



概述和要求

BeeGFS on NetApp with E-Series Storage

NetApp

January 27, 2026

目录

- 概述和要求 1
 - 解决方案概述 1
 - NVA计划 1
 - 设计概述 1
 - 用例 1
 - 优势 1
 - 架构概述 2
 - 构建块架构 2
 - 文件系统服务 2
 - HA架构 3
 - 已验证节点 4
 - 经验证的硬件设计 4
 - Ansible 6
- 技术要求 6
 - 硬件要求 6
 - 软件和固件要求 7

概述和要求

解决方案概述

NetApp解决方案 上的BeeGFS将BeeGFS并行文件系统与NetApp EF600存储系统相结合、打造了一个可靠、可扩展且经济高效的基础架构、可满足苛刻的工作负载要求。

NVA计划

NetApp解决方案 上的BeeGFS是NetApp验证架构(NVA)计划的一部分、该计划为客户提供针对特定工作负载和使用情形的参考配置和规模估算指导。NVA解决方案经过全面测试和设计、可最大限度地降低部署风险并加快上市速度。

设计概述

基于NetApp的BeeGFS解决方案是一个可扩展的组件架构、可针对各种要求苛刻的工作负载进行配置。无论是处理多个小文件、管理大量大文件操作、还是混合工作负载、都可以对文件系统进行自定义以满足这些需求。高可用性内置在设计中、采用双层硬件结构、支持在多个硬件层进行独立故障转移、并确保性能稳定一致、即使在部分系统降级期间也是如此。BeeGFS文件系统支持在不同的Linux分发版之间实现高性能和可扩展的环境、并为客户端提供一个易于访问的存储命名空间。要了解更多信息，请访问 ["架构概述"](#)。

用例

以下使用情形适用于NetApp解决方案 上的BeeGFS：

- NVIDIA DGX SuperPOD系统采用采用DGX的A100、H100、H200和B200 GPU。
- 人工智能(AI)、包括机器学习(ML)、深度学习(DL)、大规模自然语言处理(NLP)和自然语言理解(NLU)。有关详细信息，请参见 ["BeeGFS for AI：事实与虚构"](#)。
- 高性能计算(HPC)、包括通过MPI (消息传递接口)和其他分布式计算技术加速的应用程序。有关详细信息，请参见 ["BeeGFS为何超越HPC"](#)。
- 应用程序工作负载的特征为：
 - 读取或写入大于1 GB的文件
 - 多个客户端(10、100和1000)读取或写入同一文件
- 多TB或多PB数据集。
- 需要一个存储命名空间的环境、可以优化大小文件的组合。

优势

在NetApp上使用BeeGFS的主要优势包括：

- 提供经验证的硬件设计、可实现硬件和软件组件的全面集成、以确保可预测的性能和可靠性。
- 使用Ansible进行部署和管理、以实现规模化的精简性和一致性。
- 使用E系列性能分析器和BeeGFS插件提供监控和可观察性。有关详细信息，请参见 ["介绍用于监控NetApp E系列解决方案的框架"](#)。

- 高可用性采用共享磁盘架构、可提供数据持久性和可用性。
- 支持使用容器和Kubernetes进行现代工作负载管理和流程编排。有关详细信息，请参见 ["Kubernetes认识BeeGFS：一个适应未来需求的投资故事"](#)。

架构概述

NetApp解决方案 上的BeeGFS包括用于确定支持经验证的工作负载所需的特定设备、布线和配置的架构设计注意事项。

构建块架构

根据存储要求、可以使用不同的方式部署和扩展BeeGFS文件系统。例如、以大量小文件为主要特征的使用情形将受益于额外的元数据性能和容量、而大文件较少的使用情形则可能有利于提高实际文件内容的存储容量和性能。这些多种注意事项会影响并行文件系统部署的不同维度、从而增加了文件系统设计和部署的复杂性。

为了应对这些挑战、NetApp设计了一个标准的构建块架构、用于横向扩展其中的每个方面。通常、BeeGFS组件部署在以下三种配置文件之一中：

- 一个基础组件、包括BeeGFS管理、元数据和存储服务
- BeeGFS元数据加上存储组件
- 仅供BeeGFS存储使用的组件

这三个选项之间的唯一硬件更改是使用较小的驱动器来获取BeeGFS元数据。否则、所有配置更改将通过软件应用。借助Ansible作为部署引擎、为特定组件设置所需的配置文件可以使配置任务变得简单明了。

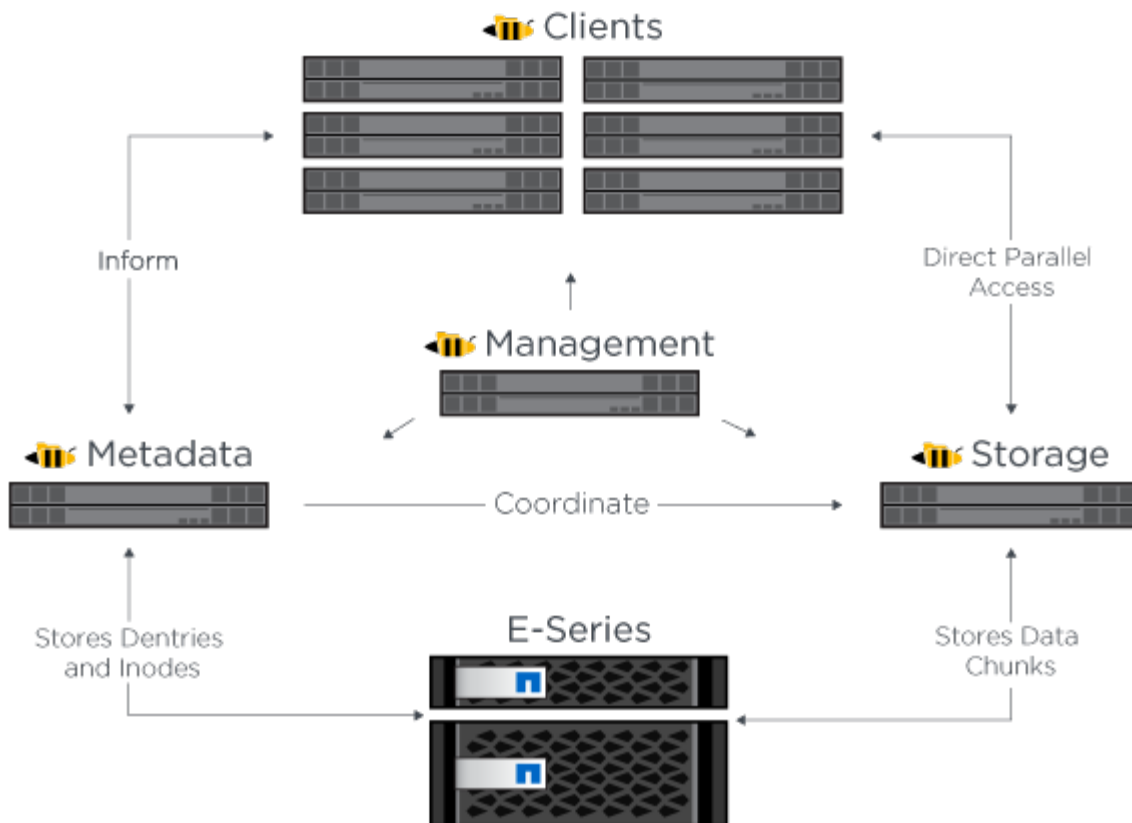
有关更多详细信息、请参见 [\[经验证的硬件设计\]](#)。

文件系统服务

BeeGFS文件系统包括以下主要服务：

- *管理服务。*注册并监控所有其他服务。
- *存储服务。*存储称为数据区块文件的分布式用户文件内容。
- *元数据服务。*跟踪文件系统布局、目录、文件属性等。
- *客户端服务。*挂载文件系统以访问存储的数据。

下图显示了NetApp E系列系统中使用的BeeGFS解决方案 组件和关系。



作为一个并行文件系统、BeeGFS会将其文件条带化到多个服务器节点上、以最大限度地提高读/写性能和可扩展性。服务器节点协同工作、可提供一个文件系统、其他服务器节点(通常称为_clients_)可以同时挂载和访问该文件系统。这些客户端可以查看和使用分布式文件系统、类似于NTFS、XFS或ext4等本地文件系统。

这四项主要服务在广泛支持的Linux分发版上运行、并通过任何支持TCP/IP或RDMA的网络进行通信、包括InfiniBand (IB)、OMNI-Path (OPA)和基于融合以太网的RDMA (RoCE)。BeeGFS服务器服务(管理、存储和元数据)是用户空间守护进程、而客户端是原生 内核模块(无修补)。所有组件均可在不重新启动的情况下安装或更新、您可以在同一节点上运行任何服务组合。

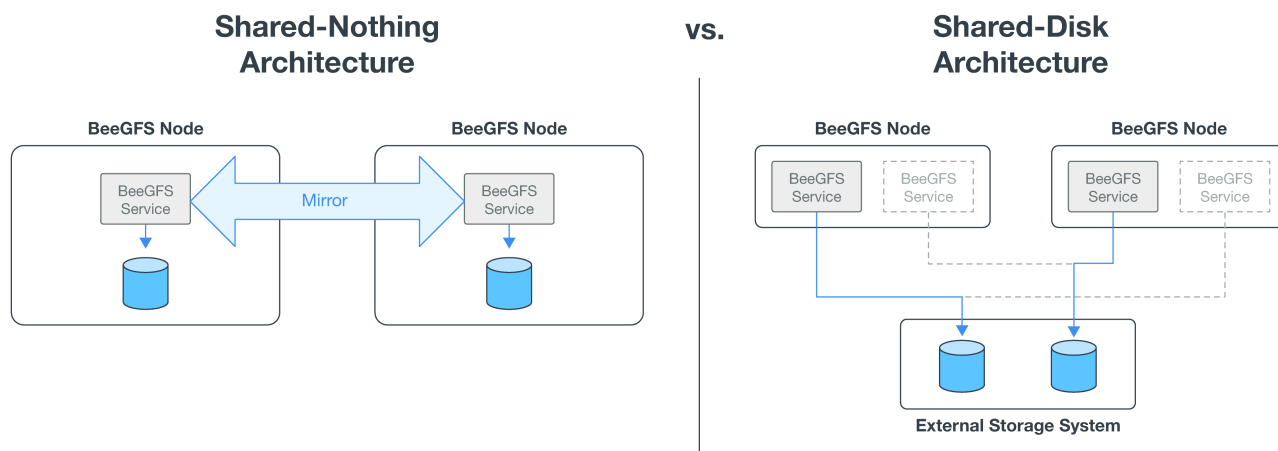
HA架构

NetApp上的BeeGFS通过创建与NetApp硬件完全集成的解决方案 来扩展BeeGFS企业版的功能、从而实现共享磁盘高可用性(HA)架构。



虽然BeeGFS社区版可以免费使用、但企业版要求从NetApp等合作伙伴购买专业支持订阅合同。企业版支持使用多项附加功能、包括故障恢复能力、配额强制实施和存储池。

下图对无共享和共享磁盘HA架构进行了比较。



有关详细信息，请参见 ["宣布NetApp支持的BeeGFS高可用性"](#)。

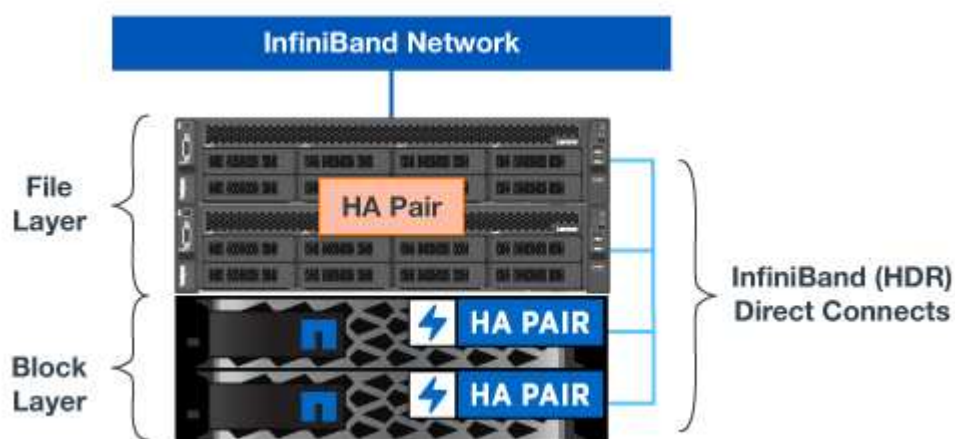
已验证节点

NetApp上的BeeGFS解决方案已验证以下节点。

节点	硬件	详细信息
块	NetApp EF600存储系统	一款高性能全NVMe 2U存储阵列、专为要求苛刻的工作负载而设计。
文件	Lenovo ThinkSystem SR6SR V3服务器	双插槽2U服务器、采用PCIe 5.0、双AMD EPYC 9124处理器。有关Lenovo SR更多功能的详细信息，请参阅 "联想的网站" 。
	Lenovo ThinkSystem SR6SR服务器	双插槽2U服务器、采用PCIe 4.0和双AMD EPYC 7003处理器。有关联想SR使用的详细信息，请参阅 "联想的网站" 。

经验证的硬件设计

该解决方案的组件(如下图所示)使用经过验证的文件节点服务器作为BeeGFS文件层、并使用两个EF600存储系统作为块层。



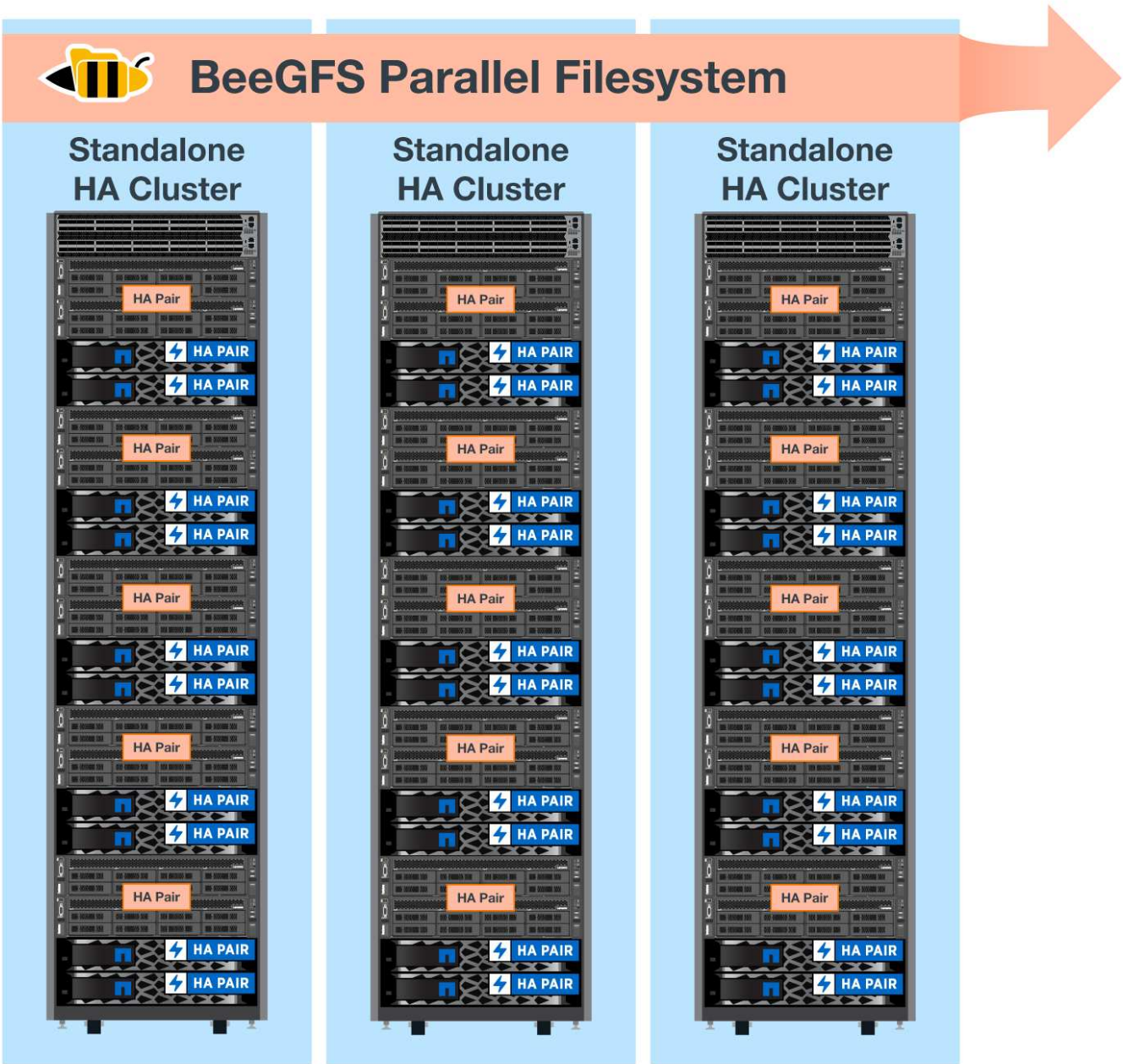
NetApp解决方案 上的BeeGFS可跨部署中的所有组件运行。部署的第一个组件必须运行BeeGFS管理、元数据

和存储服务(称为基础组件)。所有后续组件均可通过软件进行配置、以扩展元数据和存储服务、或者专门提供存储服务。通过这种模块化方法、可以根据工作负载的需求扩展文件系统、同时使用相同的底层硬件平台和组件设计。

最多可以部署五个组件、以构成一个独立的Linux HA集群。这样可以优化PacMaker的资源管理、并保持与Corosync的高效同步。将一个或多个独立BeeGFS HA集群组合在一起、以创建一个BeeGFS文件系统、此文件系统可作为一个存储命名空间供客户端访问。在硬件方面、一个42U机架最多可容纳五个组件、并可为存储/数据网络配备两个1U InfiniBand交换机。有关可视化表示、请参见下图。



要在故障转移集群中建立仲裁、至少需要两个组件。双节点集群存在一些限制、可能会阻止成功进行故障转移。您可以通过将第三个设备整合为Tiebreak 机来配置双节点集群；但是、本文档不会介绍这种设计。



Ansible

NetApp上的BeeGFS是使用Ansible自动化交付和部署的、Ansible自动化托管在GitHub和Ansible GALAXY上(BeeGFS集合可从获取 "[Ansible Galaxy](#)" 和 "[NetApp的E系列GitHub](#)") 。虽然Ansible主要是使用用于组装BeeGFS组件的硬件进行测试的、但您可以将其配置为使用受支持的Linux版本在几乎任何基于x86的服务器上运行。

有关详细信息，请参见 "[使用E系列存储部署BeeGFS](#)"。

技术要求

要实施基于NetApp的BeeGFS解决方案、请确保您的环境满足本文档中概述的技术要求。

硬件要求

开始之前、请确保您的硬件满足基于NetApp的BeeGFS解决方案的第二代组件设计的以下规格。根据客户要求、特定部署的确切组件可能会有所不同。

数量	硬件组件	要求
2.	BeeGFS文件节点	每个文件节点都应满足或超出建议文件节点的规格、以实现预期性能。 建议的文件节点选项: • *Lenovo ThinkSystem SR国际 电脑公司 ◦ 处理器： 2个AMD EPYC 9124 16C 3.0 GHz (配置为两个NUMA分区)。 ◦ 内存： 256 GB (16个16 GB TruDR5 4800MHz RDIM-A) ◦ *PCIe扩展： *四个PCIe Gen5 x16插槽(每个NUMA分区两个) ◦ 其他： ▪ 在RAID 1中为操作系统配置两个驱动器(1 TB 7.2K SATA或更高) ▪ 用于带内操作系统管理的1GbE端口 ▪ 采用Redfish API的1GbE BMC、用于带外服务器管理 ▪ 双热插拔电源和性能风扇
2.	E-Series块节点(EF600阵列)	内存： 256 GB (每个控制器128 GB)。适配器： 2端口200GB/HDR (NVMe/IB)。*驱动器： *配置为匹配所需元数据和存储容量。
8.	InfiniBand主机卡适配器(适用于文件节点)。	主机卡适配器可能根据文件节点的服务器型号而有所不同。针对经验证的文件节点的建议包括： • Lenovo ThinkSystem SR6SR V3服务器: ◦ MCX755106AS-热连接X-7、NDR200、QSFP112、双端口、PCIe Gen5 x16、 InfiniBand适配器

数量	硬件组件	要求
1.	存储网络交换机	存储网络交换机必须支持200 Gb/秒的InfiniBand速度。建议的交换机型号包括： * NVIDIA QM9700昆特2 NDNS InfiniBand交换机* * NVIDIA MMM4700多模HDR InfiniBand交换机*

布线要求

从块节点直接连接到文件节点。

数量	部件号	长度
8.	NVIDIA 1500-H001E30 (无源铜缆、QSFP56、200 GB/秒)	1 个月

*从文件节点连接到存储网络交换机。*根据您的InfiniBand存储交换机、从下表中选择适当的缆线选项。+建议的电缆长度为2米；但是、这可能会因客户环境而异。

交换机型号	电缆类型	数量	部件号
NVIDIA QM9700	有源光纤(包括收发器)	2.	MMA4Z00-NS (多模式、IB/ETH、800 Gb/秒2x400 Gb/秒双端口OSFP)
		4	MFP7E20-Nxxx (多模式、4通道到2通道分路器光缆)
		8.	MMA1Z00-NS400 (多模式、IB/ETH、400 Gb/秒单端口QP-112)
	无源铜线	2.	NVIDIA 7Y40-N002 (InfiniBand无源铜缆分路器电缆、InfiniBand 800Gb/s到4个200Gb/s、OSFP到4个QSFP112)
NVIDIA MqM8700	活动光纤	8.	MFS1S00-H003C (NVIDIA有源光纤电缆、InfiniBand 200Gb/秒、QFP56)
	无源铜线	8.	NVIDIA 1500-H002E26 (无源铜缆、InfiniBand 200Gb/秒、QSFP56)

软件和固件要求

为了确保可预测的性能和可靠性、基于NetApp的BeeGFS解决方案版本会使用特定版本的软件和固件组件进行测试。实施该解决方案需要这些版本。

文件节点要求

软件	version
Red Hat Enterprise Linux (RHEL)	RHEL 9.4 高可用性物理服务器（双插槽）。*注：*文件节点需要有效的 Red Hat Enterprise Linux Server 订阅和 Red Hat Enterprise Linux 高可用性附加组件。
Linux内核	5.14.0-427.42.1.el9_4.x86_64

软件	version
HCA固件	ConnectX-7 HCA 固件 FW: 28.45.1200 + PXE: 3.7.0500 + UEFI: 14.38.0016 ConnectX-6 HCA 固件 FW 20.43.2566 + PXE 3.7.0500 + UEFI 14.37.0013

EF600块节点要求

软件	version
SANtricity 操作系统	11.90R3
NVSRAM	N6000-890834-D02.DLP
驱动器固件	适用于所用驱动器型号的最新版本。请参见" E-Series磁盘固件站点 "。

软件部署要求

下表列出了在基于Ansible的BeeGFS部署中自动部署的软件要求。

软件	version
BeeGFS	7.4.6
核心同步	3.1.8-1.
起搏器	2.1.7-5.2.
件	0.11.7-2.
Fence Agent (红鱼/APC)	4.10.0-62.
InfiniBand / RDMA驱动程序	MLNX_OFED_LINUX-23.10-3.2.2.1-LTS

Ansible控制节点要求

NetApp解决方案 上的BeeGFS可从Ansible控制节点进行部署和管理。有关详细信息，请参见 "[Ansible文档](#)"。

下表中列出的软件要求特定于下面列出的NetApp BeeGFS Ansible集合版本。

软件	version
Ansible	10.x
Ansore-core	> 2.13.0
Python	3.10
其他Python软件包	加密法-43.0.0、netaddr-1.3.0、ipaddr-2.2.0
NetApp E-Series BeeGFS折叠资料集	3.2.0

版权信息

版权所有 © 2026 NetApp, Inc.。保留所有权利。中国印刷。未经版权所有者事先书面许可，本文档中受版权保护的任何部分不得以任何形式或通过任何手段（图片、电子或机械方式，包括影印、录音、录像或存储在电子检索系统中）进行复制。

从受版权保护的 NetApp 资料派生的软件受以下许可和免责声明的约束：

本软件由 NetApp 按“原样”提供，不含任何明示或暗示担保，包括但不限于适销性以及针对特定用途的适用性的隐含担保，特此声明不承担任何责任。在任何情况下，对于因使用本软件而以任何方式造成的任何直接性、间接性、偶然性、特殊性、惩罚性或后果性损失（包括但不限于购买替代商品或服务；使用、数据或利润方面的损失；或者业务中断），无论原因如何以及基于何种责任理论，无论出于合同、严格责任或侵权行为（包括疏忽或其他行为），NetApp 均不承担责任，即使已被告知存在上述损失的可能性。

NetApp 保留在不另行通知的情况下随时对本文档所述的任何产品进行更改的权利。除非 NetApp 以书面形式明确同意，否则 NetApp 不承担因使用本文档所述产品而产生的任何责任或义务。使用或购买本产品不表示获得 NetApp 的任何专利权、商标权或任何其他知识产权许可。

本手册中描述的产品可能受一项或多项美国专利、外国专利或正在申请的专利的保护。

有限权利说明：政府使用、复制或公开本文档受 DFARS 252.227-7013（2014 年 2 月）和 FAR 52.227-19（2007 年 12 月）中“技术数据权利 — 非商用”条款第 (b)(3) 条规定的限制条件的约束。

本文档中所含数据与商业产品和/或商业服务（定义见 FAR 2.101）相关，属于 NetApp, Inc. 的专有信息。根据本协议提供的所有 NetApp 技术数据和计算机软件具有商业性质，并完全由私人出资开发。美国政府对这些数据的使用权具有非排他性、全球性、受限且不可撤销的许可，该许可既不可转让，也不可再许可，但仅限在与交付数据所依据的美国政府合同有关且受合同支持的情况下使用。除本文档规定的情形外，未经 NetApp, Inc. 事先书面批准，不得使用、披露、复制、修改、操作或显示这些数据。美国政府对国防部的授权仅限于 DFARS 的第 252.227-7015(b)（2014 年 2 月）条款中明确的权利。

商标信息

NetApp、NetApp 标识和 <http://www.netapp.com/TM> 上所列的商标是 NetApp, Inc. 的商标。其他公司和产品名称可能是其各自所有者的商标。