

Mangstor 低延迟存储和 Mellanox 低延迟网络上的 MySQL 数据库加速

概述

每个企业都具有改善 TCO 的目标。在 MySQL 数据库应用程序中，目标是提高业务效率。实现此目标的方式可以是支持提供最大每秒事务数、缩短用户响应时间并更快执行数据库查询和分析的联机事务处理 (OLTP) 服务。此外，解决方案必须可进行扩展，因为大数据的容量、种类和速度导致传统数据库应用程序速度减慢。这因此会降低整体组织竞争力并增加用户挫折感。

超低延迟 MySQL 加速

对灵活性和实时数据访问的需求促使企业要求其 MySQL 群集环境提供显著提高的性能。通过添加 Mangstor NX6325 固态阵列以提供第 0 层块存储卷，可以在虚拟化环境中显著提高性能并提高 VM 性能密度，从而针对 OLTP 和联机分析处理 (OLAP) 工作负载支持更多且更大的横向扩展 MySQL 实例。

通过将 Mellanox 的端到端 100Gb/s 低延迟以太网解决方案（包括 Spectrum 交换机、ConnectX-4 100Gb/s RoCE（融合以太网上的 RDMA）网卡和 LinkX 电缆）与 Mangstor 的 NX6325 可扩展 NVMe 固态存储阵列相结合，可加速实时数据访问以便缩短获得见解的时间并提高业务工作效率。在 StorageReview 实验室进行的测试中，MySQL 数据文件、tempDB 文件和日志文

件全部位于 32 个 NX6325 卷上，生成 38K 以上的每秒事务数 (TPS)。这样可降低总体拥有成本 (TCO) 并提高业务智能 (BI) 效率。

低延迟解决方案

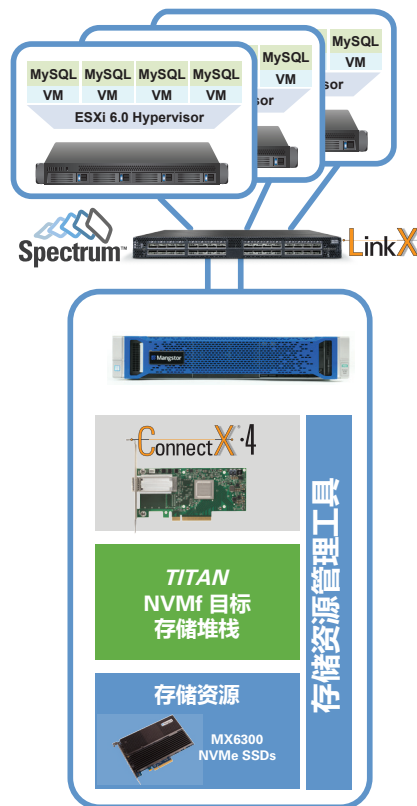
来自存储或其他服务器的低延迟数据请求是在缩短用户响应时间的情况下实现更多 MySQL OLTP TPS 的关键。Mangstor Inc. 和 Mellanox Technologies 是低延迟存储和网络解决方案领域的领先者，将 Mangstor 存储解决方案与 Mellanox 高性能网络结合在一起提供 NVMe over Fabric 存储阵列解决方案。Mangstor 的

MX6300 NVMe 固态驱动器可使用服务器的 PCIe 插槽提供高带宽低延迟闪存存储连接。采用远程直接内存访问 (RDMA) 技术的 Mellanox 100GbE 可提供低延迟网络访问，并通过最少服务器计算资源实现最快且最可靠的数据传输。

NVMe over Fabric 存储

NVMe over Fabric 存储使具有网络接口卡的主机服务器可以与远程闪存存储子系统连接，以便在物理和虚拟数据中心中实现更短的应用程序响应时间和更好的可扩展性。虽然现在已存在光纤通道 (FC) 和 iSCSI 网络结构上的 SSD 远程访问，不过该功能基于针对提供有限性能和延迟的 HDD 而设计的旧式 SCSI 协议。

通过在高速 RDMA 横向扩展应用程序中使用融合以太网上的 RDMA (RoCE) 将数据从应用程序服务器传输到存储服务器，可绕过较低级别的 SCSI 传输层，从而可加快数据吞吐量并减少数据访问延迟。与直接连接存



储 (DAS) 不同, NVMe over RDMA 技术使闪存存储的共享池可以独立于计算资源进行扩展, 反之亦然, 以便服务器可以访问数据中心中任何位置的闪存, 而不会以形成额外延迟为代价。

RoCE 利用了可高效实施以太网上的 RDMA 的以太网发展, 从而可以在主流数据中心应用程序中广泛部署 RDMA 技术。基于 RoCE 的网络管理也与任何以太网网络管理相同, 从而使 IT 经理无需了解新技术。

用于数据中心的 NVMe over RoCE 的业务优势包括:

- 无需更改数据中心基础架构
- 采用 10/25/40/50/100 GbE 以太网上的单个线路实现 I/O 统一
- 继续采用现有的数据中心管理基础架构
- 降低功耗并节省成本
- 维持现有和未来的应用程序兼容性
- 通过实现 I/O 统一的单芯片解决方案来显著节省资本支出和运营支出

高性能可以经济实惠

为了使用 Mangstor 的存储阵列和 Mellanox 的端到端 RoCE 解决方案评估 MySQL OLTP 性能, 使用了搭载两个 Xeon E5-2630v3 处理器并运行 TITAN NVMe over Fabric (NVMe) 目标软件存储堆栈的 HP DL380 G9 服务器。TITAN 软件通过将 NVMe SSD 与 RDMA 网络接口卡进行紧密集成, 来提供领先的性能和超低延迟。

采用 1U 尺寸规格的 Mellanox SN2700 32 端口 100GbE 交换机允许使用 RoCE 的多台 ESXi 服务器与 NX6325 存储阵列通信。Mellanox ConnectX®-4 适配器支持服务器之间的 10/25/40/56/100Gb/s RoCE 存储连接 (全部连接到 Mellanox 的 LinkX 电缆)。

评估 MySQL OLTP 性能

通过使用 StorageReview.com Sysbench 基础架构 MySQL OLTP 负载生成器生成使用 Percona MySQL 应用程序的典型 DB 用户的实际工作负载, 度量了 MySQL OLTP 性能。下面的图表显示在 8 台 Dell R720 服务器上并行运行的 4 到 32 个虚拟机 (每个客户端具有 27 毫秒响应时间)。

这些结果表示与利用相同设置的竞争对手解决方案相比, 效率显著提升。

• 每个 VM 的事务处理能力最高

在 Storage Review 进行的 Sysbench OLTP 工作负载测试中, 32 个 VM 的 NX6325 总分数达到 38K 以上的每秒事务数 (TPS), 处理能力比最接近的测试固态阵列高三倍。

• 最低的一致延迟和响应时间

在 Storage Review 进行的 Sysbench 按比例平均延迟测试中, NX6325 在所有 VM 间的延迟小于 27 毫秒, 总分数表示响应时间比最接近的测试固态阵列快两到三倍。

• 可扩展的经济高效的固态存储解决方案

NX6325 存储阵列采用 2U 机架封装提供, 具有 10.8TB 或 21.6TB 可用闪存存储, 只需通过添加更多阵列便可扩展到数百 TB 容量。Mangstor NX6325 是支持当前存在的数据库大小以及客户在未来需要的大小的存储解决方案。Mangstor 先进的 NVMe over RDMA 技术提供的性能比直接连接服务器中闪存提高三倍, 同时仍可提供本机 MySQL OLTP 功能。

» **应用程序加速** — 通过更快地响应时间获得更好的客户体验

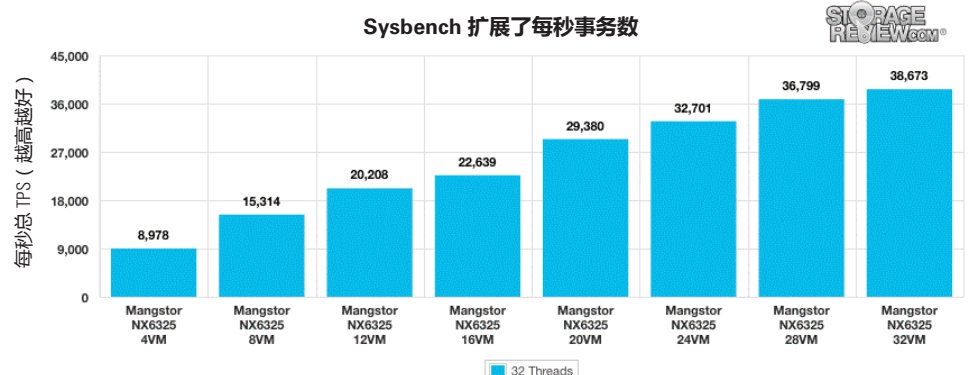


图 1. 通过 Mangstor NX6325 存储和 Mellanox 的端到端 100Gb/s RoCE 解决方案运行 Percona MySQL 应用程序的 storagereview.com 性能

- » **提高业务工作效率** — 通过增强分析快速进行决策
- » **虚拟机群集可扩展性** — 大规模的较高应用程序性能
- » **固态存储可扩展性** — 通过四个 NX6325 将容量向上扩展到 100TB

解决方案摘要

使用 Mangstor 的 NX6325 阵列和 Mellanox 端到端 RoCE 解决方案（包括 Spectrum 以太网交换机、ConnectX-4 网卡和 LinkX 电缆）的高性能存储子系统可提升存储系统效率；从而最大程度实现投资回报 (ROI) 并向旧有的基于光纤通道的存储子系统提供经济的备选方法。

关于 Mellanox

Mellanox Technologies (NASDAQ: MLNX) 是针对服务器、存储和超融合基础架构的端到端以太网和 InfiniBand 智能互连解决方案和服务的领先提供商。Mellanox 智能互连解决方案可提供最高吞吐量和最低延迟、更快地向应用程序传递数据并充分发挥系统性能，从而提高数据中心效率。Mellanox 提供一系列高性能解决方案：网络和多核处理器、网络适配器、交换机、电缆、软件和芯片，它们可针对广泛的市场（包括高性能计算、企业数据中心、Web 2.0、云、存储、网络安全、电信和金融服务）加快应用程序运行时间并最大程度实现业务成果。www.mellanox.com 上提供了详细信息。



北京迈络思科技有限公司

咨询电话：+86-10-57892000

销售咨询：china_sales@mellanox.com

市场合作：marketing_cn@mellanox.com

*欲了解更多欢迎登陆www.mellanox.com

