

为下一代数据中心打造更快的互连

摘要

随着数据洪流冲刷当今的数据中心，IT 基础架构可以从更快速的互连中受益。更快的存储需要更快的网络。通过使用 iSER —— 一项通过 RDMA（远程直接内存访问）延伸 iSCSI 的成熟标准，可以实现更高的性能。使用 iSER，高性能存储可以通过 iSCSI 连接到快速以太网链路，从而加快从网络到服务器和存储系统的数据传输。这些技术可以合并使用，以取代老化的高速互连，例如光纤通道链路和更旧的以太网链路。

Insight

作者：Jean S. Bozman

简介

一波变化浪潮正在袭击数据中心，随之而来的将是多 TB 数据洪流。云计算、移动性、社交媒体和大数据/分析，这些新数据源的出现正在推动着这一快速变化。现在，除了事务处理、企业应用和数据仓库，还必须支持所有这些工作负载。

对于许多 IT 专业人士来说，新需求将需要他们“重新思考”其数据传输基础架构。光纤通道互连和较慢的以太网链路渐显老态，这些老化的链路必须尽快更新或更换。否则，这些慢链路将成为“瓶颈”，阻碍依赖于该 IT 基础架构的企业或组织的发展。

这一数据中心转换过程要求进行周密的计划，但是 IT 组织等不起太长的时间，否则任务关键型应用和数据库的性能将受到影响。

闪存存储和您的网络

在此过程中，许多技术都需要更加有效，包括存放数据的存储系统，并且需要将其发送到计算引擎进行处理。

如下图 1 所示，存储介质正在发生改变，更多的数据中心安装了固态驱动器 (SSD)，甚至转为使用 Non-Volatile Memory Express (NVMe) 技术来优化 SSD 性能。如图所示，随着这些新技术的使用，数据访问时间显著降低，并且 SSD 的容量逐渐变大，而不仅仅是依赖于旋转的机械硬盘驱动器 (HDD)。



在服务器和存储系统中快速采用 SSD 可以获得更快的数据传输速率（每秒 I/O，或 IOPS），但是网络互连本身必须跟上节奏。

清点您的数据中心互连

现在 IT 组织是时候清点自己目前的互连情况了，评估一下需要对哪些互连进行更新，以便跟上日益增长的需求，快速传输大型数据集和支持高 IOPS 要求的应用。

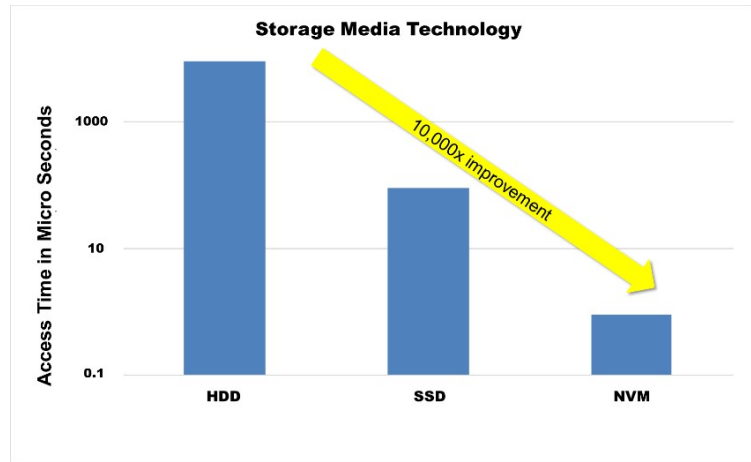


图 1：存储介质技术发展所带来的访问时间缩短
(Mellanox 2015)

SSD（固体驱动器）现在已经变成了主流，地位介于 DRAM 内存和 HDD（硬盘驱动器）之间。值得注意的是，DRAM 技术快于 SSD，而 SSD 又快于 HDD – 而 DRAM 与基于旋转机械驱动器的 HDD 相比，速度要快出更多。

新的 NVMe 技术甚至比许多闪存存储系统中使用的 SAS/SATA SSD 更快。今天，最快的闪存设备支持微秒的数据访问时间，而不是毫秒 – 如果将 NVM 技术与古老的 HDD 技术相比，访问时间剧减了 10,000 倍。



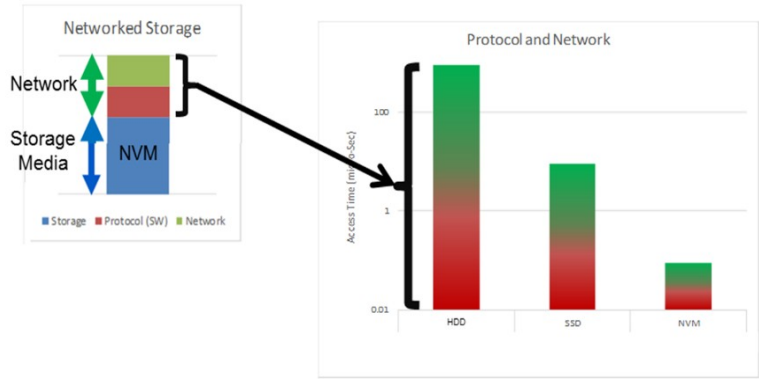


图 2：提高后的存储速度与目前的网络速度相比
(Mellanox 2015)

这一存储技术的改进将焦点转到了网络上，将它作为下一个改进领域。要检查的网络技术包括协议堆栈软件，以及网络组件的连线和交换速度。

上面的图 2 显示了提高后的存储速度与目前的网络速度选项之间的差距。此差距导致应用无法完全利用现在数据中心内可以获得的更快的存储访问时间。

为了匹配存储的改进，数据中心必须配备更快的网络组件和更快的协议软件，通过硬件卸载过程来加速。值得注意的是，如果不能通过更快的互连来加速网络连接速度，那么 IT 组织为提高计算和存储速度而做的努力最终可能会被浪费掉。

更快的以太网 iSCSI 和 iSER 技术的影响

过去，为了消除与较慢网络关联的瓶颈，许多组织都愿意为其数据中心存储支付高额的光纤通道 (FC) 技术费用。虽然 FC 过去为任务关键型系统提供高速数据传输发挥了很大的作用，但是计算和存储正在发生改变，如今越来越多的应用依赖于通过以太网连接的较小联网服务器集群。

为新工作负载（大数据、云计算、超融合基础架构和软件定义存储）相关的系统部署使用单独的 FC 网络通常并不实用，并且造价也经常过于高昂。另外，已经安装的许多 FC 基础架构仍在以每链路 4 或 8 Gbps 的速度运行，或者最大 16 Gbps（吉比特每秒）。

这就是为什么许多 IT 组织希望通过添加升级的技术来加速广泛安装的以太网的原因所在。现在，以太网的速度为 10、25、40、50 或 100 Gbps，要大大超过 FC。可使用 iSCSI 技术将新的高性能存储连接到以太网，该技术是一个行业标准接口，它允许通过以太网局域网 (LAN)、广域网 (WAN) 或互联网发送 SCSI 数据。

值得注意的是，在所有服务器操作系统和大多数存储阵列设备上本机支持 iSCSI 协议。升级支持 SSD 的系统的网络将给基于网络的应用带来显著的性能提高。

iSER：它是什么？

为了进一步提高数据中心内数据密集型工作负载的性能，您可以在更高速度的以太网上使用 iSER（针对 RDMA 的 iSCSI 扩展）协议。iSER 是一个使用远程直接内存访问 (RDMA) 扩展 iSCSI 的新标准。

iSER 可提高性能并减少网络延迟，因为不再需要 TCP/IP（互联网传输协议），它在互联网存在之前就一直在使用了。iSER 技术的发明背景为：由于以太网在全世界企业数据中心内的使用更加广泛，因此需要针对以太网修改最初针对高性能计算中所使用的 InfiniBand 高速互连而开发的 SRP（SCSI RDMA 协议）技术。

图 3 概述了 iSER 的工作原理。添加 iSER 技术可减少与 TCP/IP 堆栈相关的延迟，因此比 iSCSI 独自使用甚至更能提高闪存存储的性能。



特别是，iSER 将 iSCSI 协议扩展为可使用 RDMA。RDMA 由 RoCE（融合以太网上的 RDMA）提供，它不需要 TCP 层，因为可提供更低的延迟。它允许将数据直接传输到 SCSI 计算机内存缓冲区（将计算机连接到存储设备）和从中传输出来，无需中间进行数据复制。

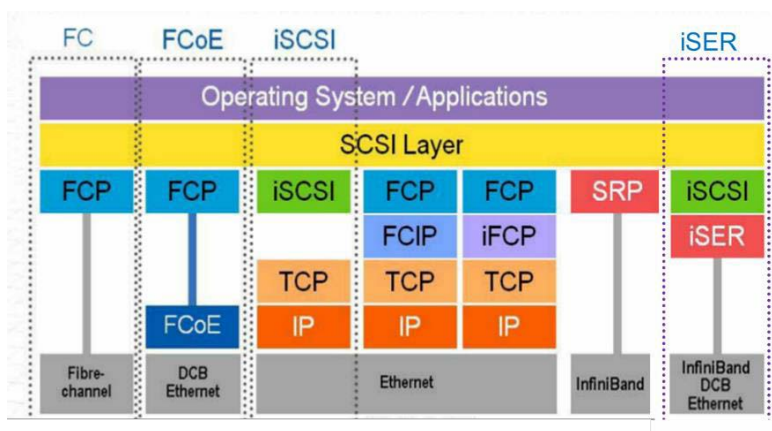


图 3：iSER 的工作原理 (Mellanox 2015)

要在数据中心内部署 iSER，IT 组织需要支持使用基于 SCSI 的块存储的应用、支持 iSER 启动器的操作系统或 Hypervisor、支持 iSER 的目标，以及支持 RDMA 的网络。

如今数据中心内的存储网络

许多数据中心都为其任务关键型系统准备了 FC 存储区域网络 (SAN)。它们长期用来对数据中心内的企业工作负载提供存储支持。今天，随着 IT 经理寻求支持更高性能存储所需的替代方案，FC 互连即将被以太网上的高速 iSCSI 或 iSER 所取代，它们向数据中心内的企业系统提供了另一种方式支持高数据传输速率。

下面的图 4 显示了如今数据中心内各种网络技术选择之间的重叠情况。图 4 中的红色标签表示存储协议，图中的紫色标签表示数据传输选项。



此边栏描述了最近几年出现的许多重要以太网技术增强功能。总的来说，这些增强功能让高性能的以太网链路成为一种替代老化光纤通道互连的可行方案。

- VLAN 和质询握手身份验证协议 (CHAP) 可确保隔离存储发起程序和目标，并确保它们像光纤通道网络上的分区和 CHAP 一样相互信任。VLAN 还支持服务质量 (QoS) 功能，以避免数据传输中的带宽瓶颈和过度抖动。
- 一个称为“数据中心桥接”的以太网协议提供了许多的可靠性优势。通过启用通道协议选项（它们让数据在网络没有空间之前先不进入网络），这些优势让以太网成为一个无损网络。
- 现在，iSCSI 和 iSER 支持在光纤通道上运行的基于服务器、设备以及阵列的存储管理软件。这包括能够从 SAN 以及 VMware 阵列集成 API (VAAI) 基元启动，支持原子测试/设置、全复制、块置零、精简配置和其他存储管理功能。
- 自 2013 年就已经推出的 40 Gbps 以太网技术已经像光纤通道一样将以太网的性能发展到了一定的高度。如今，40 Gbps 以太网技术正在以超过两倍的 16 Gbps 光纤通道线速的速度运行。

以太网技术广泛应用于全球数据中心，性能仍在提升中。随着在 2015 年引入的 50 和 100Gb 以太网升级，40Gb 以太网不再是地球上最快的以太网。而且值得注意的是 iSCSI 能够轻松地匹配光纤通道 IOPS 水平。通过在 iSCSI 上更大幅度地减少 IOPS 和延迟，使用 iSER 甚至可以获得更高的性能。

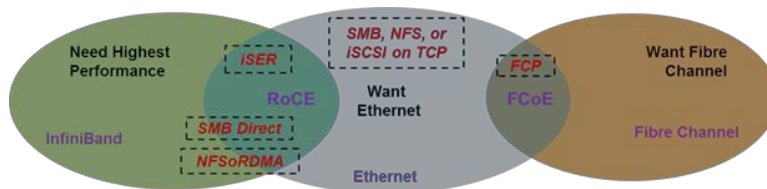


图 4：数据中心内的网络技术选择 (Mellanox 2015)

通过在使用 iSCSI 或 iSER 时利用现有的以太网技术集合，数据中心可以在已安装很长时间的旧数据中心基础架构与新兴技术之间实现更高水平的标准化和互操作性。

考虑抛弃光纤通道的客户可以评估以太网现在提供的企业级功能 [请参见本页左侧的边栏]。自光纤通道早期就出现的以太网协议增强功能可提供高水平性能、可靠性、安全性、服务质量 (QoS) 和存储管理，满足企业任务关键型工作负载预期所需。

在您的数据中心内是什么样子

将这种方法带到数据中心互连中将给 IT 组织带来许多业务优势，如下所示：

- 这种方法与他们已配备的技术相“配合”，即以以太网和 iSCSI。不必进一步改变互连基础架构，从而可避免淘汰并更换项目，造成业务中断。
- 可以利用管理软件和现有 IT 技术集合来使用更快的互连技术，从而降低运营成本。
- 任务关键型数据的安全保护将被保留，从而无需提供新的和/或专门的数据安全软件。
- 可用性/可靠性基础架构将保持原样，因为这些全都能够利用基于现有以太网的基础架构和技术。



用于高速 iSCSI 和 iSER 的 Mellanox 产品

Mellanox 提供了一套完整的互连解决方案产品组合，可以针对 iSCSI 和 iSER 标准支持高效、高性能的存储网络。Mellanox 还在扩大存储合作伙伴的生态系统，支持其产品采用 iSER 和 iSCSI 标准。

Mellanox 产品组合包括：

- 支持 iSER 和 iSCSI 的适配器，提供各种速度和端口配置，以太网从 10 到 100 Gbps。
- 适配器软件驱动程序可用于大多数的流行操作系统和 Hypervisor，包括 Linux、Microsoft Windows、VMware 和 FreeBSD，在 Linux、VMware 和 FreeBSD 中具有 iSER 开源支持。
- 在以太网上支持 iSCSI 和 iSER 的交换机，借助于行业中最高的吞吐量、最低的延迟和最低的功耗支持无损网络。
- 具有最高质量且错误率低于标准的铜缆、光缆和收发器，从而确保对一系列的以太网速度提供行业领先的支持：10、25、40、50 和 100 Gbps。

结论

数据洪流导致了数据中心内存储容量的快速上升。为了避免高性能应用出现存储瓶颈，IT 组织正在采用固态闪存存储 – 但是这造成了网络瓶颈。更快的存储需要有更快的网络。

数据中心内的这些开发项目使得许多组织开始“清点”其网络互连，从而确定需要将哪些链路升级或更换为更快的技术。否则，较慢的网络基础架构将妨碍数据访问速度从更快的存储中获益。因此，引入了采用 iSCSI 或 iSER 的更快以太网，以启用固态闪存存储性能。



数据中心内的 IT 专业人士发现，iSER 可以利用熟悉的基于 iSCSI 存储设备路径的兼容性、安全性和易用性。重要的是，iSER 所在的基础架构支持其他类型的存储和服务器流量，从而使它得到广泛应用，并且不会中断现有的业务。

通过提供技术配合基于以太网的广泛网络产品，Mellanox 在现有的数据中心基础架构上进行构筑，让提供的产品既可以提高性能，同时又能够避免通过淘汰并更换的方式进行技术更新。



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

本出版物中的信息按“原样”提供。Neuralytx, Inc. 对本出版物中的信息不做任何形式的任何表示或保证，并且明确表示，对特定用途的适销性或适用性不做任何默示保证。Neuralytx 相信，本文所包含的信息截止到出版日期为止准确无误。信息如有变更，恕不另行通知。

Neuralytx、Neuralytx 徽标、Hex 徽标、Neuralytx iQ 是 Neuralytx, Inc. 的注册商标或商标。本文所使用的所有其他商标均为其各自所有者的资产。

