

DEWCloud x-Archive 云归档解决方案

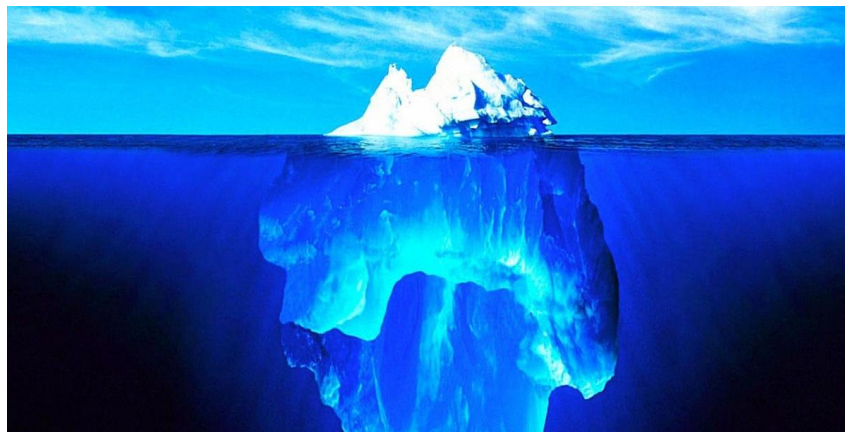
摘要

根据数据存储的“二八定律”，几乎不被访问的“冷数据”，占据了企业存储系统 80% 的空间。在新数据时代，传统的归档方式面对海量冷数据显得力不从心，DEWCloud 的 x-Archive 提供了更加经济高效的云归档解决方案，满足海量冷数据的廉价存储、云端迁移以及离线归档的需要。

背景概述（沉睡的数据巨人）

在新技术的持续催动下，全球数据继续保持高速增长。据预测，到 2025 年中国的数据量将增至 48.6ZB，企业数据中心将承受巨大的数据存储管理压力。响应滞后的 IT 预算让企业的存储资源变得十分紧张，极端情况下，不得不删除很多较少用到的数据，给企业未来发展带来不可估量的损失。

数据中心存储的各类数据中，一部分会被业务频繁调用访问，这类数据称之为“热数据”；部分数据的使用频率较低，这类数据一般被称为是“温数据”；有些数据几乎不被系统访问，处于常年沉睡状态，这类数据则被称为“冷数据”。据调研机构统计，企业数据中心里只有 10%~15% 的数据会被经常访问的，而其他全部是“温冷数据”。“温冷数据”使用的频率非常低，如互联网企业几年前用户上传的照片资料，其他行业如医院的病历档案与医疗图片、档案馆的文史资料等。这些数据虽很少被访问，但因业务需要、法规遵循等相关要求，需要被长久保存以便随时回溯使用。大数据时代的冷数据有着庞大的体量，如果仍采用传统存储设备来存储会带来高昂的数据存储成本，拉低业务访问性能。



x-Archive 云归档解决方案

目前解决“冷数据”存储问题较主流的做法是采用磁带归档，磁带归档可以大大降低数据中心的能耗和存储成本，但其实也加剧了整个 IT 系统的复杂度，且大幅的牺牲了数据的访问性能。

磁盘归档需要通过备份软件才能将数据转入磁带设备，传统备份软件需要历经的文件扫描打包、数据

写入、校验等多个步骤，归档效率非常低下；离线归档完成后，用户需要几个小时、几天甚至更长的恢复时间，才能访问到这些长久保持的冷数据。而在海量非结构数据应用场景下，以上动作所花费的时间对大部分用户而言是无法忍受的。用户希望能以更高效的归档迁移机制、更低的单位存储成本、更低的功耗，实现海量冷数据的可靠存储与快速访问。

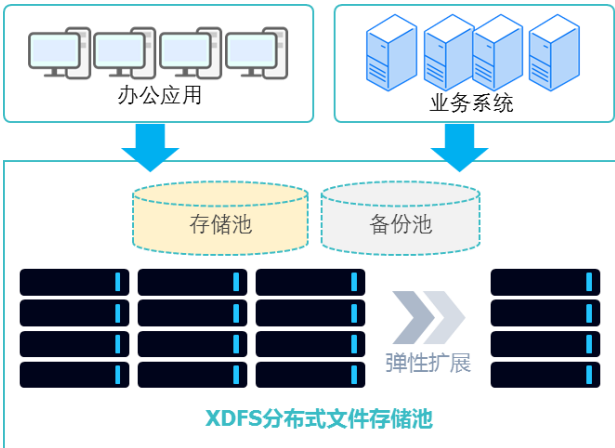
针对企业海量数据存储与归档需求，DEWCloud 采用 XDFS 分布式文件存储打造了面向海量数据应用的云归档解决方案。方案基于软件定义文件存储系统，结合蓝光/公有云归档二级存储架构，实现海量温冷数据的迁移管理，将冷数据始终保存在经济、大容量的存储空间，为热数据释放出更高性能的存储资源，并实现数据在两级存储设备之间的自动迁移，快速访问，从而达到最高的性价比。



1、海量数据存储、备份

XDFS 是 DEWCloud 自主研发的分布式文件存储系统，基于标准 X86 硬件（支持利旧），可构建高性价比的大规模文件存储池。本方案中 XDFS 既可以做为用户文件类数据的主存储系统，通过标准 NAS 或 POSIX 客户端实现高效的数据读写，也可以作为大数据类应用的备份空间，借助分布式的并发读写性能优势，实现快速的海量数据备份恢复。

XDFS 可从 2 节点起步，动态扩展至 1024 个节点，单文件系统可达 192PB 的海量存储规模，真正做到按需部署，按需扩容。XDFS 提供 Posix/NAS/S3/iSCSI 等存储接口，可满足大规模文件存储、高性价比海量数据备份、私有云/混合云等场景。通过副本和纠删码两种数据冗余保护，保证业务数据的安全可靠。



通过 XDFS 构建大容量存储池，用于存储用户边缘数据，或作为海量数据的在线备份存储池

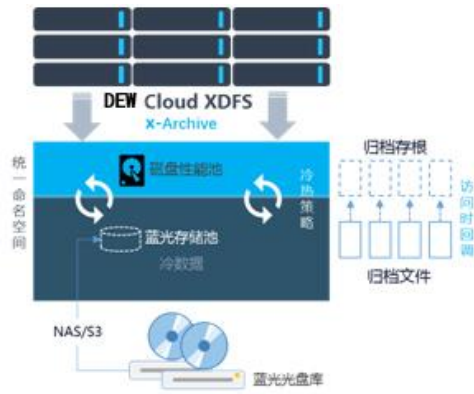
2、冷数据蓝光归档

因为存储容量大、使用寿命长等特性，很长一段时间内磁带都是备份归档的主力存储介质。随着蓝光存储技术的发展，磁带归档依赖于备份软件、对保存环境要求苛刻、长时间保存安全系数低、访问速度慢、核心技术完全掌握在国外厂商等，这些缺点逐渐暴露出来。相对磁带存储，蓝光光盘存储在普通环境下即可轻松达到 50 年的数据据保存时长，介质廉价轻便，可以很方便的存储管理和读写访问。相对于仍控制在国外厂商的磁盘和磁带技术，蓝光存储技术国内厂商已经完全掌握，在推行全国产化的信息安全背景下，蓝光归档拥有巨大的市场前景。

	磁盘/磁盘阵列	磁带/磁带库	光盘/光盘库
数据安全性	低（可修改或删除）	低（可修改或删除）	高（不可修改删除）
可靠性	硬盘损坏率极高，12 块硬盘同时工作 5 年存活率仅 1%	易胶磁、受潮、接触性读取，读写上百次后有断带风险，划伤无法修复	耐高低温、不怕潮、抗摔、抗振、抗压、防尘、非接触式读写无磨损
信息保存期	短（5-10 年）	短（10 年内）	长（大于 50 年）
总体拥有成本	高	中高	低
数据访问速度	快（随机读取）	慢（顺序访问）	中（随机读取）
强磁影响	大	大	无

在 XDFS 作为数据主存储的应用场景，XDFS 可结合蓝光归档，构建多层存储架构，以实现更低的 TCO，更高的数据可靠性以及更好的业务性能。

XDFS 首先通过 S3 或 NAS 接口与蓝光存储对接，将蓝光存储资源作为其冷数据存储资源池。内置的 X-Archive 归档迁移功能可智能检测数据的冷热属性，将访问次数少、存储占比最大的的冷数据迁移归档到蓝光存储设备进行长久保存。被迁移的数据会在磁盘存储池保留“存根”，一旦被归档数据的“存根”被访问触发，后台会自动将相应数据从蓝光介质中回调，实现归档数据的快速访问。在 XDFS 统一元数据管理下，蓝光资源作为 XDFS 存储资源池的一部分，以统一的命名空间的方式对外提供存储服务。目前 DEWCloud x-Archive 方案已支持紫晶、华录、松下、索尼等厂商主流蓝光库，提供多种蓝光归档选择。

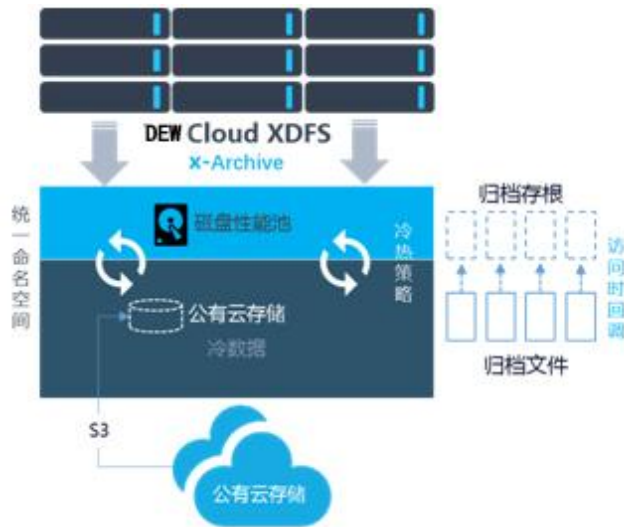


使用 XDFS 本连接蓝光设备，XDFS 磁盘存储池与蓝光存储池创建成统一命名空间的分层存储池，冷数据根据智能策略自动归档到蓝光介质

3、公有云归档、混合云网关

据相关统计，采用混合云进行分级存储可降低企业 80%的存储成本。对大部分企业而言，现有数据中心的IT资源是有限的，客户希望在不增加硬件设备采购的基础上，可以在本地数据中心的基础之上打通公有云资源，将海量温冷数据迁移归档到公有云空间，充分地利用云存储的便捷性与成本优势。

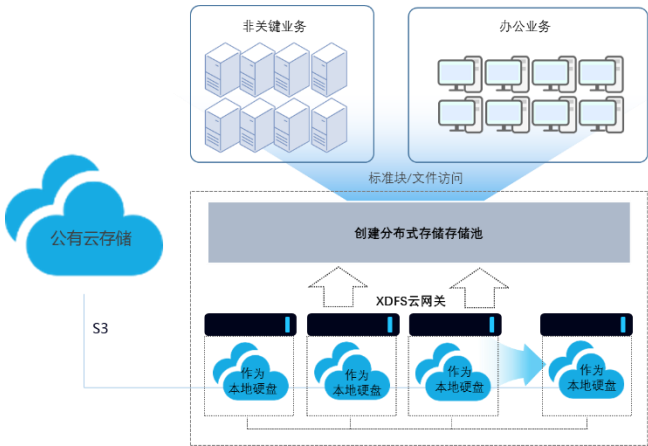
与 XDFS 的蓝光归档机制类似,XDFS 支持将公有云存储资源通过 S3 接入本地,并将其空间作为云端归档池，与 XDFS 本地存储池构建冷热分层存储，对外提供统一命名空间。用户的非敏感数据都可以存储到该分层存储池，在 x-Archive 的策略管理下，热数据保存在本地实现高效访问，冷数据自动流动到云端，如需被访问，数据自动从云端在线回调，实现私有云与公有云的无缝对接，构建经济立体的数据存储体系。



使用 XDFS 挂接公有云存储空间，XDFS 磁盘存储池与云存储池创建成统一命名空间的分层存储池，冷数据根据智能策略自动归档到云端

大部分企业数据数据中心都部署有本地存储系统和备份归档类软件，在海量冷数据不断产生的背景下，用户既希望利用混合云来实现数据分层管理，也希望充分利用现有的存储系统与备份归档工具，节约成本，避免过度建设。但传统应用和软件工具大多不支持云存储协议，无法直接使用公有云存储，因而需要将云资源转化为“本地存储”，才能打通混合云，被传统应用识别并无缝使用。

在这一应用场景下，XDFS 提供的混合云存储网关功能，可直接将公有云对象存储资源作为本地存储介质，使用这些“本地介质”创建分布式存储资源池，继而对外提供标准的 NAS、iSCSI 等存储访问，从而实现传统应用与云端的无缝对接。新创建的存储资源池可以作为业务数据存储空间，也可以配合用户的数据备份、迁移软件，将数据转移到公有云空间。



XDFS 作为云网关连接公有云存储，并将云存储作为本地硬盘创建本地存储池，通过标准 POSIX、NAS、iSCSI 等协议对外提供存储服务，传统应用可以透明使用云存储资源，实现数据存储归档等

方案优势价值



未来很长一段时间，数据高速增长都是各大数据中心的共同趋势，冷数据的积累和沉睡占据了用户越来越多的存储空间、电力能耗与经济投入，给企业带来巨大的管理压力。在大数据时代，这些数据都是企业的宝贵财富，里面埋藏了解决问题、优化业务和预测发展的数字密码。如何将这些海量规模的数据实现经济、可靠、高效的长久存放，是每个企业都需要面对的难题。

DEWCloud 的 x-Archive 云归档解决方案与时下兴起的“第二存储”理念不谋而合，在当前数据环境下，采用经济弹性的软件定义存储，无疑是企业海量非结构数据存储的首选。随着云存储的越来越成熟经济，打通公有云，构建从本地到云端的立体式存储体系，成为最具性价比的温冷数据治理方案。保证高效的本地数据访问的同时，也能满足海量边缘数据的长久存放，按需获取。