



使用LightOS 部署可靠和高性能的存储

第一个支持集群的软件定义NVMe/TCP 存储解决方案

目录

云属性Cloud-Native块存储革命	2
达到云级可靠性	2
Light OS 解决方案	3
LightOS 集群支持以防止数据丢失 LightOS	3
LightOS 的部署	5
其他 优势	5
性能和延迟	5
灵活性和效率	6
云扩展 Cloud Scale	6
结论	6

云属性Cloud-Native块存储革命

在部署任何规模云属性应用程序时，访问可靠、可执行和易于部署的基础设施是业务敏捷性和应用程序体验的关键，无论您部署的是5个节点、500个节点还是数百个节点。

众所周知，诸如Facebook、Google 和 AmazonWeb服务(AWS) 这样的云服务提供商(CSP) 破解了基础设施效率的代码，因为这对他们的业务至关重要。不太为人所知的是，他们是如何在不损害以下任何一项的情况下实现高效率的：

- 出色的性能、低延迟和可扩展性
- 部署的存储的灵活性和效率
- 很容易适应于云扩展
- 与现有数据中心基础架构兼容的情况下运行，而无需删除和更换

以上四个标准是任何云属性存储解决方案的核心需求。挑战在于，由于需要工程人才，构建这样一个存储解决方案成为一项复杂而艰巨的任务

好消息是，现在这个障碍已经被 Lightbits 开发的 LightOS 软件定义的、支持集群的存储解决方案清除了。LightOS 为云属性块存储部署设立了一个新的标准，为您的基础设施部署提供了下一步。

实现云级可靠性

可靠性是云属性块存储解决方案的关键要素，尽管它是由标准 COTS 服务器、NIC 和 NVMe SSD 等不可靠组件构建的，但解决方案必须保证具有可靠性。SSD 服务器集群必须在内部复制数据，并在出现故障时保持数据的完全一致性和可用性。

从访问数据的客户端的角度来看，数据复制必须是透明的，服务器故障转移必须是无缝的。此外，任何云属性块存储解决方案都应减少应用程序延迟，提高应用程序性能，提高可扩展性，并为现有数据中心基础设施提供传统存储区域网络(SAN)的可靠性。

大型云提供商极好的扩展性和可靠性并不是通过构建一个无限增大的大型系统来实现的，而是通过使用许多松散聚集的较小系统而构建更大的系统来实现的。类似地，任何云属性存储解决方案在由许多较小的系统组成、松散聚合和作为一个整体管理的情况下取得最好的性能。

对于云提供商来说，如果硬件故障或软件故障都将导致整个系统失效。在过去，我们用高度可用和可靠的组件建立存储区域网络(SAN)，并尽可能小心地编写无bug的软件。

在云中，您要接受一切可能失效的风险，利用不可靠的组件构建可靠的系统来实现可靠性。

LightOS - 解决方案

LightOS 是 Lightbits Labs™ 软件定义的存储解决方案，它为云和数据中心基础设施的分离存储提供了一场革命性的方法：一种只用软件或者软件加硬件加速的解决方案。

LightOS 利用高性能的 NVMe 解决方案，最大限度地提高资源利用率和灵活性，挖掘出分离存储的潜力。它使数据中心能够从一个低效的直接附加的 SSD 模型转移到一个共享模型，在这个模型中，计算和存储可以独立地进行有效的扩展，通集群功能使得这样做并不会降低系统的可靠性。

LightOS 安装在大型数据中心的标准服务器上，为 I/O 密集型计算集群如 Cassandra、MySQL、FDB、MongoDB 和时间序列数据库进行了优化。它基于有专利技术的多 NVMe SSD 管理和创新的 NVMe over TCP (NVMe/TCP) 存储通信协议提供高性能和低延迟。Lightbits NVMe/TCP 不需要对网络基础设施或应用服务器软件进行任何更改。LightOS 可以提供数百万 IOPS，时延可以低到几十微秒。

LightOS 非常适合于诸如 Kubernetes 之类的容器环境，其需要具有持久存储的大规模群集，以实现快速节点迁移、工作负载重新平衡或从故障恢复，而无需在网络上复制数据

LightOS 群集最大限度地提高了资源利用率和灵活性，并提供了高可用性和耐用性，每个 volume 副本数可单独配置。Light OS 群集可以由驻留在不同故障域中并位于数据中心内任何位置的服务器组成。

此外，LightOS 还提供了丰富的存储服务，例如数据压缩、thin provisioning 以实现 SSD 的最高利用率、以及用于保护服务器和 SSD 故障的复制和擦除编码 (EC)，以确保在故障下的服务数据可用性。

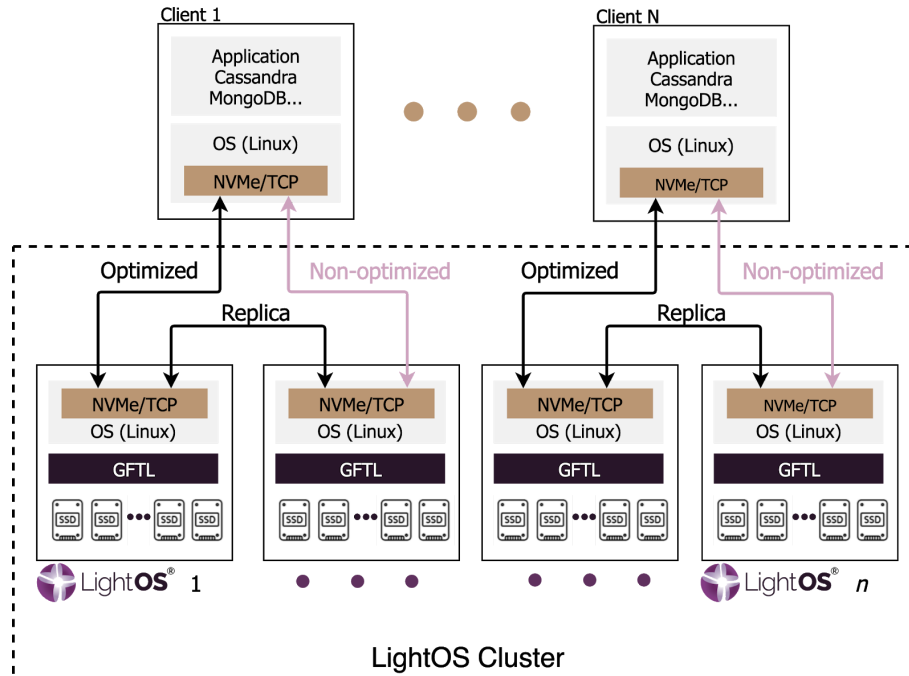
LightOS 集群支持以防止数据丢失

通常，为了防止数据丢失，软件定义的存储解决方案在不同的存储服务器上复制数据，并在存储服务器在经历短暂或者永久故障的情况下保证服务和数据的可用性。请记住，在云规模下，任何东西都可能失效。一组 Light OS 服务器在内部复制数据，并保持数据完全一致和可用，从而保证服务的可靠性和数据可用性，即使一个或多个存储服务器出现临时故障或永久性失效。

从客户端访问数据的角度来看，Light OS 集群体系结构是透明的，在发生故障时服务器失效接管是无缝的。数据保持一致和可用。客户端不需要任何特殊的驱动程序或代理，只要使用所有主要 Linux 发行版中现在可用的标准 NVMe/TCP 驱动程序。在 Light OS 集群支持下，客户端 CPU 用于应用程序，而不是用于存储处理。

LightOS 群集可以由驻留在不同故障域中并位于数据中心内任何位置的服务器组成。该体系结构最大限度地提高了资源利用率和灵活性，并提供了高可用性和耐用性，每个卷具有可配置的副本数。

图1：LightOS 集群多路径复制设计

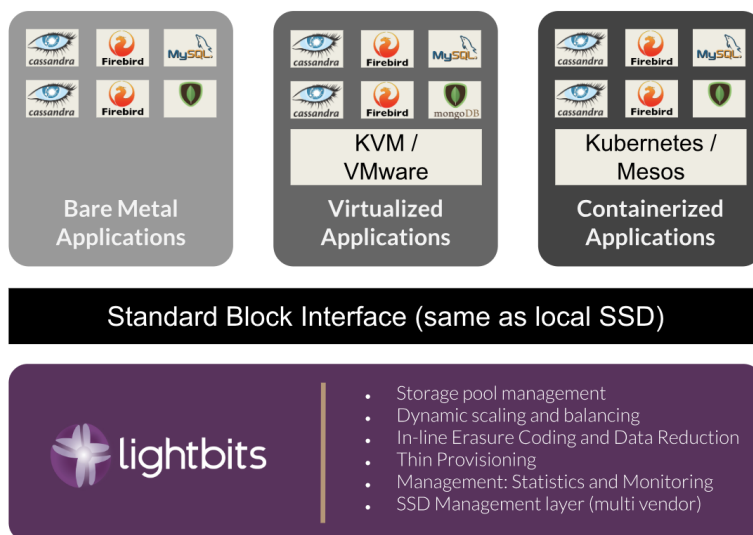


除了它的可靠性外，**LightOS** 为云数据中心提供了与直接连接的 **NVMe SSD** 相同的每秒输入输出操作 (**IOPS**)，并且与直接连接存储**DAS** 相比，尾部延迟减少了 **50%**。它还可以独立扩展存储和计算，从而提高基础架构利用率和效率，降低总拥有成本 (**TCO**)，并提供无与伦比的灵活性。

通过提供高度可用且持久的存储层，应用程序团队可以专注于开发新服务，而**LightOS** 确保数据的可用性和高性能访问。

LightOS群集支持临时禁用对存储服务器的访问的无中断维护例程（例如，TOR 固件升级）。在这种情况下，客户端继续以降级模式与存储服务器的子集透明地工作，直到维护例程完成为止，并且所有存储服务器都变得可访问。

图2 LightOS 框图



LightOS 的部署

它的核心是一个软件定义的分层存储解决方案，它是现有数据中心基础结构的自然发展。

LightOS使用 NVMe/TCP 块存储访问协议，用于应用服务器到存储服务器和存储服务器到存储服务器之间的通信链路。NVMe/TCP 运行在任何 TCP/IP 网络上，并且已包括在某些版本的操作系统和 hypervisor 中不久将包括在每个操作系统和 hypervisor 中。

该软件只需使用标准的NVMe/TCP 驱动程序就可以在现有的应用服务器和现有的 TCP/IP 网络上简单方便地部署。此外，在同一个云数据中心中也可以存在多个 LightOS 集群，这些数据中心可以很容易地作为一个大型云块存储解决方案进行聚合和管理。

LightOS是在用现成的零部件和产品制造的商品硬件上运行的。它作为一个持久的在数据和控制路径上无单点故障点的解决方案，可以帮助云供应商在任何时候都保持他们的服务不中断。

其他优势

除了其可靠性外，具有集群特性的 LightOS 还为构建云属性块存储基础结构提供了以下几个优点：

性能和延迟

当今云中流行的存储架构使用直接连接的硬盘驱动器和SSD。特别是以合理的成本提供高性能和低延迟的 NVMe SSD。

任何云属性 块存储 解决方案都不能在应用程序当前所经历的级别之外降低性能或增加延迟。

LightOS 给应用程序提供具有与直接附加结构相同的性能级别和类似的延迟。

灵活性和效率

当前直接附加的存储部署常常受到利用率低和缺乏灵活性的影响。低使用率意味着购买的基础设施处于闲置状态。缺乏灵活性意味着运营费用高于应有的水平，因为很难适应不断变化的工作负载条件，并且很难根据需要增加存储或计算能力用直接附加的存储，你必须存储和计算始终一起扩展。

低效率和僵化的基础设施都是高TCO 的原因。

LightOS 使用 分离 存储模型，通过将计算和 存储分离来提供灵活性和效率。这种效率的提高会降低您的基础设施的 TCO。

云扩展 Cloud Scale

云实现其大规模 地 扩展不是通过构建一个无限期增长的大系统，而是通过从许多松散聚集的较小系统中构建更大的系统来实现。

LightOS 集群支持从许多较小的存储服务器构建存储集群，松散聚合并作为一个整体进行管理。

结论

对于企业和云提供商来说，构建云属性 块存储集群的时机已经到来。 LightOS 集群体系结构为您的基础结构提供了更高的灵活性、效率和可靠性，同时提供了更好的延迟和更低的整体TCO 。 LightOS 群集可以帮助您管理一个或多个服务器故障，而不会丢失任何服务或数据。将在 2019 年后期为生产部署提供 LightOS 集群支持。

About Lightbits Labs™

当今的存储方法是为企业设计的，无法满足不断发展的云规模基础架构要求。例如，SAN因缺乏性能和控制能力而广为人知。从规模上讲，直接连接的固态硬盘（DAS）变得过于复杂，无法顺利运行，成本也很高，并且固态硬盘利用率低下。

云规模的基础架构需要对存储和计算进行分解，顶级云巨头从低效的直接连接SSD架构过渡到低延迟共享NVMe闪存架构证明了这一点。

与其他NVMe-oF方法不同，Lightbits NVMe / TCP节省成本的解决方案将存储和计算分开，而不会影响网络基础架构或数据中心客户端。Lightbits团队成员是NVMe标准的主要贡献者以及NVMe over Fabrics (NVMe-oF) 的发起者之一。

现在，Lightbits正在制定新的NVMe / TCP标准。作为该领域的开拓者，Lightbits解决方案已经在行业领先的云数据中心中成功进行了测试。

该公司的共享NVMe架构提供了有效而强大的分解。如此平稳的过渡过程，您的应用团队甚至都不会注意这一变化。他们现在可以以比本地SSD更好的尾部延迟发狂！最后，您可以将存储与计算分开，而不会造成麻烦。

Finally, you too can separate storage from compute without drama.

 www.lightbitslabs.com

 info@lightbitslabs.com

US Office
1830 The Alameda,
San Jose, CA 95126, USA

Israel Offices
17 Atir Yeda Street,
Kfar Saba, Israel 4464313
3 Habankim Street,
Haifa, Israel 3326115

The information in this document and any document referenced herein is provided for informational purposes only, is provided as is and with all faults and cannot be understood as substituting for customized service and information that might be developed by Lightbits Labs Ltd for a particular user based upon that user's particular environment. Reliance upon this document and any document referenced herein is at the user's own risk.

The software is provided "As is", without warranty of any kind, express or implied, including but not limited to the warranties of merchantability, fitness for a particular purpose and non-infringement. In no event shall the contributors or copyright holders be liable for any claim, damages or other liability, whether in an action of contract, tort or otherwise, arising from, out of or in connection with the software or the use or other dealings with the software.

Unauthorized copying or distributing of included software files, via any medium is strictly prohibited.

COPYRIGHT (C) 2019 LIGHTBITS LABS LTD. - ALL RIGHTS RESERVED