

联想 ThinkAgile HX 技术白皮书

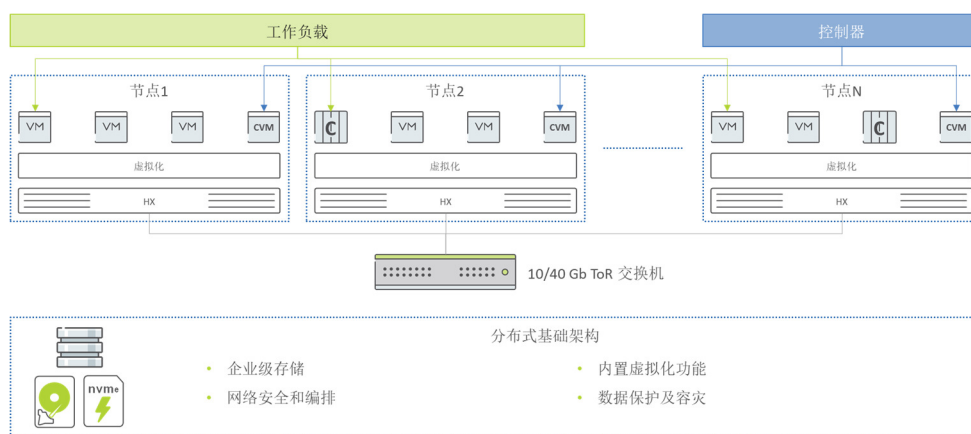
1. 软硬件一体化，超融合设备出厂预装超融合软件和虚拟化软件。

联想 ThinkAgile HX 系列是一款一流的超融合系统，在联想平台上预装了业界领先的 Nutanix 软件，可大大简化数据中心管理、减轻 IT 员工压力和加快部署进度，从而最终降低总体拥有成本。

联想 ThinkAgile HX 系列超融合设备可以根据需要出厂预装 Nutanix AHV 虚拟化或者 VMware vSphere ESXi 虚拟化软件。

2. 企业级分布式存储系统架构。

联想 ThinkAgile HX 系列超融合拥有自主知识产权，提供企业级分布式文件解决方案，不是基于 GlusterFS 或者 Ceph 等开源产品的二次开发。无需专用硬件，也不依赖于传统的 RAID 技术，通过软件技术将多个计算存储一体化服务器所有节点的本地磁盘空间整合成一个整体存储空间提供给虚拟化软件使用。单节点故障不会影响到整体存储空间的使用，不会造成数据错误或者数据丢失。



3. 企业级虚拟化平台 AHV

联想 ThinkAgile HX 系列超融合平台与计算无关，可以灵活地选择适合您当前需求的虚拟机管理程序和云服务组合，并可以自由地按需移动工作负载。联想 ThinkAgile HX 系列超融合平台的原生虚拟化平台是 AHV，它允许管理员开箱即用，立即开始加载具有企业级虚拟化功能的虚拟机，无需额外费用。

在 AHV 部署中，控制器虚拟机（CVM）作为虚拟机运行，磁盘使用 PCI 直通呈现。这允许完整的 PCI 控制器（和连接的设备）直接传递到 CVM 并绕过虚拟机管理程序。AHV 基于 CentOS KVM。使用完全硬件虚拟化技术用于虚拟机服务（HVM）。

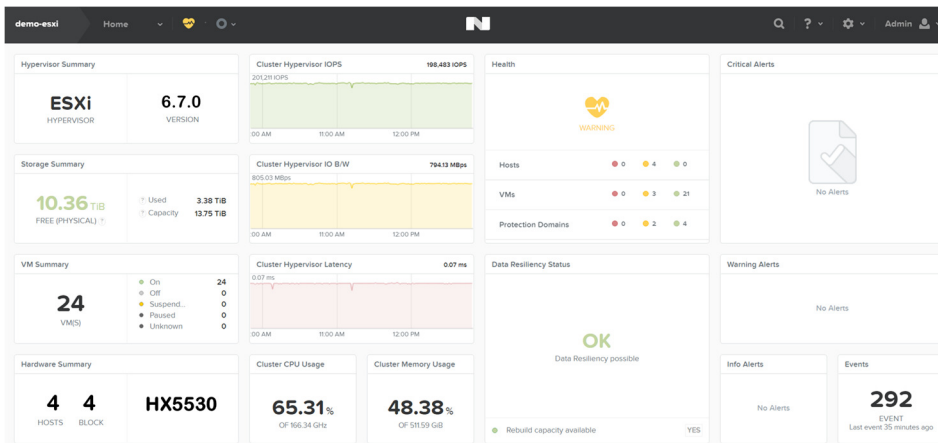
4. 超融合软件对虚拟化平台的支持

联想 ThinkAgile HX 系列超融合软件支持多种主流虚拟化平台，包括自带虚拟化 AHV，VMware vSphere ESXi 和 Microsoft Hyper-V。

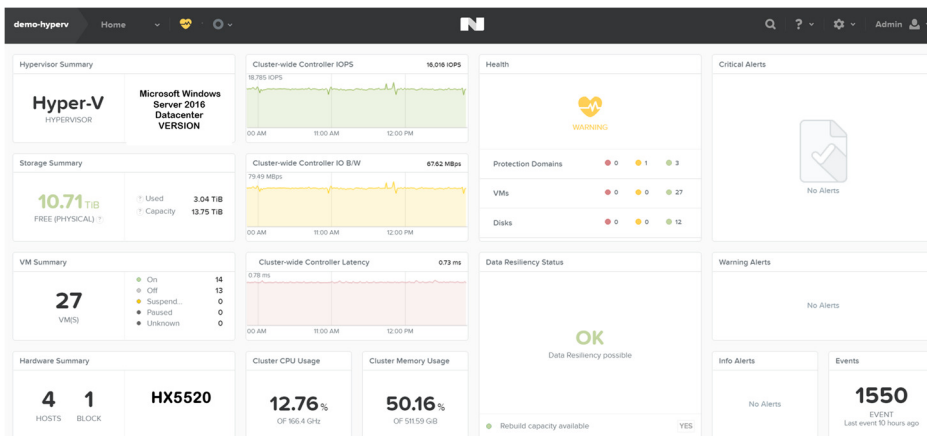
下图是自带虚拟化 AHV 的截图



下图是VMware vSphere ESXi 的截图



下图是Microsoft Hyper-V 的截图



5. 支持不同型号、不同年代的计算存储一体超融合设备的混用和兼容。
联想 ThinkAgile HX 系列超融合将服务器、存储与虚拟化软件整合到一个资源池中，在可横向扩展的集群内通过单个管理界面轻松进行管理。该系列给客户交付联想提供的完全集成且经过验证的基础架构单元，可提供极高的可靠性、安全性、可扩展性和简化的管理。

规格

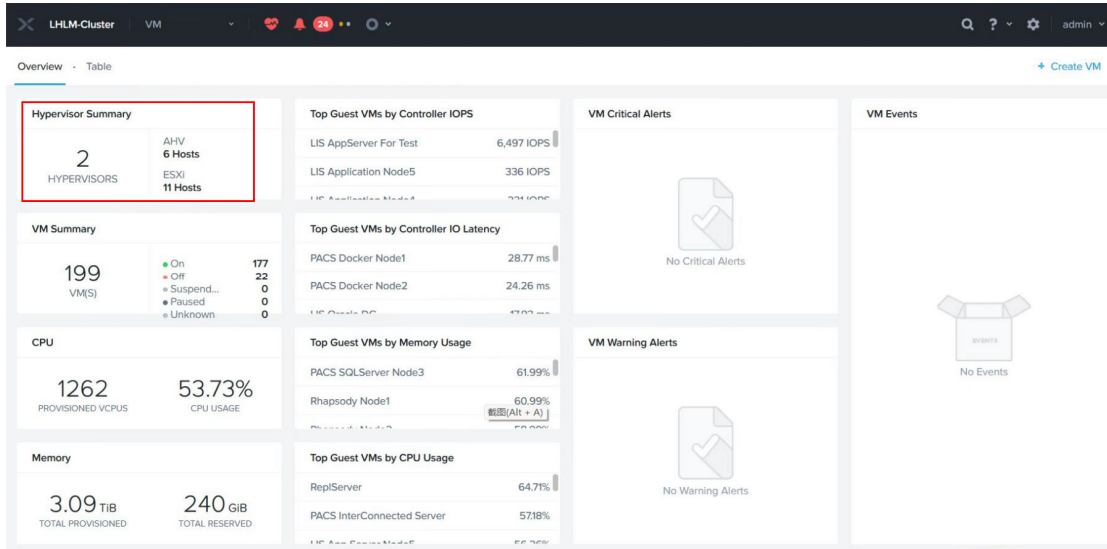
型号	HX5530	HX5520
工作负载		大数据和高容量
外形	2U	2U
处理器	2个第三代 Intel® Xeon® Silver、Gold 或 Platinum 处理器，最高可达 220W	2个第二代 Intel® Xeon® Silver、Gold 或 Platinum 处理器，最高可达 165W
内存	64GB - 4TB	128GB - 3TB
驱动器托架	12个 3.5 英寸 + 4个 2.5 英寸 (可选)	12 或 14 个 3.5 英寸
SSD 选项 (混合配置)	2-6 个 1.6TB, 1.92TB, 3.84TB, 7.68TB	2-6 个 480GB, 800GB, 960GB, 1.92TB, 3.84TB, 7.68TB
HDD 选项 (混合配置)	6-12 个 4TB, 6TB 6-10 个 8TB 8 个 10TB, 12TB, 16TB	6-12 个 4TB, 6TB, 6-10 个 8TB 8 个 10TB, 12TB, 16TB
SSD 选项 (全闪存配置)	2-16 个 1.6TB, 1.92TB, 3.84TB, 7.68TB	6-14 个 480GB, 800GB, 960GB, 1.6TB, 1.92TB, 3.84TB, 7.68TB
NVMe 选项 (全闪存配置)	无	无
GPU 选项	0 - 6 个单宽或 2 个双宽	不支持

联想 ThinkAgile HX 系列能够在同一软件管理界面管理不同型号、不同年代的计算存储一体超融合设备、进行资源调度分配等功能，不需要跳转或重新登录，这些设备可以具有不同的 CPU、内存、SSD 和磁盘配置，甚至可以将全闪节点加入到混闪集群。参见下面的 5 个基于 Intel Cascade Lake 处理器的 HX5520 和 3 个基于 Intel Ice Lake 处理器的 HX5530 形成的单集群 8 节点的统一管理模式。



6. 异构虚拟化平台支持

联想 ThinkAgile HX 超融合支持将 VMware vSphere ESXi 和 AHV 混合在同一个集群中，其中 AHV 源于 KVM 虚拟化，可以大大降低商用虚拟化许可费用。



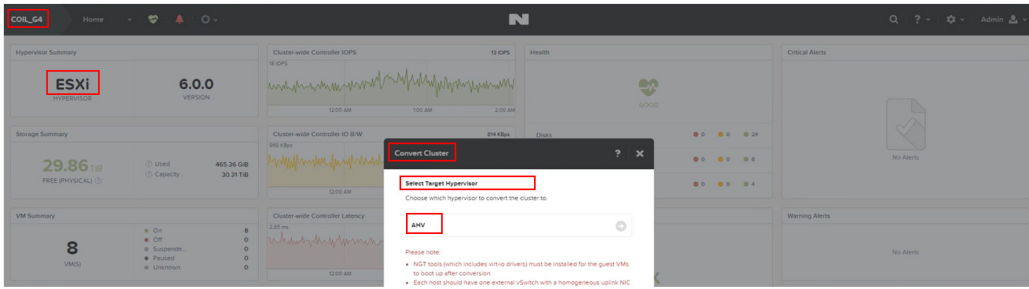
Host Name	Host IP	CVM IP	Hypervisor	CPU Usage	CPU Capacity	Memory Usage	Memory Capacity	Total Disk Usage	Disk Usage	Disk IOPS	Disk IO B/W	Disk IO Latency
NTNX-J30100R6	10.254.145.21	10.254.145.141	ESXi	65.01%	50.28 GHz	66.36%	383.6 8 GiB	19.12 TiB of 50.92 TiB	37.55%	700	46.9 MBps	2.98 ms
NTNX-J3010F3B	10.254.145.22	10.254.145.142	AHV	34.96%	16.8 GHz	44.85%	93.87 GiB	20.44 TiB of 51.97 TiB	39.32%	335	36.02 MBps	3.25 ms
NTNX-J3010F3C	10.254.145.23	10.254.145.143	AHV	31.98%	16.8 GHz	44.86%	93.87 GiB	17.11 TiB of 50.79 TiB	33.68%	283	31.97 MBps	4.21 ms
NTNX-J3010F3D	10.254.145.24	10.254.145.144	AHV	39.09%	16.8 GHz	44.86%	93.87 GiB	17.88 TiB of 50.75 TiB	35.23%	235	26.88 MBps	4.41 ms
NTNX-J3010F3E	10.254.145.25	10.254.145.145	AHV	31.53%	16.8 GHz	44.86%	93.87 GiB	22.52 TiB of 50.96 TiB	44.19%	712	24.16 MBps	1.15 ms
NTNX-J3010F3F	10.254.145.26	10.254.145.146	AHV	31.35%	16.8 GHz	44.85%	93.87 GiB	17.57 TiB of 50.75 TiB	34.61%	317	31.27 MBps	3.97 ms
NTNX-J3010F3G	10.254.145.27	10.254.145.147	AHV	41.58%	16.8 GHz	44.85%	93.87 GiB	19.34 TiB of 50.96 TiB	37.95%	242	23.64 MBps	3.5 ms

7. 一键式虚拟化平台转化

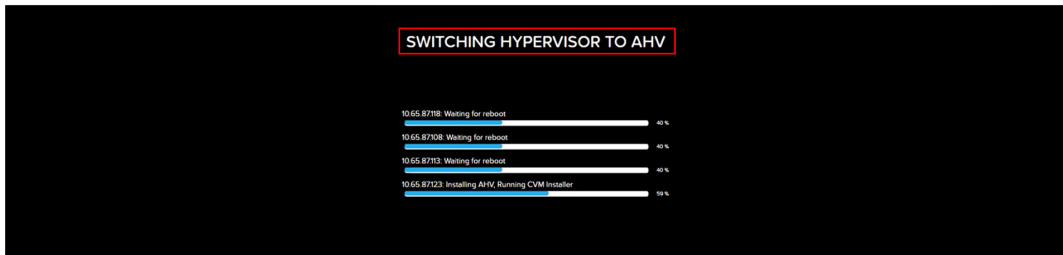
联想 ThinkAgile HX 超融合支持将 VMware vSphere ESXi 一键式转化成 AHV 虚拟化，虚拟机无需导入导出，虚拟磁盘格式自动转化。



在集群中选择 Convert Cluster



集群提示转化为 AHV 虚拟化



集群开始自动转化



转化完成后集群虚拟化转为 AHV

8. 本地集群规模支持。

联想 ThinkAgile HX 超融合在软件架构层面没有节点数量强制性限制，集群节点数量按照应用需求和维护最佳实践进行调配。

AOS 软件版本

	入门级	专业级	旗舰级
核心软件功能	核心软件功能	丰富的数据服务、恢复能力和管理功能	全套 Nutanix 软件功能，以应对复杂的基础架构挑战
适用于	运行有限工作负载的小规模部署 (仅限本地)	运行多款应用或大规模单一工作负载的部署 (本地或公有云)	多站点部署和高级安全要求 (本地或公有云)
企业存储			
集群规模	12	本地没有强制性限制。在 AWS 中验证的 16 个节点	本地没有强制性限制。在 AWS 中验证的 16 个节点在 AWS 中验证的 16 个节点
异构集群	✓	✓	✓
以虚拟机为中心的快照和克隆	✓	✓	✓
数据分层	✓	✓	✓
在线集群扩展/收缩	✓	✓	✓

9. 超融合集群支持以 1 个节点为单位进行扩容，自动发现新增节点，实现集群计算和存储资源的无缝扩展，在不中断业务的情况下将新节点/磁盘自动加入现有集群中。

联想 ThinkAgile HX 超融合集群支持以 1 个节点为单位进行扩容，自动发现新增节点，实现集群计算和存储资源的无缝扩展，参考如下集群扩展，7 节点集群在线动态扩展 1 个新节点，计算资源扩展为 8 节点，存储资源扩展为 48 块磁盘。

3.4.4 集群扩展(增加节点)

Acropolis 集群动态扩展能力是核心的功能。要扩展 Acropolis 集群，可以将节点上架、摆放和插线，然后启动节点。一旦节点被启动后，其会以 mDNS 方式自动被当前集群发现。

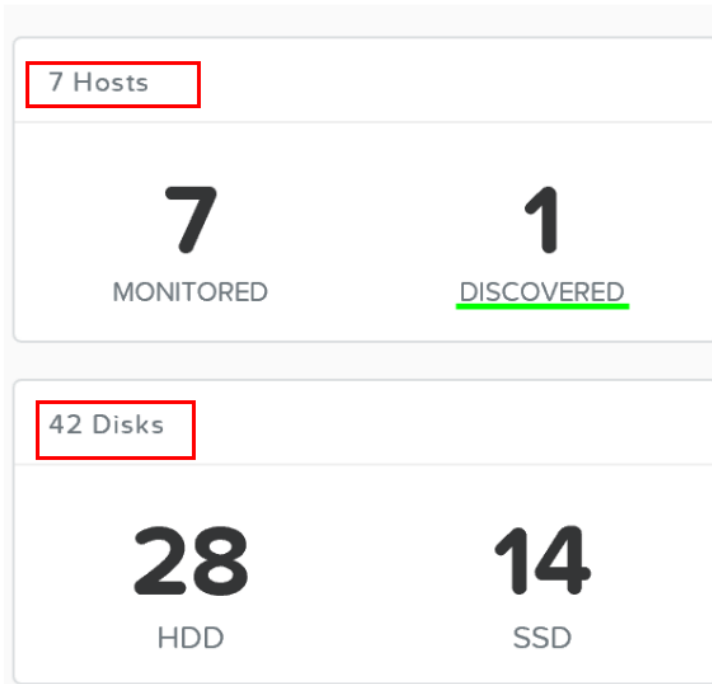


图. 动态发现节点

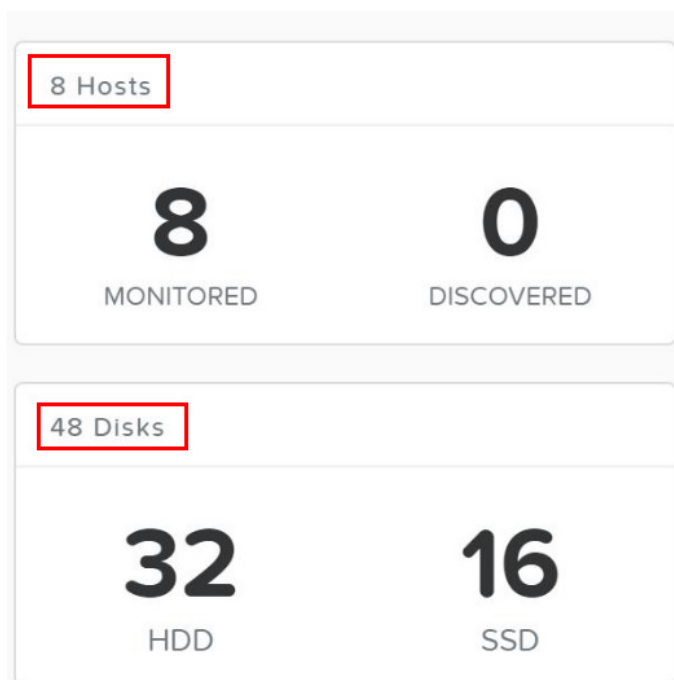


图. Expand Cluster—执行

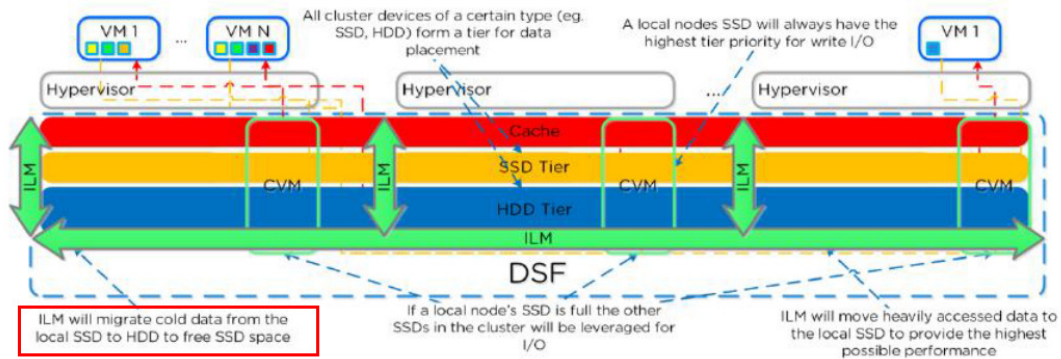
- SSD 和 HDD 混插配置中实现智能数据分层，SSD 同时可作为缓存层和持久数据存储层，软件自动将经常访问的热数据缓存到 SSD 磁盘上，而将不常用的数据放在大容量磁盘上，无需手工干预，充分利用 SSD 性能。

联想 ThinkAgile HX 超融合系统中 Extent Store 是 DSF 中持久性大容量存储组件，它横跨 SSD 和 HDD，并且能扩展到其他节点的存储设备上。有两种数据流进入 Extent Store，一种是从 Oplog 中被合并的数据顺序写入到 Extent Store，另外一中是绕过 Oplog 的顺

序写操作直接写入 Extent Store 中。超融合的 ILM (information lifecycle management) 基于 IO 类型 (IO Pattern) 动态决定数据保存的热分层位置, 并且在不同热分层之间移动数据。

Oplog 类似于文件系统的日志 (journal), 用来处理突发的写操作, 并聚合这些写操作顺序地写到 Extent Store 中。为了保证数据高可用性的要求, 在写操作完成 (Ack) 之前, 数据写入 Oplog 的同时被同步复制到另一个 CVM 的 Oplog 中。所有 CVM 的 Oplog 都参与复制操作, 并依据 CVM 负载情况自动选择。Oplog 保存在 CVM 的 SSD 层以提供极高的 IO 写性能, 特别是随机 IO 写操作。

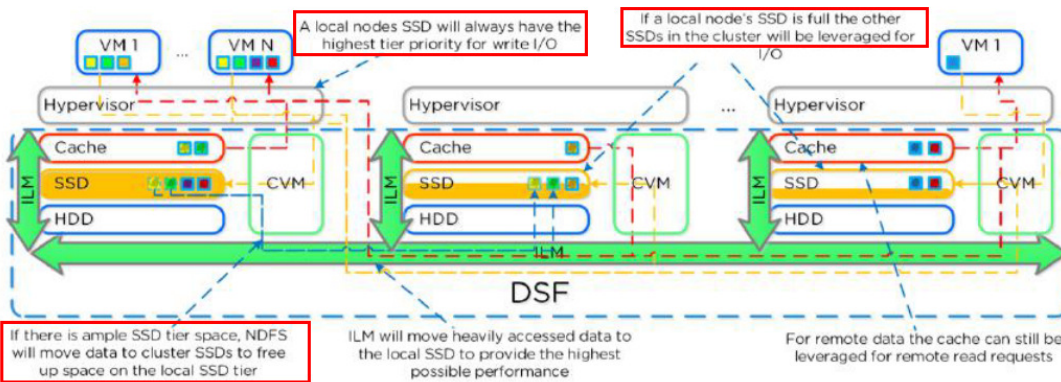
Unified Cache 是可消重的读缓存, 它横跨 CVM 的内存和 SSD。当读取的数据不在缓存中 (或基于一个特定的指纹), 数据会放到 Unified Cache 的 Single-touch 池, 该池完全保留在内存中, 并且使用 LRU (最近最少使用算法) 直到数据被弹出。



11. 节点优先使用本地 SSD 磁盘进行 IO 处理, 其次使用集群范围内其他节点 SSD 磁盘进行 IO 吞吐, 不会由于某个 SSD 容量问题而导致性能降级。

联想 ThinkAgile HX 系统中磁盘的分层是实现在集群内所有节点的 SSD 和 HDD 上, 并且由 ILM 负责触发数据迁移。本地节点的 SSD 层总是最高优先级的, 负责所有本地虚拟机 I/O 的读写操作。并且还可以使用集群内所有其他节点的 SSD, 因为 SSD 层总是能提供最好的读写性能, 并且在混合阵列中尤为重要。

当本地 SSD 利用率非常高时, 磁盘平衡功能会自动迁移 SSD 中最冷的数据到集群内另一个节点的 SSD 上。这将释放本地 SSD 的空间, 数据会继续写到本地 SSD 上, 而不会因为通过网络写到另一个节点的 SSD 上而导致性能下降。一个关键点需要提及的是所有 CVM 和 SSD 都参与这种远程 IO 操作, 可以消除潜在的瓶颈, 并且通过执行远程 IO 操作可以自愈一些命中率的问题。



12. 超融合软件在发现故障后立刻进行数据重建。集群规模越大, 数据重建越快。

联想 ThinkAgile HX 超融合系统出现计划外故障时 (在某些情况下, 如果它们无法正常工

作，系统将主动将其置于脱机状态)，立即开始重建过程。

与其他一些等待 60 分钟开始重建并且在此期间仅保留一份副本的供应商不同（风险很高，如果出现任何类型的故障，可能会导致数据丢失），我们不愿意牺牲这种风险来提升存储利用率。

联想 ThinkAgile HX 超融合系统依靠保证元数据和数据跨所有节点和所有磁盘设备分布，可以在通常的数据处理和重新保护时保证最高性能。

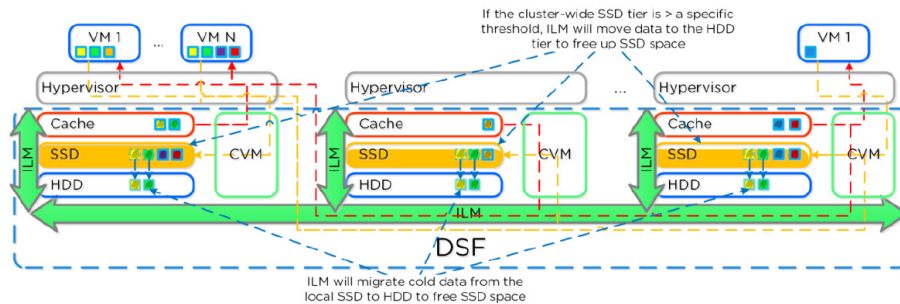
当数据进入系统，主拷贝和副本拷贝会分布在本地和所有其他远程节点。这样就消除了潜在热点（例如运行缓慢的节点和磁盘），并确保一致的写性能。

在数据磁盘或者节点故障时数据必须被重新保护，集群的全部能力都可以被用来重建。在重建事件里，元数据扫描（找出故障设备上的数据和副本的位置）平均分布在所有 CVM 上。一旦数据副本被所有健康 CVM 发现，磁盘设备（SSD+HDD）和主机网络上联链路可以被用来并行重建数据。

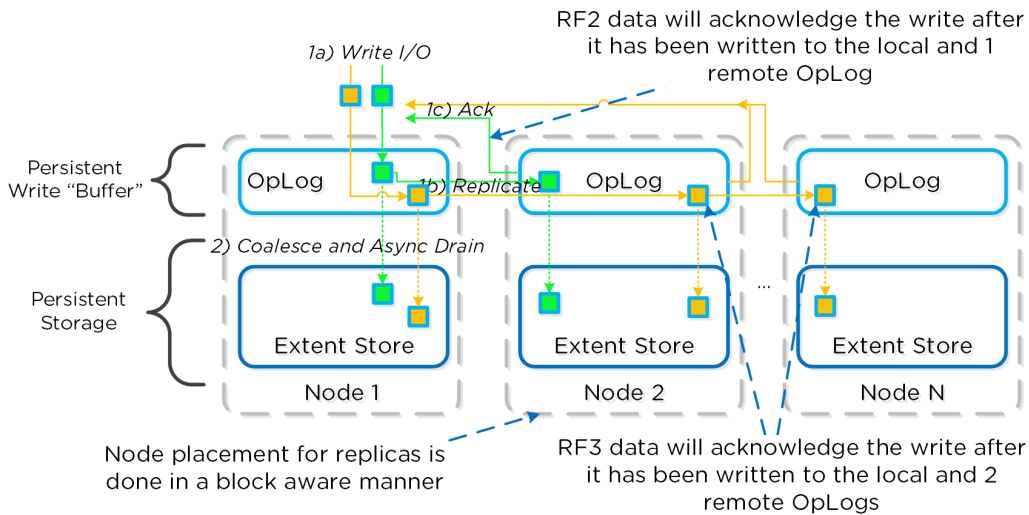
例如，在一个有磁盘故障的 4 节点集群中，每个 CVM 会处理 25%的元数据扫描和数据重建。在一个 10 节点集群中，每个 CVM 处理 10%的元数据扫描和数据重建。在一个 50 节点集群中，每个 CVM 会处理 2%的元数据扫描和数据重建。

13. 当 SSD 磁盘出现故障时，仅需重建 SSD 中的数据，无需重建 SSD+HDD 磁盘组中的所有数据，减少重建时间，降低对整体集群的影响。

联想 ThinkAgile HX 超融合系统采用分层存储架构，SSD 故障不影响同节点上的其他 SSD 和 HDD，其他 SSD 和 HDD 任然可以正常工作。系统仅需要重建故障 SSD 上的用户数据。例如一台主机 2*800G SSD+8*4T HDD，50%使用率，如果一个 SSD 故障，HX 只需要重建 400G 数据



联想 ThinkAgile HX 超融合系统使用副本技术保护用户数据，可以提供 RF2 两份数据或者 RF3 三份数据分布在不同节点不同磁盘，防止磁盘/节点故障对用户数据的影响。当有磁盘/节点故障时，其上的数据会在整个集群的其他磁盘/节点上重建，不像传统 Raid，不使用 Raid 技术或分布式 Raid 技术保护数据，所以无需配置热备盘，可以避免由于磁盘故障导致整个 Raid 组性能降级，避免影响业务正常运行。



14. 数据本地化技术

联想 ThinkAgile HX 超融合系统作为一个融合（计算+存储）平台，I/O 和数据本地化对于集群和虚拟机性能至关重要。所有读/写 I/O 都由本地控制器虚拟机（CVM）提供服务，该虚拟机位于与普通虚拟机相邻的每个虚拟化主机上。虚拟机的数据从 CVM 本地提供，并位于 CVM 控制的本地磁盘上。当虚拟机从一个虚拟化节点移动到另一个虚拟化节点时（或在 HA 事件期间），新迁移的虚拟机的数据将由现在的本地 CVM 提供。读取旧数据（存储在现在的远程节点/CVM 上）时，I/O 会由本地 CVM 转发到远程 CVM。所有写入 I/O 将立即在本地发生。AOS 将检测到 I/O 是从不同的节点发生的，并将在后台迁移数据到现在的节点，从而允许现在在本地提供所有读取 I/O。数据只会在读取时迁移，以免占用网络资源。

数据本地化有两种主要形式：

- 缓存位置

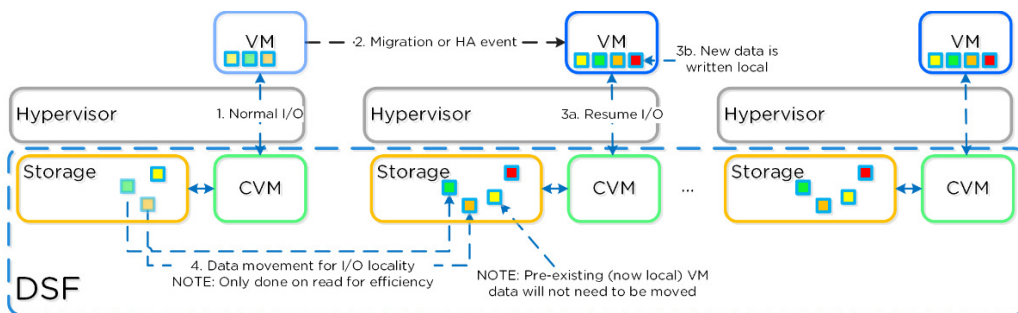
将远程数据拉入本地 Stargate 服务的统一缓存中。这是以 4K 粒度完成的。

对于没有本地副本的情况，请求将被转发到包含副本的 Stargate 服务，副本将返回数据，本地 Stargate 服务将在本地存储这些数据，然后返回 I/O。对该数据的所有后续请求都将从缓存中返回。

- 数据块位置

如果数据副本已在本地的，则无需移动。

在此方案中，实际的数据副本将在满足某些 I/O 阈值后重新本地化。系统不会在不需要数据的时候重新本地化/迁移数据副本，以确保有效地利用网络。



15. 超融合的自动化安装部署工具

联想 ThinkAgile HX 超融合系统提供自动化安装部署工具 Foundation，Foundation 可以自动发现网络内的出厂节点，按照用户选择自动在每个节点上部署超融合软件和虚拟化软件，

实现超融合和虚拟化的融合部署。用户只需要为每个节点输入网络信息即可完成超融合和虚拟化的安装

New Cluster Setup
Set up general information to create and connect your cluster to the network.

Network information
This is some basic information about your Hypervisor, CVM, IPMI IPs.

CVM	Hypervisor	IPMI (Optional)
NETMASK: 255.255.255.0	NETMASK: 255.255.255.0	NETMASK: 255.255.255.0
GATEWAY: 10.2.100.1	GATEWAY: 10.2.100.1	GATEWAY: 10.2.100.1
MEMORY: 32 GB	DNS SERVER IP: 10.1.1.00	

Post Imaging Tests
This enables a series of tests to ensure that the cluster has been correctly configured and everything is running smoothly.

ENABLE TESTING

< Prev Next >

Foundation 支持用户按需选择虚拟化平台进行统一安装，无需手工安装虚拟化。

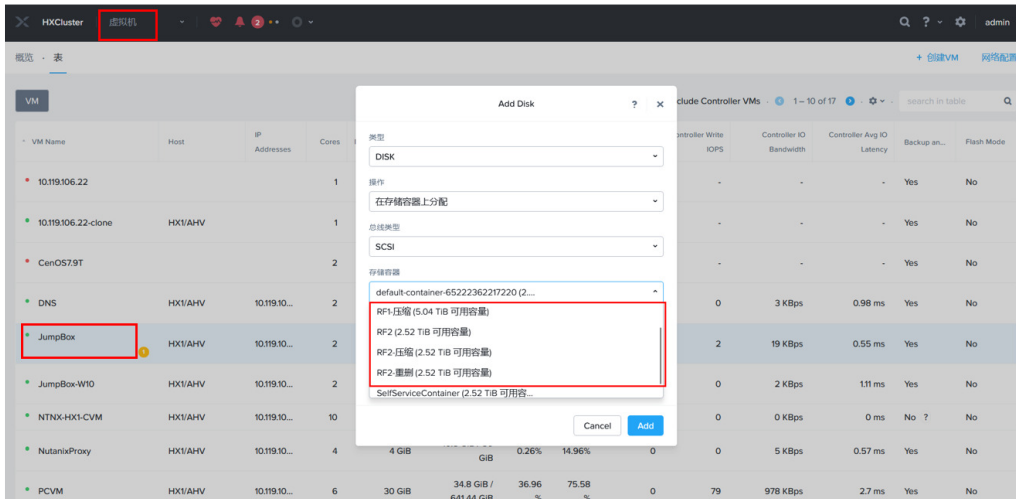
Image Uploads
You can select Acropolis and Hypervisor ISO images for your nodes.

Acropolis	Hypervisor	SKU
Installed Version 4.7 Uploaded nutanix_installer_package- release-danube-4.7.0.1- stable.tar.gz Upload Tarball	AHV ESX HYPER-V CPS Available hypervisors <input type="radio"/> host-bundle-el6.nuta nix.20150921.tar.gz <input type="radio"/> host-bundle-el6.nuta nix.20160217.2.tar.gz <input checked="" type="radio"/> kvm_host_bundle_2 0160601.6.tar.gz Upload RPM Content	Choose Hyper-V SKU <input type="radio"/> Free <input type="radio"/> Standard <input checked="" type="radio"/> Datacenter <input type="radio"/> Standard with GUI <input type="radio"/> Datacenter with GUI

< Prev Skip Create >

16. 存储空间优化技术

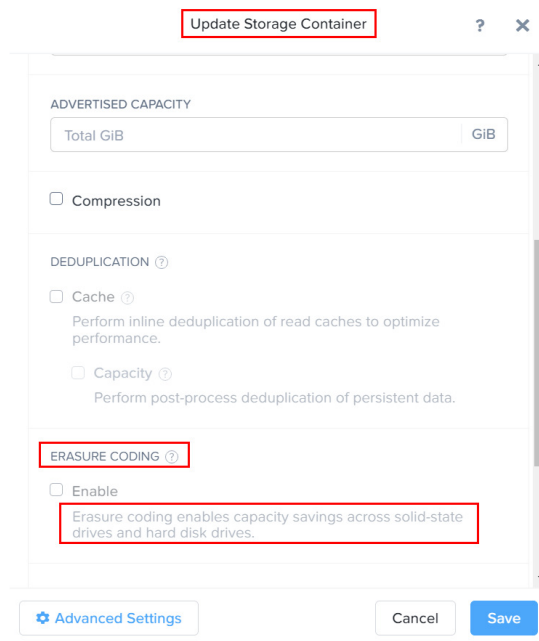
联想 ThinkAgile HX 超融合系统可以提供重删、压缩、纠删码技术优化存储空间使用。ThinkAgile HX 超融合系统通过软件方式（不使用专用硬件），提供内存、SSD、SATA 等不同层面的重复数据删除；通过软件方式（不使用专用硬件），提供数据压缩、纠删码等功能，且可根据不同的业务应用特点，针对不同的存储器或卷组，独立开启以上功能，提升存储的效率。



重删、压缩和纠删码可以针对存储器独立开启或者关闭，针对不同的业务特点优化使用。

名称	复制因子	Node	压缩	缓存重删 数据删除	容量重删 数据删除	纠删码	可用容量 (逻辑)	已用	Reserved Capacity	最大容量	控制器 IOPS	控制器 I/O BW	控制器 I/O 延迟
flow_data	1	-	关闭	Off	Off	Off	36775 GiB	0 GiB	0 GiB	36775 GiB	-	-	-
kafka_data	1	-	打开 (60分钟)	Off	Off	Off	36775 GiB	25.55 GiB	0 GiB	393.31 GiB	0	6 KBps	7.04 ms

为了在可用性与减少所需存储量之间取得平衡，联想 ThinkAgile HX 超融合系统提供了使用纠删码（EC）对数据进行编码的功能。EC 对不同节点上的数据块条进行编码并计算奇偶校验。在主机和/或磁盘发生故障时，可以利用奇偶校验来计算任何丢失的数据块（解码）。



17. 存储快照、克隆功能

联想 ThinkAgile HX 超融合系统在存储层面提供基于虚拟机的快照和克隆，并且具备自动快照管理功能，可以通过图形化界面定制时间间隔自动拍摄快照，并且自定义保存最新的多个快照(实现核心业务系统的逻辑备份)。

AOS 为卸载的快照和克隆提供本机支持，可以通过 VAAI, ODX, ncli, REST, Prism 等利用。快照和克隆都利用了最有效和高效的写入时重定向算法。如上面的数据结构部分所述，虚拟机由文件 (vmdk/vhdx) 组成，这些文件是超融合平台上的虚拟磁盘。

虚拟磁盘由盘区组成，盘区是逻辑上连续的数据块，这些数据块存储在盘区组中，这些盘区组是作为文件存储在存储设备上的物理连续数据。创建快照或克隆时，基础虚拟磁盘将标记为不可变，并创建另一个虚拟磁盘作为读/写。此时，两个虚拟磁盘具有相同的块映射，即虚拟磁盘到其相应盘区的元数据映射。与需要遍历快照链(可能会增加读取延迟)的传统方法相反，每个虚拟磁盘都有自己的块映射。这消除了大型快照链深度通常看到的任何开销，在拍摄连续快照或者克隆时不会对性能产生任何影响。

The image shows two overlapping configuration windows from a management interface. The top window is titled "Clone VM" and contains the following fields:

- General Configuration**
 - Number Of Clones: 1
 - Name: MSSQL16_SOURCE_VM
- Compute Details**
 - vCPU(s): 2
 - 每个vCPU的核数: 1
 - 内存: 4 GIB
- 启动配置

Buttons: 取消 (Cancel), Save

The bottom window is titled "Take Snapshot" and contains the following field:

- 名称: 数据库快照

Value must only contain alphanumeric characters, dots, dashes, or underscores.

Buttons: 取消 (Cancel), Submit

Database Snapshot Management Summary:

- Name: 数据库快照管理
- Protection Domain Mode: Active
- Next Snapshot Time: 07/25/23, 01:12:00 PM
- VM Count: 1
- Volume Group Count: 0
- File Count: 0
- Remote Site(s):

ID	Create Time	Reclaimable Space	Expiry Time	应用程序一致性	Actions
662313	07/25/23, 01:12:02 PM	Processing	07/25/23, 11:12:02 PM	错误	Details · Restore · X
662303	07/25/23, 01:10:47 PM	Processing	07/26/23, 01:10:47 PM	错误	Details · Restore · X
662291	07/25/23, 01:10:08 PM	Processing	Forever	错误	Details · Restore · X

18. 远程数据容灾

联想 ThinkAgile HX 超融合系统可以提供原生的跨地域的远程数据容灾能力。支持 1 对 1 远程异步数据复制功能，提供应用级别的一致性快照，保证存储间数据一致性。

Update Protection Domain: 数据库容灾保护

Entities | **Schedule**

Configure your local schedule

- Repeat every minute(s) ?
- Repeat every hour(s) ?
- Repeat every day(s) ?
- Repeat weekly
- S M T W T F S
- Repeat monthly
- Day of month: ?
- Start on at
- End on at

Create application consistent snapshots

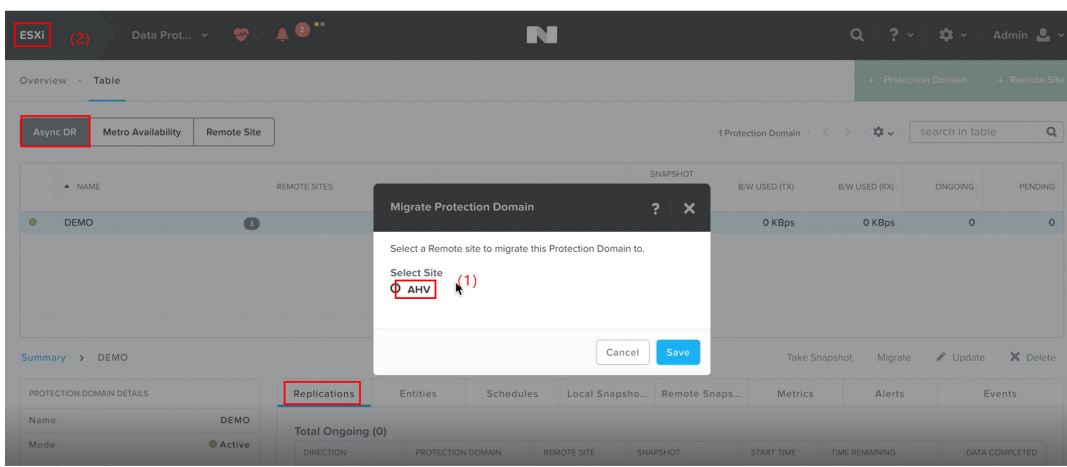
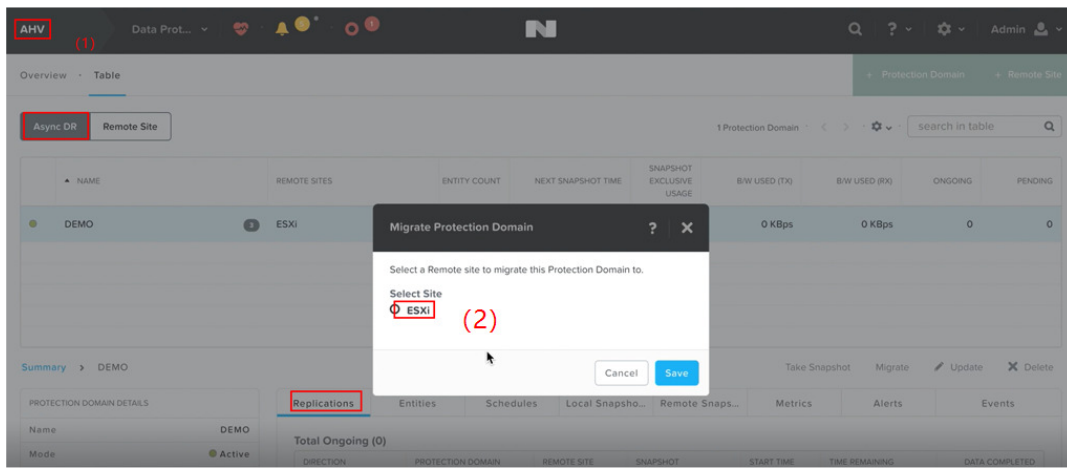
Retention policy

- Local keep the last snapshots
- Up to 10 快照 will be retained for this schedule and retention policy (10 快照 for each interval).
- Remote Sites**
- Shanghai keep the last snapshots
- Shenzhen keep the last snapshots
- Zhuhai keep the last snapshots

取消 **Create Schedule**

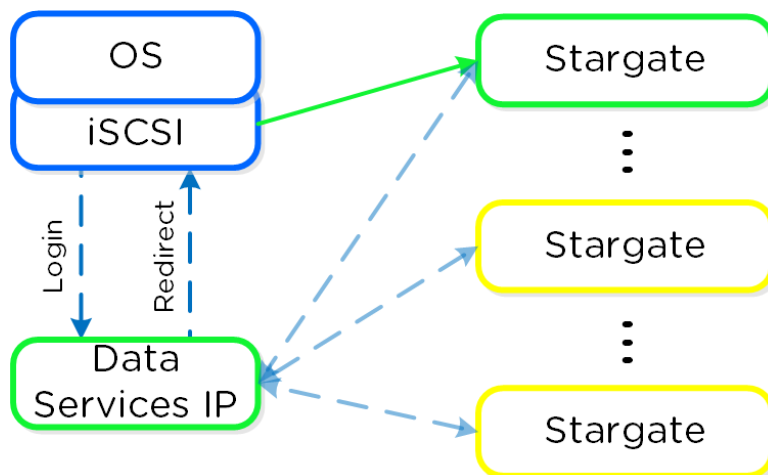
19. 虚拟机跨虚拟化平台容灾功能

联想 ThinkAgile HX 超融合系统支持跨虚拟化平台容灾，HX 可以将应用和数据从自有 AHV 虚拟化平台 (1) 容灾到 ESXi 虚拟化平台 (2) 或者从 ESXi 虚拟化平台 (2) 容灾到自有 AHV 虚拟化平台 (1)，提供方便快捷的容灾保护能力。

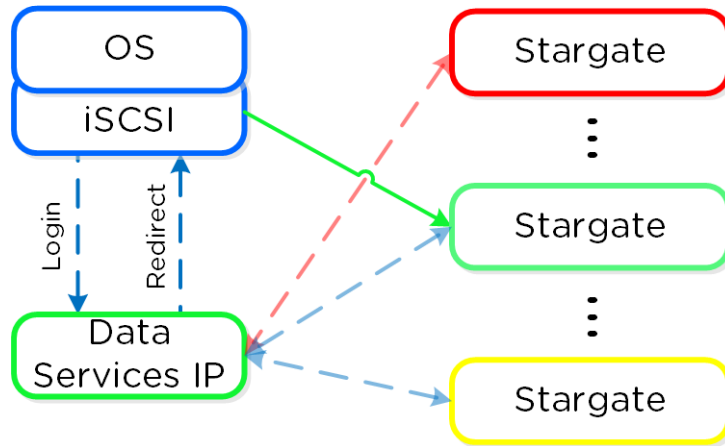


20. 分布式块存储服务

联想 ThinkAgile HX 超融合系统可以为内部虚拟机、外部物理主机和容器提供分布式块存储服务，且无需手动配置 OS 多路径软件即可自动实现多路径高可用和负载均衡。并可支持基于此分布式块存储服务构建 Oracle RAC，Windows Failover Cluster 等应用集群。分布式块存储服务使用数据服务 IP 进行 Target 发现。这允许使用单个地址，而无需知道单个 CVM IP 地址。数据服务 IP 将分配给当前的 iSCSI 领导者。如果失败，将选出新的 iSCSI 领导者并分配数据服务 IP。这可确保发现门户始终保持可用。iSCSI 发起程序配置为数据服务 IP 作为 iSCSI 目标门户。在收到登录请求后，平台将执行 iSCSI 登入重定向到健康的 Stargate。



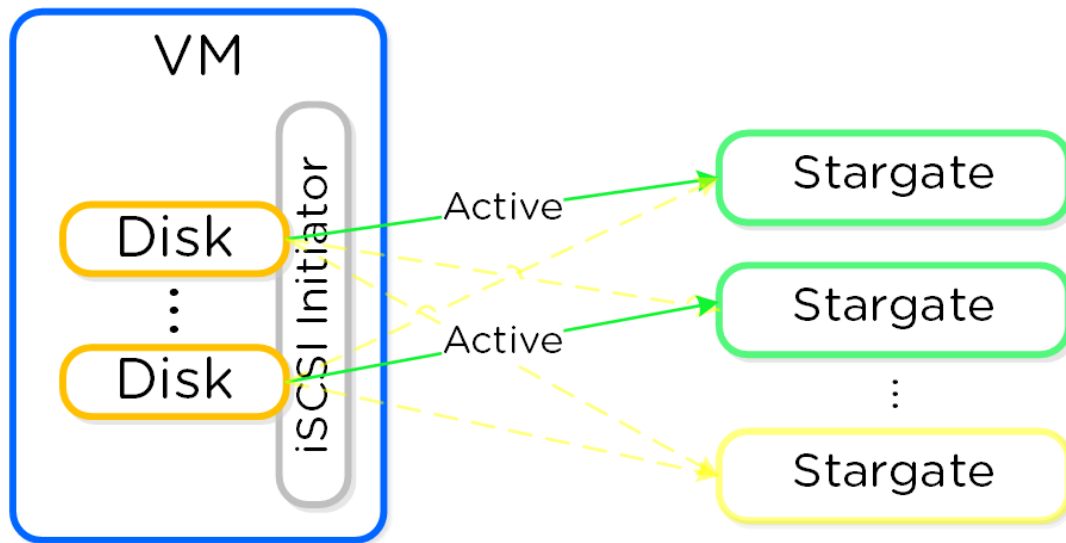
如果活动（关联）Stargate 关闭，发起方将重试对数据服务 IP 的 iSCSI 登录，然后重定向到另一个健康的 Stargate。



这种设计无需手动配置 OS 多路径软件即可自动实现多路径高可用。

iSCSI 协议规范要求每个目标在发起方和目标之间有一个 iSCSI 会话（TCP 连接）。这意味着 Stargate 和目标之间存在 1:1 的关系。

块存储服务将为每个连接的 Initiator 自动创建 32 个（默认）虚拟目标，并将其分配给添加到卷组（VG）的每个磁盘设备。这将为每个磁盘设备提供一个 iSCSI 目标。以前，这将通过创建多个 VG 来处理，每个 VG 都有一个磁盘。这允许每个磁盘设备拥有自己的 iSCSI 会话，并且能够跨多个 Stargate 托管这些会话，从而提高可扩展性和性能。在 iSCSI 会话建立（iSCSI 登录）期间，对每个目标进行负载平衡。



21. 简便强大的管理能力

联想 ThinkAgile HX 超融合系统支持端对端的消费级管理解决方案 Prism，在虚拟化数据中心环境下，将多种管理与报告工作融为一体，实现前所未有的简化管理。Prism 采用全新设计，打造流畅而丰富的使用体验，提供直观的用户界面，旨在简化及优化常见数据中心 workflow，使用户无需再为不同的任务而采取不同的管理解决方案。Prism 提供软硬件监控、风险分析、健康检查、错误告警、备份容灾等，而不是多种监控管理、容灾等软件的组合。

软硬件监控



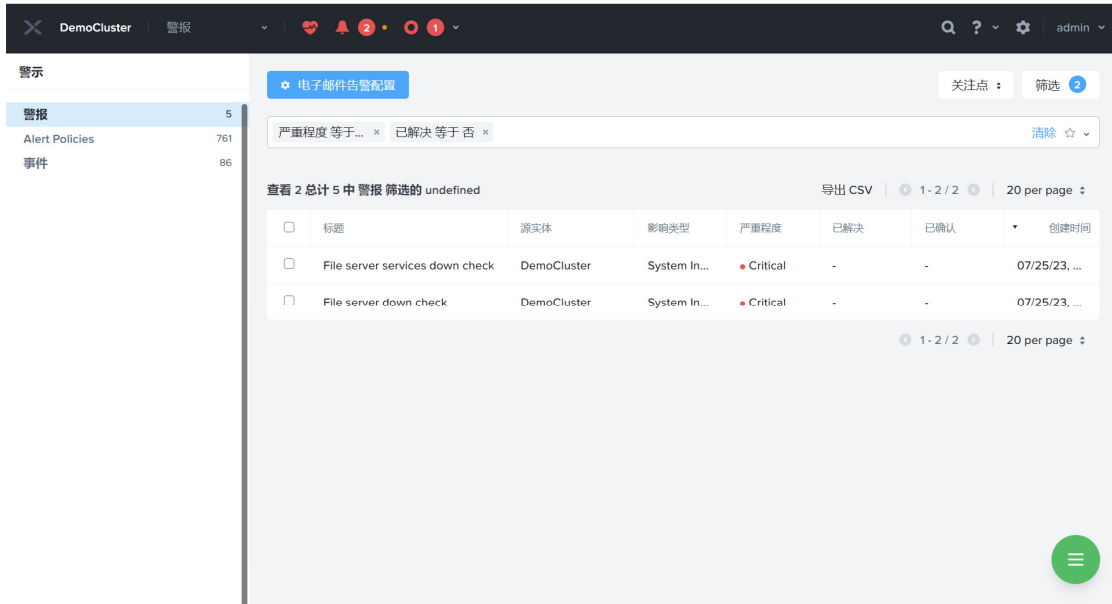
风险分析



健康检查



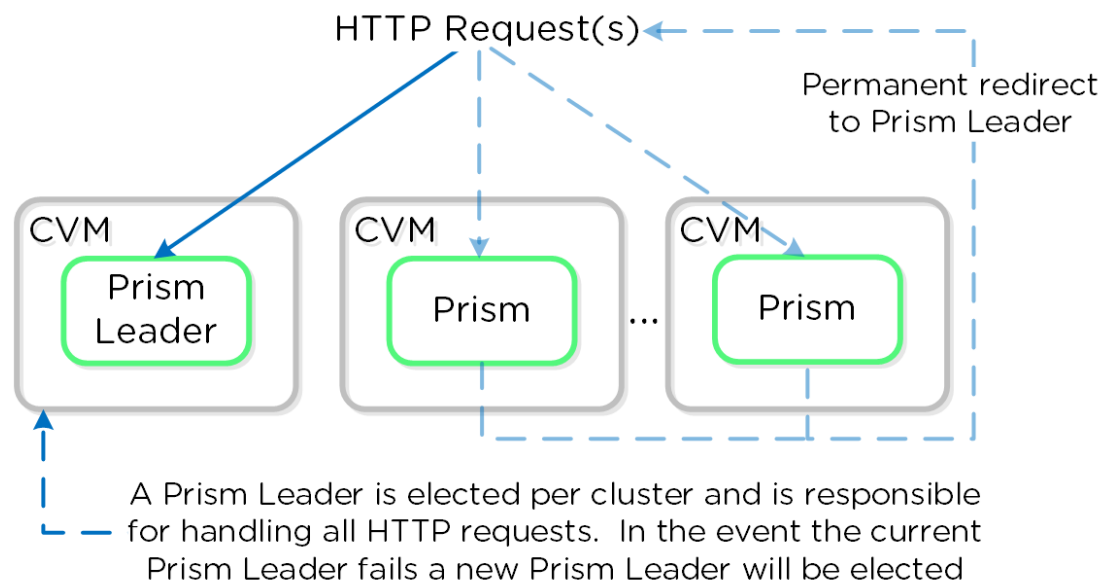
错误告警



备份容灾

名称	Remote Sites	Entity Count	Next Snapshot Time	Snapshot Exclusive Usage	B/W Used (Tx)	B/W Used (Rx)	Ongoing	Pending
era-server-pd-223787e6	-	1	-	9.69 GiB	0 KBps	0 KBps	0	0
ERA_Demo-Fiesta-Source_93ae9910	-	0	-	3.34 MiB	0 KBps	0 KBps	0	0
ERA_DEMO_MSSQL_ERA_CREATE_D_PATCHED_101fb817	-	0	-	10.55 GiB	0 KBps	0 KBps	0	0
ERA_DRIVE_023c8569-5216-438f-921d-0b75b8bb1d53_PD	-	1	-	164.97 MiB	0 KBps	0 KBps	0	0
ERA_DRIVE_ca39657d-f997-439f-b003-bc442cee4bf1_PD	-	1	-	152.34 MiB	0 KBps	0 KBps	0	0
ERA_ERA_DEMO_MSSQL_SOFTW ARE_PROFILE_2043c2c3_ba50ce8c	-	0	-	-	0 KBps	0 KBps	0	0

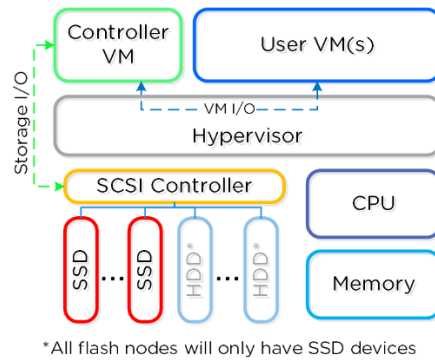
Prism 统一管理平台也是分布式架构，而不是主备（高可用）架构，访问统一服务域名（IP 地址）或任意硬件节点服务域名（IP 地址）都可以管理整个集群；任意管理节点及组件出问题，都可通过访问其他健康节点实现对整个平台的管理。



22. 一键式在线升级服务

联想 ThinkAgile HX 超融合系统提供 LCM (Life Cycle Manager) 跟踪和升级集群中所有软件和部件的源码。

由于在用户空间中作为 VM 运行的性质，超融合平台软件升级或 CVM “故障”之类的事情不会对用户虚拟机和工作负载产生影响，因为它们在虚拟机管理程序之外。例如，如果出现某个 CVM 出现故障的灾难性问题，整个节点仍会继续运行，存储 I/O 和来自集群中其他 CVM 的服务。在 AOS (Nutanix 核心软件) 升级期间，我们可以重新启动 CVM，而不会对该主机上运行的工作负载产生任何影响。HX 不需要关闭或迁移业务虚拟机和主机节点，即可实现超融合平台软件的在线升级。



LCM 可以升级 AOS 超融合平台软件，虚拟化软件，BIOS、磁盘控制卡、磁盘微码以及其他相关功能模块，不会带来业务中断。

The screenshot shows the LCM 'Inventory' view. It displays installed software and firmware versions for a cluster and three hosts. The interface includes a sidebar with navigation options like 'Best Practices', 'Inventory', 'Updates', 'Software', and 'Firmware'. The main content area shows two tables: 'Installed versions on 1 cluster' and 'Installed versions on 3 hosts'.

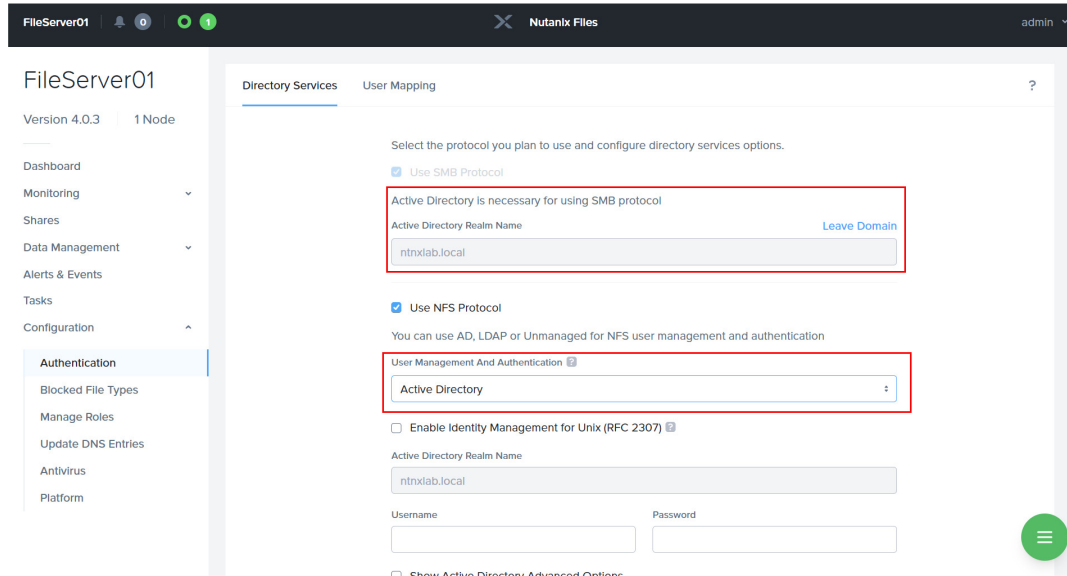
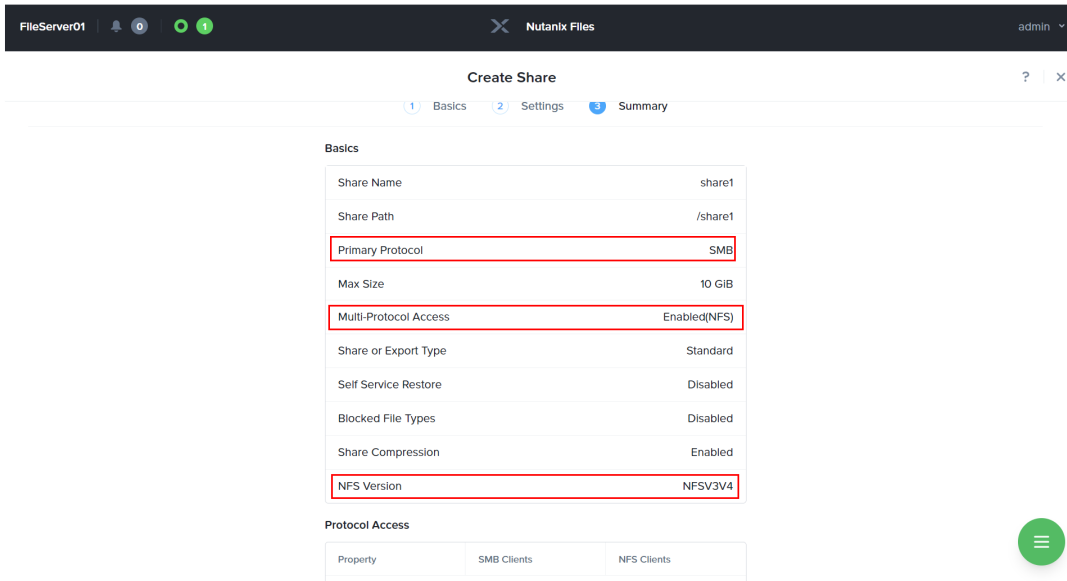
Cluster	AOS	Cluster Maintenance Utilities	ESX hypervisor	FSM	Foundation
rp-esx	5.18.1.2	1.0.0	Various on 3 entities	1.3.1	4.5.4.2

Host	BIOS	BMCs	HBAs	BIOS (Redfish)	BMCs (Redfish)
rp-esx-1			MPTFW-14.00.00.00-IT	PB42.602	710.00
rp-esx-2	PB42.602	9.09	MPTFW-14.00.00.00-IT	System does not meet Redfish requirements	System does not meet Redfish requirements
rp-esx-3	PB42.300	9.09	MPTFW-10.00.07.00-IT	System does not meet Redfish requirements	System does not meet Redfish requirements

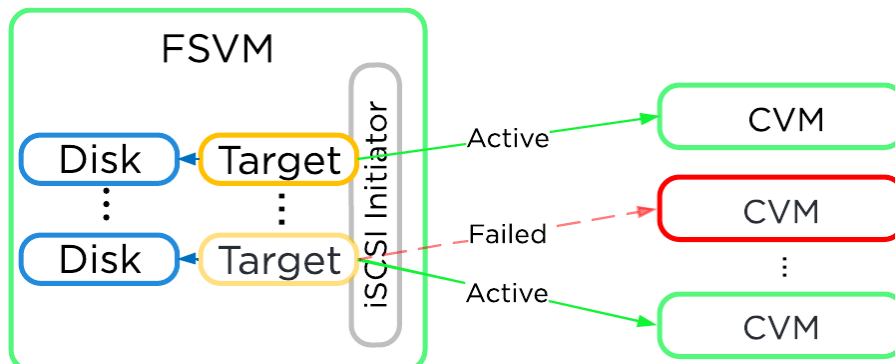
23. NAS 文件服务

联想 ThinkAgile HX 超融合系统可以同时提供 NAS 文件服务 Files，Files 是采用完全分布式架构，软件定义可横向扩展文件存储解决方案，支持按需自动扩展文件服务器，可解决包括用户配置文件、主目录和部门共享等各种用例。联想 ThinkAgile HX 超融合系统将服务器虚拟化的灵活性与企业云平台的分布式存储相结合，提供融合 NAS 和公有云服务最佳属性的文件服务，同时克服了两种方案的局限性。

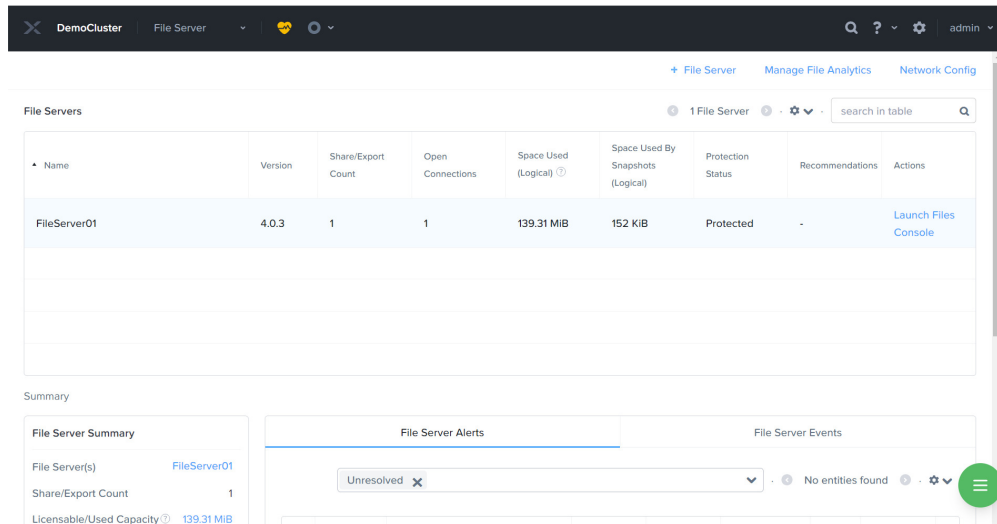
联想 ThinkAgile HX 超融合系统 NAS 文件服务 Files 提供 SMB 2/SMB 3/NFS v4/NFS v3 协议支持。文件服务特性与微软活动目录（AD）和 DNS 完整集成，这样就可以利用 AD 中所有安全的和已有的认证和授权功能。所有的共享权限、用户和组管理也可以在传统的负责 Windows 文件管理的 MMC 中完成。



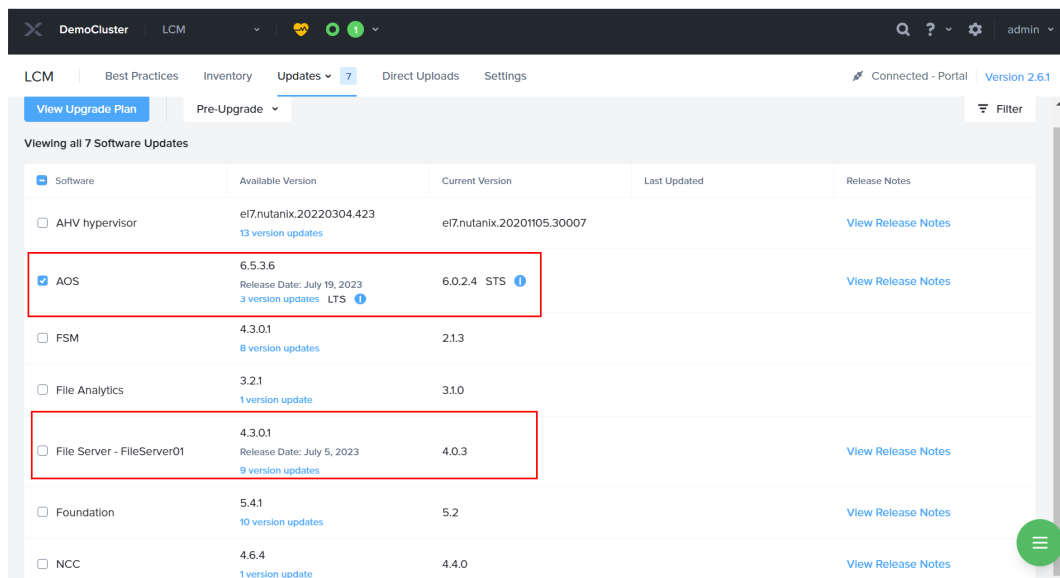
Files 遵循与 HX 超融合平台相同的分布式架构，以确保可用性和可扩展性。作为文件服务器部署的一部分，将至少部署 3 个 FSVM，支持存储路径冗余，无需配置支持文件服务器高可用。每个 FSVM 都利用卷 API 进行数据存储，该数据存储可通过客户机内 iSCSI 访问。这允许任何 FSVM 在发生 FSVM 中断时连接到任何 iSCSI 目标。如果 CVM 不可用（例如，活动路径关闭），iSCSI 重定向用于将目标故障转移到另一个 CVM，然后由该 CVM 接管 IO。



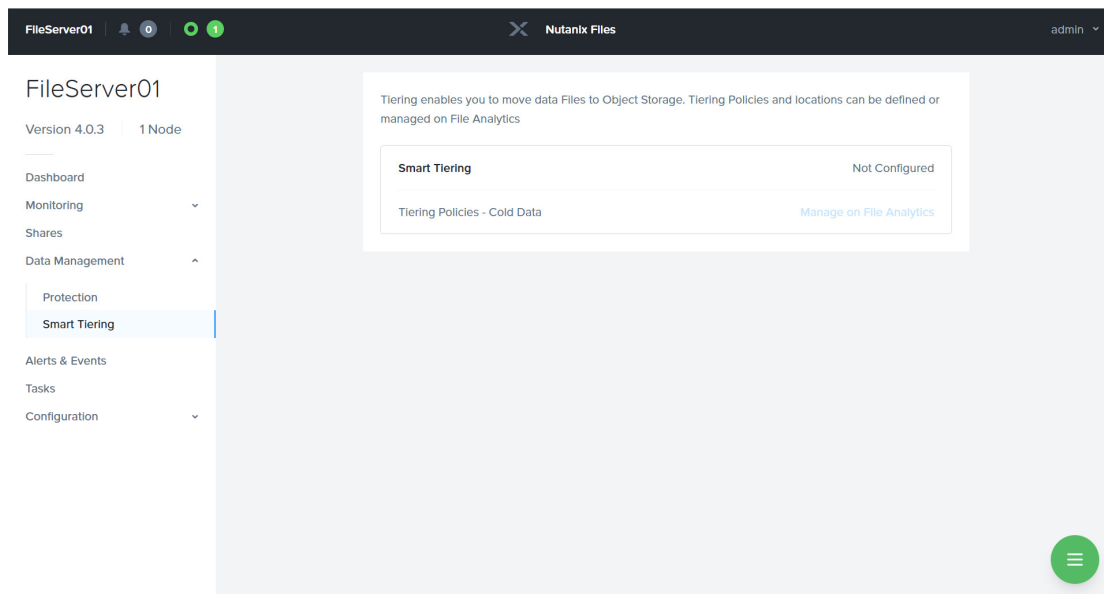
Files 在超融合管理界面中统一管理，无需安装其他管理工具。



Files 与 AOS 超融合平台独立运行，支持独立于分布式存储之外在线升级，分布式存储软件升级不影响 Files 的使用。



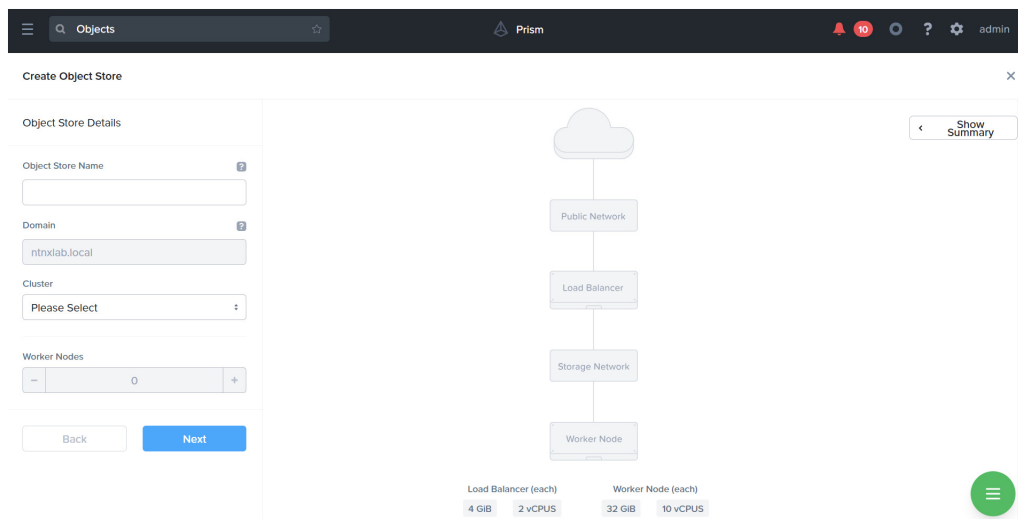
Files 支持文件数据通过 Smart Tiering 特性分层到其他对象存储，可以通过将数据分层到对象存储来释放文件服务器上的空间。执行分层有两种方法：通过文件控制台使用标准分层管理和定义策略，或者通过 Data Lens 使用高级分层。



24. 对象存储服务

联想 ThinkAgile HX 超融合系统可以同时提供对象存储服务 Objects，对象功能通过符合 S3 的 API 提供高度可扩展且持久的对象服务。鉴于对象部署在 HX 超融合平台之上，它可以利用 AOS 功能，如压缩、纠删码、复制等。从上面的描述可知，HX 超融合系统支持在同一平台上运行对象存储服务以及虚拟机、文件存储和块存储服务。当然，HX 超融合系统也支持单独部署专用的对象存储服务集群。

联想 ThinkAgile HX 超融合系统可以在几分钟内按需一键式启用对象存储服务，所有操作均为向导式界面，方便快捷。



对象存储服务 Objects 支持无中断的横向扩展模块化设计，且任何组件均可独立扩展以匹配工作负载增长需求，可从 TB 级扩展到数 PB 级。此服务采用与 AWS S3 兼容的 REST API 接口设计，该接口能够处理数 PB 的非结构化和机器生成的数据。对象通过使用标准 S3 API 解决与存储相关的使用案例，用于云原生应用程序的备份、长期保留和数据存储。您不再需要引入单独管理的外部存储解决方案。对象存储服务 Objects 提供 S3 兼容的 REST API 接口，支持通过网络 HTTP 命令“PUT”和“GET”对象，支持在程序或脚本中集成 REST API 调用，无需跟踪复杂的目录结构

Name	Version	Domain	Nodes	Usage (Logical)	Bucket s	Object s	Alerts	Notifications	Objects Public IPs
ostore01	3.4.0.2	ntnrlab.local	1	53.35 MB	2	72	3	Disabled	192.168.1.15

25. 数据库生命周期管理服务

联想 ThinkAgile HX 超融合系统可以提供数据库生命周期管理服务，自动化并简化数据库管理，为数据库配置和生命周期管理带来一键式简单性和隐形运维。

数据库生命周期管理服务使数据库管理员能够执行诸如数据库注册，配置，克隆，修补和还原之类的操作。它允许管理员使用最终状态驱动的功能（包括网络分段、高可用性（HA）数据库部署等）定义配置标准。使用数据库生命周期管理服务多集群，可以使用 HX 超融合集群轻松管理多个位置的数据库，包括本地和云中的数据库。数据库生命周期管理服务支持管理多种类型数据库，比如 Oracle、SQL Server，PostgreSQL、MySQL 等等。

Time Machine	Type	Status	Size (GiB)	Nutanix Cluster
Demo-Fiesta...	SQL Server Database	●	0	DemoCluster
FiestaDB_TM	SQL Server Database	●	204	DemoCluster

数据库生命周期管理服务支持一键式置备数据库，包括自动创建虚拟机，自动安装和配置数据库等信息。

Provision a MySQL Instance

? | X

- 1 Database Server VM
- 2 Instance
- 3 Time Machine

Database Server VM

- Create New Server
- Use Registered Server

New Database Server VM

Database Server VM Name ⁱ

Description

Nutanix Cluster

Software Profile

Compute Profile

Network Profile

Version: MYSQL_... [Update Version](#)

Era Drive User

Password ⁱ

Cancel Next

Provision a MySQL Instance

? | X

- 1 Database Server VM
- 2 Instance
- 3 Time Machine

MySQL Instance Name

Description

Listener Port

Size (GiB)

ROOT password

Database Parameter Profile

Name of Initial Database ⁱ

Pre-Post Commands

Back Next

1 Database Server VM 2 Instance 3 Time Machine

Name: Description:

SLA: ⓘ

Schedule

Daily Snapshot at: ⓘ Snapshots Per Day: - + Log Catch Up Every:

Weekly Snapshot on: ⓘ Monthly Snapshot on the: ⓘ Quarterly Snapshot in: ⓘ

数据库生命周期管理服务支持统一的数据库补丁管理,使得企业数据库补丁统一,简化管理。

Software Profile

Versions

Name	Version	Owner	Status	In Use	Mutually Exclusive Clusters	Optimized	Modified On	Description
<input checked="" type="radio"/> CU5	3.0	Hary	Published	0	9	No	2020-06-30 11:57:28	SQLServer 2019 CU5
<input type="radio"/> CU2	2.0	Hary	Published	0	9	No	2020-03-25 14:25:16	SQLServer 2019 CU2
<input type="radio"/> RTM	1.0	Hary	Published	0	9	No	2019-06-12 13:42:37	SQLServer 2019 RTM

CU5 Details

Patch	Operating System	Database Software	Database Server VMs	Availability
SQLServer2019CU5-KB4552255-x64-ENU.exe				

Patch: SQLServer2019CU5-KB4552255-x64-ENU.exe
Version: 15.0.4043.16
KB Number: 4552255
More Info Url: <http://support.microsoft.com/?id=4552255>

Notes:

How to Update Databases with Software Profile Versions

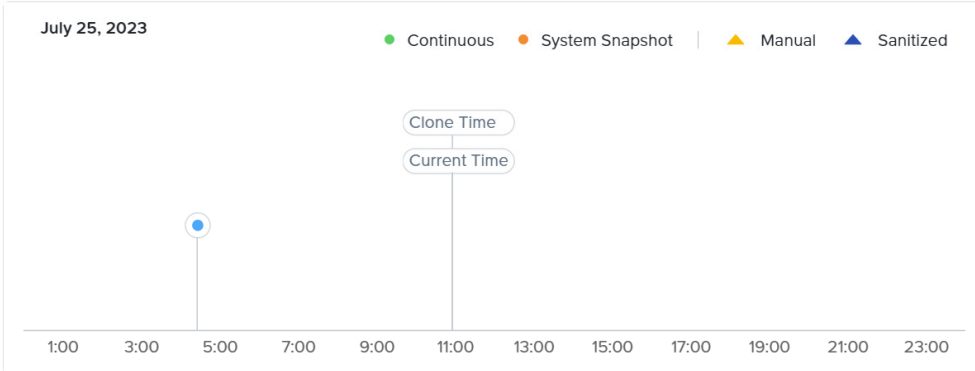
- Identify the patch file to be applied to the databases.
- Place the patch file on a shared file server that is accessible from all the Database Server VMs where the patch will be applied.
- Create a new version of the software profile. Provide the path to the patch file.
- Open the Database Server VM Homepages where the software profile is used and apply the patch.
- Publish the Software Profile Version to make it available to all users.

数据库生命周期管理服务支持针对源数据库进行一键克隆、创建计划任务刷新数据库,供给开发、测试、数据分析、批处理作业等其他场景使用。

- 1 Time/Snapshot
- 2 Database Server VM
- 3 Database

Select a point in time or a snapshot

Nutanix Cluster DemoCluster



Clone a

Point in Time

Snapshot

12:11:49

era_auto_snapshot (2023-07-25)

Cancel

Next

- 1 Time/Snapshot
- 2 Database Server VM
- 3 Database

Database Server VM

Create New Server

Use Authorized Server

Database Server VM Name

MSSQL16_SOURCE_VM_2023_Jul_2

Description

Compute Profile

DEFAULT_OOB_COMPUTE

Network Profile

DEFAULT_OOB_SQLSERVER_N

Windows License Key

Optional

Administrator Password

Join Domain

Back

Next

- 1 Time/Snapshot
- 2 Database Server VM
- 3 Database

Clone Name: Fiesta_2023_Jul_25_12_11_49

Description:

Database Name on VM: Fiesta

Instance Name: MSSQLSERVER

Schedule Data Refresh

Refresh Every (days):

Refresh Time (America/New_York):

Use Mount Point

Removal Schedule

Pre-Post Commands ▼

API Equivalent

Back

Clone