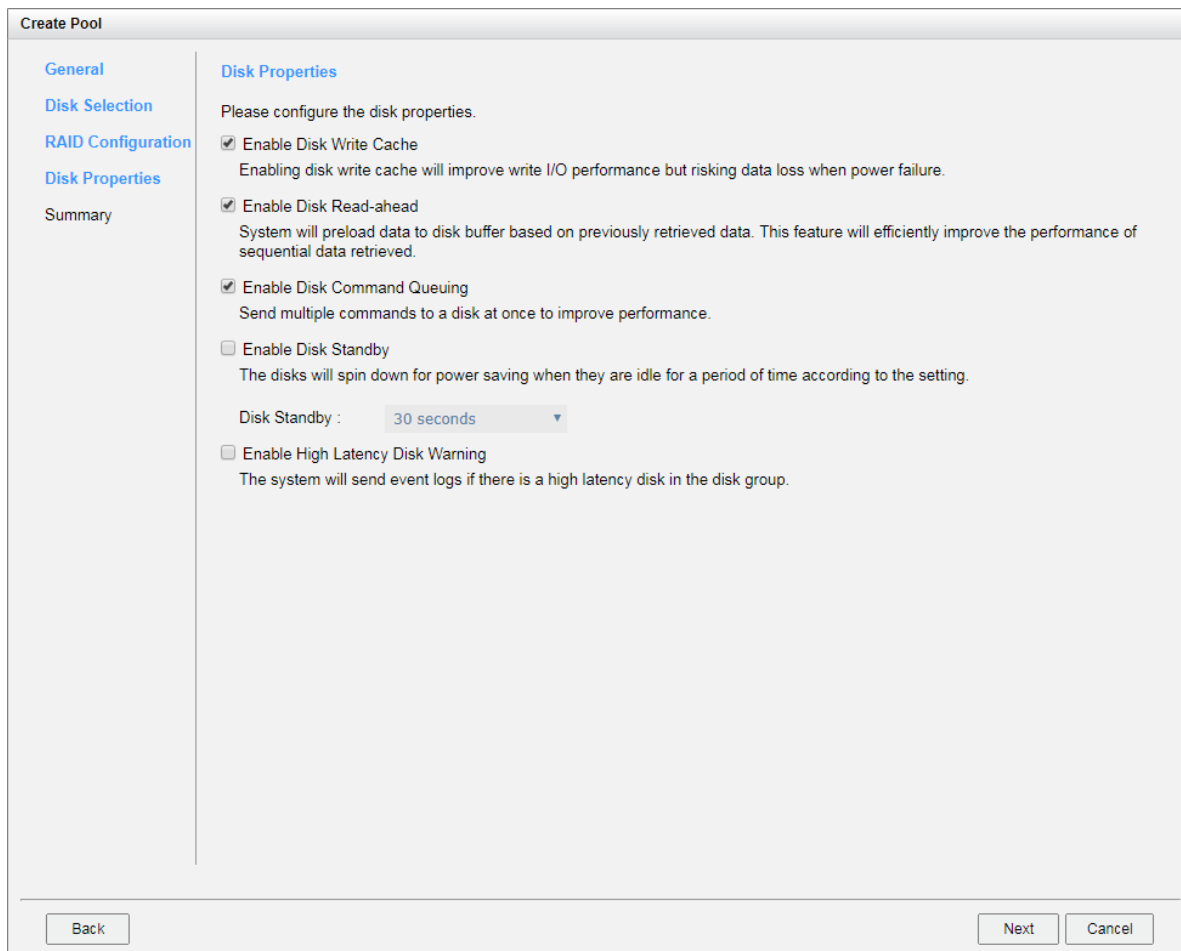
Cancel

图表 8 创建 RAID EE 存储池步骤 2

7. 请为存储池选择磁盘。磁盘组中的最大磁盘数为 64，从下拉列表中选择**机箱 ID** 以从扩展机柜中选择磁盘。
8. 单击下一步按钮继续。

图表 9 创建 RAID EE 存储池步骤 3

9. 从下拉列表中选择 **RAID 级别**，该列表仅根据选择的磁盘列出可用的 RAID 级别。如果选择了 RAID EE 级别，则需选择 **RAID EE 备用磁盘数量**。如果选择了组合 RAID 级别，还可以选择子组数量。
10. 单击下一步按钮继续。



图表 10 创建 RAID EE 存储池步骤 4

11. 磁盘属性也可以在此步骤中选择配置：

- **启用磁盘写入缓存：**点选以启用磁盘的写缓存选项。启用磁盘写入缓存可提高写入 I/O 性能，但在断电时可能会丢失数据。
- **启用磁盘预读：**点选以启用磁盘的预读功能。系统将根据先前检索的数据预先加载下一笔数据到磁盘缓冲区，此功能将有效地提高被检索的顺序数据性能。
- **启用磁盘命令队列：**点选以启用磁盘的命令队列功能。一次将多个命令发送到磁盘以提高性能。
- **启用磁盘待机：**点选以启用磁盘自动节能功能。当磁盘空闲一段时间后，磁盘将自动关闭以节省电量。
- **启用高延迟磁盘警告：**点选以在磁盘组中启用高延迟磁盘警告功能。当磁盘出现异常且延迟较高时，系统将发送事件日志以提前通知 IT 管理员，以避免影响整体性能。

12. 单击下一步按钮继续。

Create Pool

General

Pool Properties

Pool Type : Thick Provisioning

Pool Name : Pool-5

Preferred Controller : Controller 1

RAID Configuration

RAID Level : RAID 5EE

Quantity of RAID EE Spares : 1

Quantity of SAS Disks : 4

Disk Properties

Disk Write Cache : Enabled

Disk Read-ahead : Enabled

Disk Command Queuing : Enabled

Disk Standby : Disabled

Back Finish Cancel

图表 11 创建 RAID EE 存储池步骤 5

13. 在摘要页面确认后，单击完成按钮以创建存储池。

	Pool Name	Status	Health	Total	Free	Available	Thin Provisioning	Auto Tiering	Encryption	Volumes	Current Controller
▼	Pool-5	Online	Good	2.18 TB	2.18 TB	2.18 TB	Disabled	Disabled	Disabled	0	Controller 1

Create Pool

图表 12 已创建一个 RAID EE 存储池

14. 已创建存储池。如有必要，再次单击创建存储池按钮以创建其他存储池。

列出 RAID EE 存储池

单击一个存储池；它将显示相关的磁盘组。同样，单击一个磁盘组；它将显示相关的磁盘。存储池属性可以透过单击存储池左侧的功能按钮 ▼ 来配置。

	Pool Name	Status	Health	Total	Free	Available	Thin Provisioning	Auto Tiering	Encryption	Volumes	Current Controller
▼	Pool-5	Online	Good	2.18 TB	2.18 TB	2.18 TB	Disabled	Disabled	Disabled	0	Controller 1

Disk Groups

	No.	Status	Health	Total	Free	RAID	Disks Used	RAID EE Spares
▼	1	Online	Good	2.18 TB	2.18 TB	RAID 5EE	4	1

Disks

Enclosure ID	Slot	Status	Health	Capacity	Disk Type	Manufacturer	Model
0	9	Online	Good	1.09 TB	SAS HDD 12.0Gb/s	SEAGATE	ST1200MM0088
0	10	Online	Good	1.09 TB	SAS HDD 12.0Gb/s	SEAGATE	ST1200MM0088
0	11	Online	Good	1.09 TB	SAS HDD 12.0Gb/s	SEAGATE	ST1200MM0088
0	12	Online	Good	1.09 TB	SAS HDD 12.0Gb/s	SEAGATE	ST1200MM0088

Create Pool

图表 13 列出 RAID EE 存储池

此表显示列字段说明。

表格 2 存储池列字段说明

列名称	说明
存储池名称	存储池名称。
状态	存储池的状态： - 在线：存储池在线。 - 离线：存储池离线。 - 正在重建中：正在重建存储池。 - 正在迁移中：正在迁移存储池。 - 正在搬迁中：正在搬迁存储池。 - 正在 EE 重建中：正在重建 RAID EE 存储池。
健康状况	存储池的健康状况：

	- 良好：存储池良好。 - 失效：存储池失效。 - 降级：存储池不健康且不完整，原因可能是磁盘丢失或失效。
总容量	存储池的总容量。
剩余容量	存储池的剩余容量。
可用	存储池的可用容量。
自动精简配置	自动精简配置的状态： - 停用：存储池为厚配置。 - 启用：存储池为自动精简配置。
自动分层存储	自动分层的状态： - 停用：存储池已禁用自动分层。 - 启用：存储池已启用自动分层。 - 不支援：存储池含有混合磁盘类型的磁盘组。
磁盘加密	数据安全模式： - 停用：存储池未加密。 - 启用：存储池已加密。
存储卷	存储池中的存储卷数量。
当前控制器 (此选项仅在安装 双控制器时可 见。)	存储池的当前运行控制器。

表格 3 磁盘组列字段说明

列名称	说明
序号	磁盘组的编号。

状态	磁盘组的状态： • 在线：磁盘组在线。 • 离线：磁盘组离线。 • 正在重建中：正在重建磁盘组。 • 正在迁移中：正在迁移磁盘组。 • 正在搬迁中：正在搬迁磁盘组。 • 正在 EE 重建中：正在重建 RAID EE 磁盘组。
健康状况	磁盘组的健康状况： • 良好：磁盘组良好。 • 失效：磁盘组失效。 • 降级：磁盘组不健康且不完整，原因可能是磁盘丢失或失效。
总容量	磁盘组的总容量。
剩余容量	磁盘组的剩余容量。
RAID	磁盘组的 RAID 级别。
磁盘使用数量	磁盘组中的磁盘数量。
RAID EE 备用磁盘	磁盘组中 RAID EE 备用磁盘的数量，传统 RAID 级别显示为不适用。

表格 4 磁盘列字段说明

列名称	说明
机箱 ID	机箱 ID。
插槽	磁盘的位置。
状态	磁盘的状态： • 在线：磁盘在线。 • 丢失：存储池中缺少磁盘。

	- 正在重建中：正在重建磁盘。 - 正在转换中：当发生重建时，磁盘正在迁移或正在被另一个磁盘替换。 - 正在擦除中：正在擦除磁盘。 - 检查完成：已检查磁盘健康状况。 - 正在回拷中：正在回拷磁盘。
健康状况	磁盘的运行状况： - 良好：磁盘良好。 - 失效：磁盘发生故障。 - 错误提示：S.M.A.R.T.错误提醒。 - 读取错误：磁盘具有不可恢复的读取错误。
容量	磁盘的容量。
磁盘类型	磁盘的类型： [SAS 硬盘 NL-SAS 硬盘 SAS 固态硬盘 SATA 固态硬盘] [12.0Gb/s 6.0Gb/s 3.0Gb/s 1.5Gb/s]
制造商	磁盘的制造商。
型号	磁盘的型号名称。

RAID EE 存储池上的操作

[XCubeSAN SANOS 4.0 软件手册](#)中的“配置存储池”章节介绍了大多数操作。以下我们描述有关 RAID EE 存储池的限制。

存储池验证奇偶校验

单击任一个存储池中的 ▼ -> 验证奇偶校验以重新生成存储池的奇偶校验。支持 RAID 级别 3、5、6、30、50、60 和 RAID EE 级别 5EE、6EE、50EE、60EE。

新增磁盘组到存储池中

单击任一个存储池中的 ▼ -> 新增磁盘组将另一个磁盘组新增到存储池中。新增加的磁盘组可以是 RAID EE 级别或传统 RAID 级别。有关详细讯息，请参阅 [XCubeSAN SANOS 4.0 软件手册](#) 中的第 8.4.5 节 “在厚配置存储池中新增磁盘组” 一节。

Add Disk Group

Pool Type

Thin Provisioning : Disabled

Auto Tiering : Disabled ⓘ

RAID Level

Please select a RAID level.

RAID Level :

RAID 0
RAID 0
RAID 1
RAID 3
RAID 5
RAID 5EE
RAID 6
RAID 6EE
RAID 0+1
RAID 10
RAID 30
RAID 50
RAID 50EE
RAID 60
RAID 60EE

Select Disks

Please select disks to add a disk group. disk in a disk group is 64.

Enclosure ID :

☐	Enclosure ID	Slot	Health	Capacity	Type	Manufacturer	Model
☐	0	1	Good	744.96 GB	SAS SSD 12.0Gb/s	MICRON	S630DC-800
☐	0	2	Good	744.96 GB	SAS SSD 12.0Gb/s	MICRON	S630DC-800
☐	0	3	Good	744.96 GB	SAS SSD 12.0Gb/s	MICRON	S630DC-800
☐	0	4	Good	744.96 GB	SAS SSD 12.0Gb/s	MICRON	S630DC-800
☐	0	5	Good	744.96 GB	SAS SSD 12.0Gb/s	MICRON	S630DC-800
☐	0	6	Good	744.96 GB	SAS SSD 12.0Gb/s	MICRON	S630DC-800
☐	0	7	Good	744.96 GB	SAS SSD 12.0Gb/s	MICRON	S630DC-800
☐	0	8	Good	744.96 GB	SAS SSD 12.0Gb/s	MICRON	S630DC-800
☐	0	13	Good	5.46 TB	NL-SAS HDD 12.0Gb/s	SEAGATE	ST6000NM0014
☐	0	14	Good	5.46 TB	NL-SAS HDD 12.0Gb/s	SEAGATE	ST6000NM0014
☐	0	15	Good	5.46 TB	NL-SAS HDD 12.0Gb/s	SEAGATE	ST6000NM0034

OK

Cancel

图表 14 将磁盘组新增到存储池中

迁移 RAID EE 存储池中的磁盘组

迁移磁盘组功能将磁盘组更改为其他 RAID 级别，或新增磁盘组的成员磁盘以增加容量。传统的 RAID 级别可以迁移到 RAID EE 级别，但 RAID EE 级别只能迁移到 RAID EE 级别。

Migrate Disk Group

RAID Level

Please select a RAID level.

RAID Level : RAID 5EE

Quantity of RAID 2.0 Spares : RAID 5EE

Select Disks

Please select disks to migrate the disk group. The maximum quantity of disk in a disk group is 64.

Enclosure ID : 0 (Head Unit: XS5216)

☐	Enclosure ID	Slot	Health	Capacity	Disk Type	Manufacturer	Model
☐	0	5	Good	744.96 GB	SAS SSD 12.0Gb/s	MICRON	S630DC-800
☐	0	6	Good	744.96 GB	SAS SSD 12.0Gb/s	MICRON	S630DC-800
☐	0	7	Good	744.96 GB	SAS SSD 12.0Gb/s	MICRON	S630DC-800
☐	0	8	Good	744.96 GB	SAS SSD 12.0Gb/s	MICRON	S630DC-800
☒	0	9	Good	1.09 TB	SAS HDD 12.0Gb/s	SEAGATE	ST1200MM0088
☒	0	10	Good	1.09 TB	SAS HDD 12.0Gb/s	SEAGATE	ST1200MM0088
☒	0	11	Good	1.09 TB	SAS HDD 12.0Gb/s	SEAGATE	ST1200MM0088
☒	0	12	Good	1.09 TB	SAS HDD 12.0Gb/s	SEAGATE	ST1200MM0088
☐	0	13	Good	5.46 TB	NL-SAS HDD 12.0Gb/s	SEAGATE	ST6000NM0014
☐	0	14	Good	5.46 TB	NL-SAS HDD 12.0Gb/s	SEAGATE	ST6000NM0014
☐	0	15	Good	5.46 TB	NL-SAS HDD 12.0Gb/s	SEAGATE	ST6000NM0034

OK Cancel

图表 15 迁移厚配置存储池中的 RAID EE 磁盘组



提示：

传统的 RAID 级别可以迁移到 RAID EE 级别，但 RAID EE 级别只能迁移到 RAID EE 级别。

测试结果

测试案例 1: RAID 5 与 RAID 5EE

此测试提供 RAID 5 和 RAID 5EE 之间重建时间和回拷时间的比较。我们假设 RAID EE 备用磁盘越多，重建时间就越短。

测试设备和配置

- 服务器
 - 型号：华硕 RS700 X7/PS4 (中央处理器：Intel Xeon E5-2600 v2 / 内存：8GB)
 - iSCSI 主机总线适配器：Intel 82574L 千兆网络连接
 - 操作系统：Windows Server 2012 R2
- 存储设备
 - 型号：XCubeSAN XS5224D
 - 内存：16GB (2 x 8GB 在插槽 1 和 3) 每个控制器
 - 固件版本：1.4.1
 - 磁盘：24 x Seagate Constellation ES, ST500NM0001, 500GB, SAS 6Gb/s
 - 硬盘存储池：
 - RAID 5** 存储池，控制器 1 中具有 16 x NL-SAS 硬盘
 - RAID 5EE** 存储池，控制器 1 中具有 17 (16+1 x RAID EE 备用磁盘) x NL-SAS 硬盘
 - RAID 5EE** 存储池，控制器 1 中具有 18 (16+2 x RAID EE 备用磁盘) x NL-SAS 硬盘
 - RAID 5EE** 存储池，控制器 1 中具有 20 (16+4 x RAID EE 备用磁盘) x NL-SAS 硬盘
 - RAID 5EE** 存储池，控制器 1 中具有 24 (16+8 x RAID EE 备用磁盘) x NL-SAS 硬盘
 - 硬盘存储卷：1TB 在存储池中
- I/O 模式
 - 工具：IOmeter V1.1.0
 - 工作：1
 - 突出 (列队深度)：128
 - 访问规则：
 - 备份模式 (顺序读取 / 写入，256KB (MB/s))
 - 数据库访问模式 (由 Intel/StorageReview.com 定义，8KB，67% 读取，100% 随机)
 - 文件服务器访问模式 (由 Intel 定义，参见图表 16，
http://www.storagereview.com/articles/200003/20000313OSandBM_5.html)

闲置

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%

图表 16 由 StorageReview.com 提供的访问模式

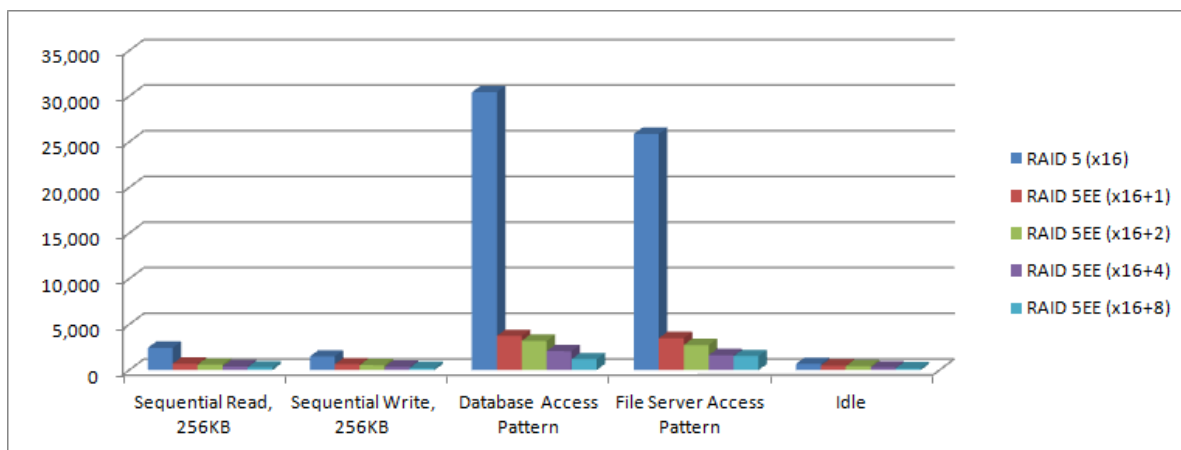
- 测试场景
 - 首先我们创建一个 RAID 5 存储池。初始化后，拔出然后插入一颗磁盘。使用不同的 I/O 访问模式计算重建时间。
 - 接下来按顺序创建带有 1/2/4/8 x RAID EE 备用磁盘的 RAID 5EE。初始化后，拔出一颗磁盘，RAID EE 开始重建。使用不同的 I/O 访问模式计算重建时间。然后插入一颗磁盘并将其设定为专用备用磁盘，它开始回拷。最后，计算回拷时间。

测试结果

表格 5 RAID 5 和 RAID 5EE 的测试结果

	RAID 5 (x16)	RAID 5EE (x16+1)	RAID 5EE (x16+2)	RAID 5EE (x16+4)	RAID 5EE (x16+8)
顺序读取， 256KB	40'43"	11'39"	10'09"	6'46"	4'58"
改善百分比		71%	75%	83%	88%
回拷时间		34'08"	30'57"	28'30"	30'56"
顺序写入， 256KB	24'54"	10'37"	9'08"	6'12"	4'02"
改善百分比		57%	63%	75%	84%
回拷时间		22'29"	30'23"	28'50"	31'54"
数据库访问模式	507'33"	62'23"	53'25"	34'37"	19'50"
改善百分比		88%	89%	93%	96%
回拷时间		1320'37"	1082'21"	829'00"	754'35"
文件服务器访问 模式	431'18"	58'24"	45'54"	27'19"	25'42"
改善百分比		86%	89%	94%	94%
回拷时间		1363'17"	1093'20"	736'87"	705'50"
闲置	11'20"	8'24"	7'20"	5'02"	3'22"
改善百分比		26%	35%	56%	70%
回拷时间		21'36"	22'26"	24'31"	26'05"

举个例子，用 256KB 顺序读取 RAID 5 的重建时间为 40 分 43 秒。与具有 1 颗 RAID EE 备用磁盘的 RAID 5EE 相比，重建时间为 11 分 39 秒。它改善 $((40 \times 60 + 43) - (11 \times 60 + 39)) / (40 \times 60 + 43) = (2443 - 699) / 2443 = 0.7138 = 71\%$ 。



图表 17 RAID 5 和 RAID 5EE 的图表

摘要

- RAID EE 可将重建时间提高多达 **96%**。
- 使用 RAID EE 的备用磁盘越多，重建时间就越少。
- 当访问为读取时，重建时间更有效。
- 如果访问模式是随机的，则回拷时间较长。
- 当数据进行回拷时，系统资源将优先保留给前端 I/O。

测试案例 2: RAID 60 与 RAID 60EE

此测试提供 RAID 60 和 RAID 60EE 之间重建时间和回拷时间的比较。同样，我们假设 RAID EE 备用磁盘越多，重建时间越短，RAID 60EE 的效率也越高。

测试设备和配置

- 服务器
 - 型号：华硕 RS700 X7/PS4 (中央处理器：Intel Xeon E5-2600 v2 / 内存：8GB)
 - iSCSI 主机总线适配器：Intel 82574L 千兆网络连接
 - 操作系统：Windows Server 2012 R2
- 存储设备
 - 型号：XCubeSAN XS5224D
 - 内存：16GB (2 x 8GB 在插槽 1 和 3) 每个控制器
 - 固件版本：1.4.1
 - 磁盘：24 x Seagate Constellation ES, ST500NM0001, 500GB, SAS 6Gb/s

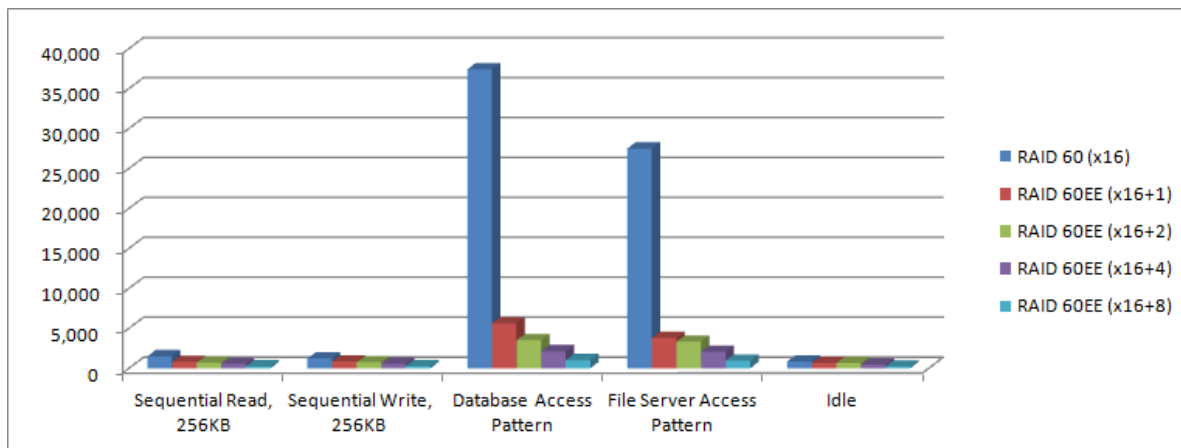
- 硬盘存储池：
 - RAID 60** 存储池，控制器 1 中具有 16 x NL-SAS 硬盘
 - RAID 60EE** 存储池，控制器 1 中具有 17 (16+1 x RAID EE 备用磁盘) x NL-SAS 硬盘
 - RAID 60EE** 存储池，控制器 1 中具有 18 (16+2 x RAID EE 备用磁盘) x NL-SAS 硬盘
 - RAID 60EE** 存储池，控制器 1 中具有 20 (16+4 x RAID EE 备用磁盘) x NL-SAS 硬盘
 - RAID 60EE** 存储池，控制器 1 中具有 24 (16+8 x RAID EE 备用磁盘) x NL-SAS 硬盘
- 硬盘存储卷：1TB 在存储池中
- I/O 模式
 - 工具：IOmeter V1.1.0
 - 工作：1
 - 突出 (列队深度)：128
 - 访问规则：
 - 备份模式 (顺序读取 / 写入，256KB (MB/s))
 - 数据库访问模式 (由 Intel/StorageReview.com 定义，8KB，67% 读取，100% 随机)
 - 文件服务器访问模式 (由 Intel 定义，参见图表 16，
http://www.storagereview.com/articles/200003/20000313OSandBM_5.html)
 - 闲置
- 测试场景
 - 首先我们创建一个 RAID 60 存储池。初始化后，拔出然后插入一颗磁盘。使用不同的 I/O 访问模式计算重建时间。
 - 接下来按顺序创建带有 1/2/4/8 x RAID EE 备用磁盘的 RAID 60EE。初始化后，拔出一颗磁盘，RAID EE 开始重建。使用不同的 I/O 访问模式计算重建时间。然后插入一颗磁盘并将其设定为专用备用磁盘，它开始回拷。最后，计算回拷时间。

测试结果

表格 6 RAID 60 和 RAID 60EE 的测试结果

	RAID 60 (x16)	RAID 60EE (x16+1)	RAID 60EE (x16+2)	RAID 60EE (x16+4)	RAID 60EE (x16+8)
顺序读取， 256KB	24'58"	13'04"	11'37"	9'23"	4'02"
改善百分比		48%	53%	62%	84%
回拷时间		46'41"	41'38"	34'01"	35'43"
顺序写入， 256KB	20'16"	13'59"	12'28"	9'40"	4'26"
改善百分比		31%	38%	52%	78%
回拷时间		33'32"	44'33"	31'54"	35'04"
数据库访问模式	623'22"	93'24"	58'17"	35'34"	16'52"
改善百分比		85%	91%	94%	97%
回拷时间		843'11"	876'59"	492'25"	383'16"
文件服务器访问 模式	458'03"	63'36"	55'23"	34'24"	15'51"
改善百分比		86%	88%	92%	97%
回拷时间		1215'34"	1087'47"	673'21"	478'57"
闲置	13'47"	10'49"	10'50"	7'24"	3'30"
改善百分比		22%	21%	46%	75%
回拷时间		26'46"	27'32"	29'30"	33'22"

举个例子，用 256KB 顺序读取 RAID 60 的重建时间为 24 分 58 秒。与具有 1 颗 RAID EE 备用磁盘的 RAID 60EE 相比，重建时间为 13 分 4 秒。它改善 $((24 \times 60 + 58) - (13 \times 60 + 4)) / (24 \times 60 + 58) = (1498 - 784) / 1498 = 0.4766 == 48\%$ 。



图表 18 The Chart of RAID 60 and RAID 60EE

摘要

- RAID EE 可将重建时间提高多达 97%。
- 使用 RAID EE 的备用磁盘越多，重建时间就越少。
- 当访问为读取时，重建时间更有效。
- 如果访问模式是随机的，则回拷时间较长。
- 当数据进行回拷时，系统资源将优先保留给前端 I/O。

结论

随着磁盘容量的增长，RAID 重建时间呈线性增长。磁盘组中包含的磁盘越多，累积故障增加的可能性就越大，磁盘容量对重建速度的影响也越大。使用 RAID EE 技术将大大降低这些风险。

适用于

- XCubeSAN XS5200 / XS3200 / XS1200 固件版本 1.4.1 及更高版本

参考

XCubeSAN SANOS 4.0 软件手册

- [XCubeSAN SANOS 4.0 软件手册](#)

附录

相关文件

有相关文件可以从网站上下载。

- [所有 XCubeSAN 文档](#)
- [XCubeSAN QIG \(快速安装指南\)](#)
- [XCubeSAN 硬件手册](#)
- [XCubeSAN 配置工作表](#)
- [XCubeSAN SANOS 4.0 软件手册](#)
- [兼容性矩阵](#)
- [技术白皮书](#)
- [应用说明](#)

技术支持

您是否有任何问题或需要帮助解决问题？请联系 QSAN 技术支持团队，我们会尽快回复给您。

- 通过网站：https://www.qsan.com/technical_support
- 通过电话：+886-2-77206355
(服务时间：09:30 - 18:00，周一至周五，UTC+8)
- 通过 Skype 通话，Skype ID：qsan.support
(服务时间：09:30 - 02:00，周一至周五，UTC+8，夏令时间：09:30 - 01:00)
- 通过电子邮件：support@qsan.com