

大型金融机构使其数据中心效率保持领先一代的水平



一家大型金融机构为了管理其存储基础架构的持续快速增长，决定探索当前基于光纤通道 (FC) 的 SAN 的替代方案。该公司向 Microsoft 和 Mellanox 寻求高级存储解决方案，希望使用更高效和可扩展的互连技术，可以更快速地处理大型数据集，并且成本更低。

执行摘要

呈指数级持续增长的数据迫使 IT 经理停止部署传统的存储区域网络 (SAN) 技术，并采用诸如横向扩展、软件定义的存储之类的现代体系架构，从而实现更程度的可扩展性，并且能够以更低的成本更快地处理更多的数据。

一家大型金融机构在管理其快速增长的存储基础架构方面遇到挑战，因此决定将扩展其当前光纤通道 (FC) SAN 解决方案的选项与融合横向扩展存储解决方案进行比较，后者基于 Microsoft 存储空间并使用 Mellanox 支持 RDMA (远程直接内存访问) 的架构。

基准测试结果显示，在使用 RoCE 端到端解决方案 (基于融合以太网的 RDMA) 的 Mellanox 40GbE 上运行基于 SMB Direct 的存储空间时，该公司能够：

- 使存储访问加速 2.5 至 9 倍
- 使虚拟机实时迁移速度加快 8.8 倍，存储迁移速度加快 12.5 倍

- 将 IO 端口数量减少到 1/4
- 使存储成本降低 50%

成本节省源于 Microsoft 的存储空间能够利用行业标准硬件，这消除了基于 FC 的传统 SAN 解决方案所需的专有硬件和软件的成本。

形势

十多年来，存储体系架构一直保持相对稳定——分别使用 SAN 和 NAS 作为块存储和文件存储的主要体系架构。然而，随着虚拟化和基于 SSD 存储等新技术的发展，以及 (近乎) 实时处理移动用户不同类型海量数据的需要，IT 经理不得不评估和采用新的存储系统。

“为降低存储成本，我们使用更多的网络！我们如何扩展 Azure 存储？RoCE (基于以太网的 RDMA) 支持 40GbE 的 Windows Azure 存储，可实现显著的 COGS 节省”
 Albert Greenberg 在 2014 年开放网络峰会上的主题演讲

例如，在“预虚拟化”时代，当每个服务器运行一个作业时，只通过处理一种数据类型的一个作业来处理存储访问。因此，很容易配置存储系统并使其高效地运行。但是，在虚拟化环境中，许多虚拟机并行运行，每个虚拟机可能需要来自磁盘上不同位置的数据，并且当数据具有不同类型（文本、视频和语音）和大小（从 1B 到 256KB）时，传统体系架构的效率要低得多，因为没有办法针对同时运行的所有工作负载优化存储系统。

最近，一家大型金融机构（银行）决定增加一套新的内部和外部服务，需要大规模拓展其存储容量，以及实时提供这些服务的能力，同时让总体拥有成本 (TCO) 保持可控。该公司认识到，使用当前的基于 FC 的 SAN 体系架构会造成管理异常复杂且成本高昂，也无法为在线和移动客户提供服务所需的响应速度。于是，该公司决定评估一款高级解决方案，该解决方案能够以传统 FC SAN 成本的几分之一来达成其目标。

解决方案

该银行评估了基于 SMB Direct 的 Microsoft 存储空间解决方案，这是 Microsoft 最新的存储解决方案。存储空间是 Windows Server 2012 R2 的一个组成部分，包含传统存储所提供的全部功能，其中包括重复数据删除、灾难恢复和高级服务质量。（有关更多详情，请参阅 Microsoft 白皮书“Windows Server 2012 R2 存储”¹）。该解决方案为客户提供了使用低成本行业标准硬件的能力，他们可以使用 SMB Direct（基于 RDMA 的 SMB 3.0）轻松进行横向扩展。在 Microsoft Azure 云中使用该解决方案时，其性能提高了一倍（与传统 SAN 体系架构相比），并且存储购置成本降低了 50%。

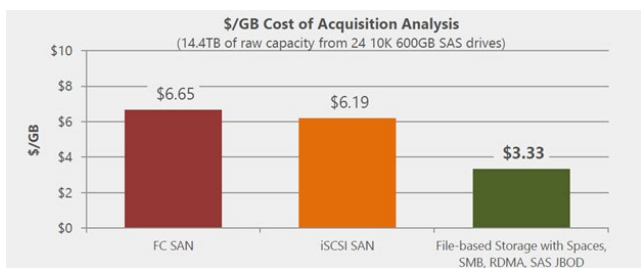


图 1. 模式方案之间的成本购置比较¹

为了将当前基于 FC 的 SAN 部署与基于存储空间的部署进行比较，该公司决定向现有基础架构添加一台文件服务器，该服务器使用 Mellanox 的端到端 40GbE 解决方案连接到 SMB Direct，然后在两种配置上运行数项基准测试。

如图 2 所示，当前的设置使用两种不同的网络类型。第一个网络是 10GbE 网卡和交换机，第二个网络是 4 个 8Gb/s FC HBA。当然，这使得该解决方案成本更高，更难以管理，而且扩展不太灵活。



图 2. 当前基于 FC 的 SAN 部署

如图 3 所示，将该设置与基于 SMB Direct 运行的存储空间进行比较，后者使用 Mellanox 端到端 40GbE 解决方案，其中包括 40GbE SX1012 交换机、ConnectX-3 40GbE 网卡和 LinkX 40GbE 铜缆。

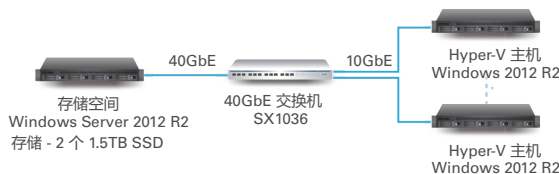


图 2. 使用基于 SMB Direct 的存储空间部署拓扑

测试结果

以下图表显示为了比较两种部署的性能而运行的各种基准测试的结果：

基准测试	基于 FC 的 SAN	存储空间
文件复制	1GB/秒	2.5GB/秒
Ntttcp	4Gb/秒	37Gb/秒
实时虚拟机迁移	90 秒	17 秒
实时存储迁移	327 秒	50 秒
虚拟机创建	13:30 分	10:45 分

这些结果表明存储空间提供的性能优于基于 FC 的 SAN。存储空间还将互连端口数量从 4 个削减到 1 个，并允许使用商用组件替代专有存储系统，使成本削减了 50%。

在审查了这些结果之后，该大型金融机构决定进行下一阶段的评估：用基于 SMB Direct 的 Microsoft 存储空间（使用 Mellanox 端到端 40GbE RoCE）替换基于 FC 的 SAN。

摘要

现代数据中心需要处理大量数据，因此必须采用新的存储技术，以消除互连瓶颈，并以更低的成本实现更高的性能和灵活的可扩展性。基于 SMB Direct 的 Microsoft 存储空间提供能够满足所有这些需求的横向扩展存储。

Mellanox 的端到端 RDMA 网络解决方案提供以下功能：

- 最快的互连适配器，每个端口为 10、25、40、56 和 100Gb/s 带宽，延迟小于 1 微秒
- 单个线路上的服务器到服务器和服务器到存储 I/O 整合
- 完全操作系统旁路和远程直接内存访问 (RDMA)
- 基于无损互连技术的卓越高可用性解决方案
- 经济高效的高密度交换机和网络体系架构

Mellanox RDMA 技术已获得 Microsoft 资格认证，并且完全集成在 Windows Server 2012 R2 中。该技术已经用于 Microsoft Azure 等大规模部署，并实现了高可靠性和无风险的部署。

使用基于 RDMA 的存储空间解决了存储瓶颈，并以 50% 的成本实现了更高的性能，这有助于企业保持数据中心效率领先一代的水平。

参考资料

¹ http://www.transtec.de/fileadmin/Medien/products/storage/Windows_Server_2012_R2_Storage_White_Paper.pdf

² http://www.mellanox.com/related-docs/applications/SB_Accelerating_Remote_Storage_Access.pdf



北京市朝阳区望京东园七区保利国际广场 T1 15 层

Tel: 010-5789 2000
www.mellanox.com

版权所有 © 2015, Mellanox Technologies. 保留所有权利。

Mellanox、Mellanox 徽标、BridgeX、ConnectX、CORE-Direct、InfiniBridge、InfiniHost、InfiniScale、PhyX、SwitchX、Virtual Protocol Interconnect 和 Voltaire 是 Mellanox Technologies, Ltd. 的注册商标。Connect-IB、CoolBox、FabricT、MetroX、MLNX-OS、ScalableHPC、Unbreakable-Link、UFM 和 Unified Fabric Manager 是 Mellanox Technologies, Ltd. 的商标。所有其他商标均为其各自所有者的资产。