

海量小文件备份效率倍增攻略



目 录

PART 01 各行业小文件现状与风险

PART 02 灾备建设与实施难点

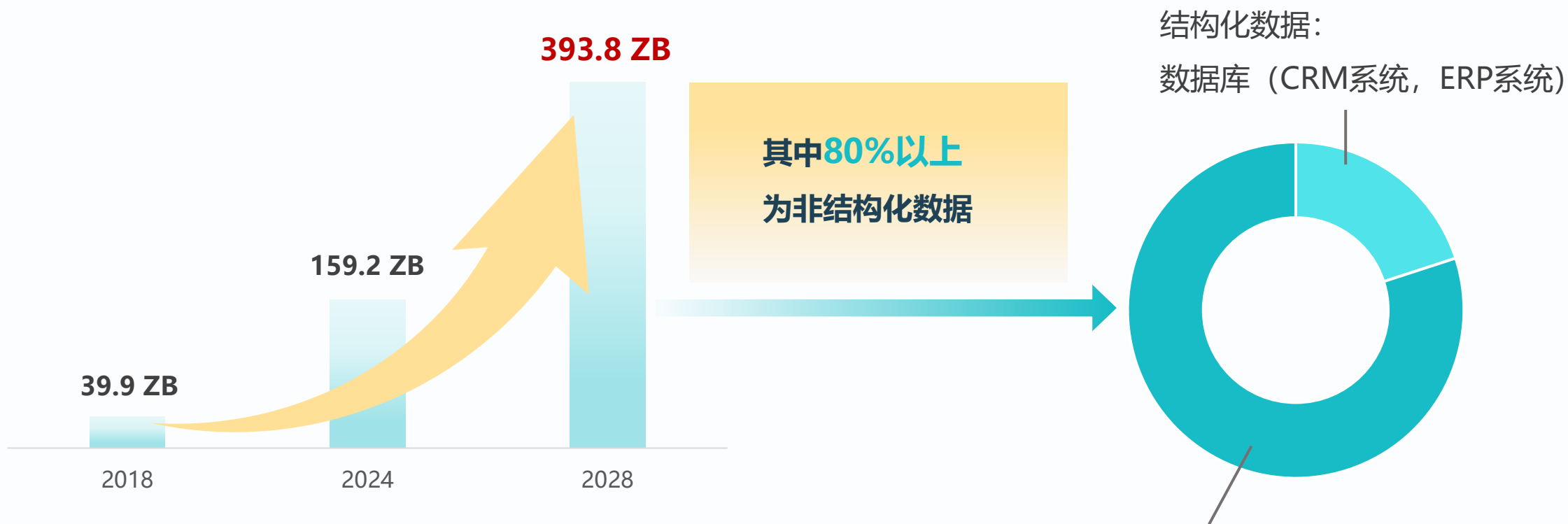
PART 03 云祺解决方案与应用

PART 01

各行业小文件现状与风险



快速增长的非结构化数据洪流



IDC 预测, 到2028年全球数据量将增长至**393.8ZB**, 相比于2018年增长**9.8**倍。
中国区数据量将达到接近**100ZB**。

非结构化数据分布情况

智慧医疗

PACS影像归档、电子病历存储、检验报告与医疗文书数据。

工程设计

CAD设计图纸、BIM模型文件、项目方案文档与渲染素材。

档案管理

数字化档案卷宗、历史文献扫描件、政府文书与凭证影像。

公共安全

执法记录仪视频、户籍电子档案、案件证据材料与监控录像。

智能制造

产线实时传感数据、设备工艺参数、质检图像与运维日志。

广电传媒

媒资影音素材、节目制作片段、直播流切片与素材库文件。

金融服务

票据扫描影像、客户身份资料、交易凭证与合同扫描件。

司法政务

电子卷宗管理、庭审录音录像、案件文书与证据档案。

自然资源

卫星遥感图像、地质勘探数据、测绘成果与地理信息文件。

智慧城市

IoT设备监控数据、城市安防视频、交通传感信息与环境监测。



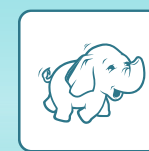
NAS



S3 对象存储



本地文件系统



Hadoop HDFS

非结构化数据存储分布

海量小文件特点

- **规模庞大：** 文件数量达千万至数十亿级
- **总量惊人：** 总数据量可达数百TB甚至PB级
- **层级复杂：** 目录嵌套深达数层甚至数十层
- **粒度极小：** 单文件大小多在几KB到几MB间

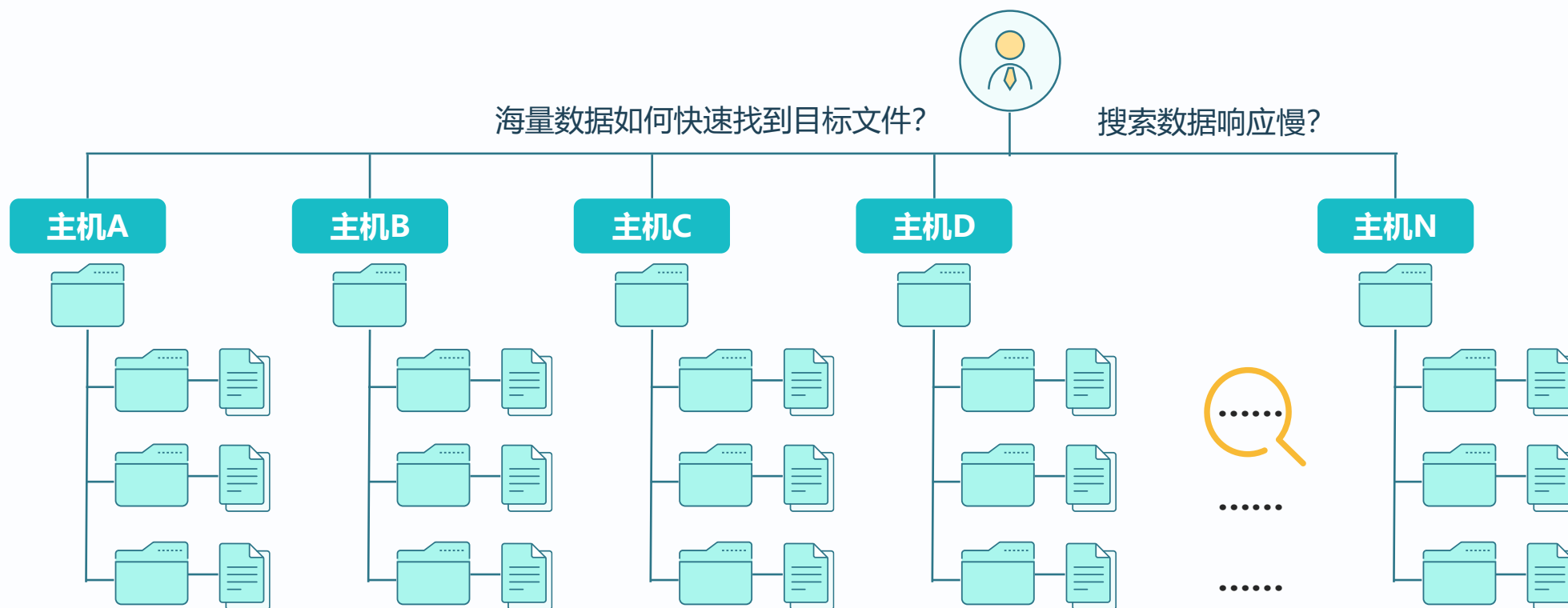
海量小文件为什么备份慢？

高频I/O开销，性能严重损耗

海量小文件需要频繁执行“打开-读取-关闭”操作，单次I/O的系统调用开销被极度放大。频繁的文件句柄操作导致扫描速度急剧下降，成为备份效率的核心瓶颈。

文件目录层级深且分布散乱

层级目录的深度及目录结构复杂性，会增加扫描的复杂性，导致扫描和传输效率低下。



目标文件匹配效率低

海量非结构数据中，因目录结构复杂和层级深等原因，定位目标文件时间久，整体效率低下。

恢复文件查找慢

备份海量数据之后，如需单独恢复某个小文件，对其查找等待时间久，极大降低恢复效率。

痛点根源：技术与管理的双重短板



传统架构性能瓶颈

传统NAS存储在面对亿级海量小文件时，读写性能急剧下降，难以支撑高频交易的低延迟需求。



权限管控与审计缺失

细粒度权限管理不足，操作日志记录不全，导致数据访问难以追溯，内部违规风险居高不下。



灾备体系响应迟缓

备份窗口过长，恢复速度慢，RTO与RPO难以满足金融级高可用标准，业务连续性面临威胁。



数据丢失与损坏

硬件故障、软件系统BUG或人为误操作，极易导致核心交易数据与客户档案损毁，直接引发业务停摆与资产损失。



勒索病毒攻击

金融机构是勒索软件的首要攻击目标，加密锁库不仅导致数据不可用，更面临巨额赎金与资金安全的直接威胁。



客户隐私泄露

内部违规操作或外部黑客入侵，导致身份证、银行卡等敏感信息泄露，触发严重的合规危机与声誉崩塌。



合规监管红线

不符合《网络安全法》、等保2.0及行业监管要求，将面临巨额行政处罚、业务暂停整改等严重后果。

核心风险与隐患

- **诊断失误：**影像损坏致误诊，直接危及患者生命安全
- **业务中断：**HIS系统故障，医院停摆影响紧急救治
- **勒索攻击：**行业重灾区，数据被加密导致巨额损失
- **合规风险：**隐私泄露，触犯法规面临严厉处罚

医疗数据保护的核心痛点

备份窗口严重不足

PACS影像数据量巨大，传统全量备份耗时极长，难以在医院业务低峰期内完成。这导致数据无法得到及时、完整的保护，极易在备份过程中出现中断或遗漏，埋下数据丢失隐患。

恢复测试缺失，风险巨大

约42%的医疗机构在备份恢复测试中失败，这意味着当灾难真正发生时，看似安全的备份数据可能无法恢复，将直接导致诊断依据丢失、业务停摆，甚至引发医疗事故。

各行业海量小文件挑战——制造行业

- **核心IP丢失：** 图纸泄露致核心竞争力丧失
- **生产中断：** MES数据丢失引发产线停摆
- **供应链风险：** PLM损坏影响上下游协同
- **勒索攻击：** 制造行业已成首要攻击目标

隐患根源：管理混乱与存储分散

设计文件版本迭代快、修改频繁，极易出现误删或版本覆盖；文件分散存储，缺乏统一管控，一旦设备故障或人员离职，随时面临丢失风险。

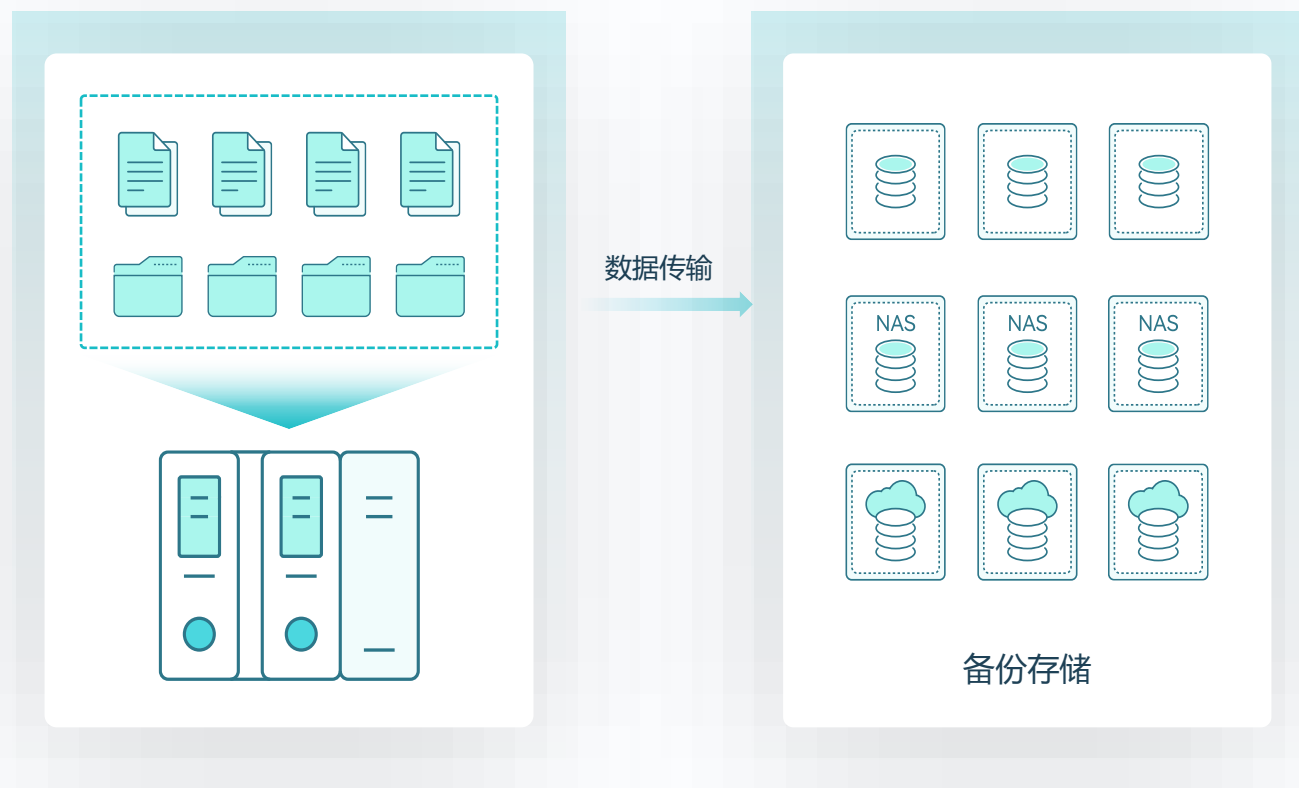
严重后果：研发停滞与生产瘫痪

研发数据一旦丢失，不仅导致项目停滞、产品迭代中断，还会引发生产计划混乱、供应链协同失效，造成巨大的经济损失和市场竞争力下滑。

图纸/参数/工艺/流程/图样/模板/记录视频等



各行业海量小文件挑战——其他行业



教育与政务：安全与稳定的双重考验

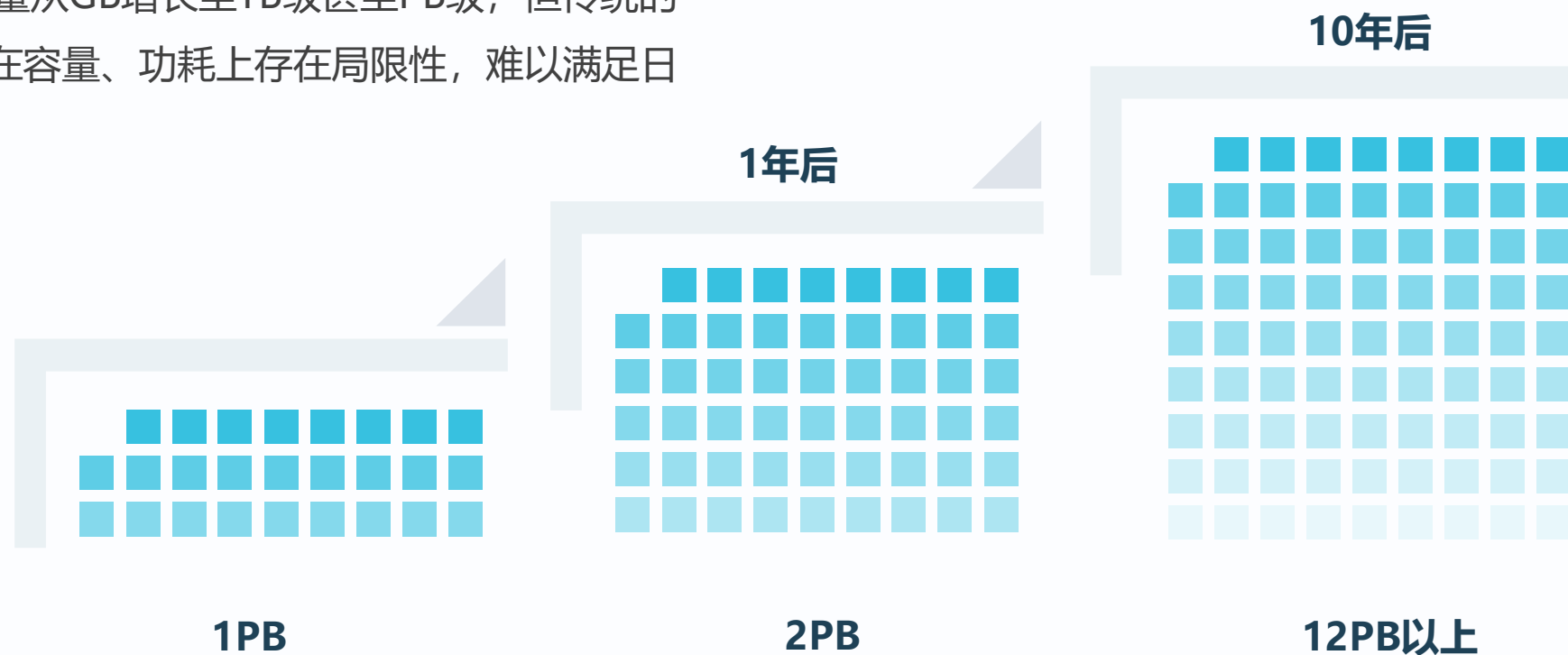
教育行业面临数字资产流失与隐私泄露风险，教学科研数据分散存储易丢失；政务系统承载民生服务，服务中断直接影响市民办事体验，公民信息泄露则损害公众利益与政府公信力，安全稳定是核心诉求。

互联网行业：性能与成本的两难困境

海量日志小文件造成磁盘I/O性能瓶颈，拖慢系统响应速度；同时，高性能SSD存储海量小文件带来高昂成本，严重挤压企业利润空间，成为规模化发展的技术与成本双重负担。

数据增长带来的存储成本急剧上升

- **存储成本增加**：随着数据量的逐年激增，企业面临着数据规模达到数百TB甚至PB级别的挑战，这导致了对存储资源的急剧增长需求。
- **传统存储局限**：数据量从GB增长至TB级甚至PB级，但传统的存储介质如HDD和SSD在容量、功耗上存在局限性，难以满足日益增长的需求。



PART 02

灾备建设与实施难点



01 备份效率低下

传统NDMP等协议备份速度仅几十MB/s，面对海量小文件时备份窗口无限延长，难以满足业务对备份时效性的严苛要求。

02 恢复性能堪忧

海量小文件恢复时产生频繁的I/O操作，导致恢复速度极慢，业务中断时间大幅增加，严重影响业务连续性与用户体验。

03 数据一致性难保障

实时备份场景下，主备数据易出现时间差与内容差异，关键业务数据的完整性与一致性难以确保，存在隐性数据丢失风险。

04 异地备份带宽瓶颈

海量数据异地传输对网络带宽要求极高，不仅带来高昂的带宽成本，还易受网络波动影响，传输效率低且稳定性差。

05 备份可用性成谜

多数企业仅执行备份操作，缺乏定期的恢复验证测试，备份数据是否完整、能否正常恢复完全未知，灾备体系形同虚设。

06 管理运维复杂度高

面对异构IT环境（NAS、数据库、云存储等），需部署多个代理，管理界面分散不统一，配置繁琐，运维成本高且易出错。



3份数据副本

避免单点故障，确保数据物理上的冗余与安全，防止因介质损坏导致的数据丢失。



2种存储介质

混合使用磁盘、磁带、云存储等不同介质，规避单一介质的技术风险与环境风险。



1份异地备份

跨地域存储，有效抵御火灾、地震等区域性灾难，保障业务连续性底线。



升级：3-2-1-1-0-0框架

在经典3-2-1原则基础上，增加**1份不可变存储**以防御勒索病毒，并要求**0数据丢失、0未授权访问**，构建更坚固的安全防线。



量化目标：RTO & RPO

根据业务等级制定恢复时间目标(RTO)和恢复点目标(RPO)。对于核心业务，应追求**RTO $\approx$ 0（零停机）与RPO $\approx$ 0（零丢失）**的极致指标。



验证：常态化恢复演练

备份不等于可恢复。必须定期进行**灾难恢复演练**，验证备份数据的完整性与可用性，确保在真实灾难发生时能快速、准确地恢复业务。

企业落地灾备方案的关键考虑因素

vinchin

01 全栈兼容性适配

方案需无缝兼容企业现有IT架构，覆盖虚拟化平台、主流数据库、异构操作系统及多云环境，消除兼容性壁垒，确保灾备体系快速部署与平滑迁移，不影响现有业务运行。

02 海量数据极致性能

针对海量小文件场景深度优化，通过多线程并发、数据重删与增量同步技术，大幅缩短备份窗口，提升恢复速度（RTO），保障业务连续性，减少意外停机带来的损失。

03 安全合规与易运维

提供统一可视化运维界面，降低管理复杂度；集成防勒索病毒、数据加密存储与防篡改校验机制，满足等保合规要求，确保备份数据的完整性与安全性。

业务连续性保障：RTO<1小时

构建“本地+异地”多级灾备架构，实现分钟级故障检测与自动切换，确保在硬件故障、自然灾害等极端场景下，核心业务可快速恢复，满足企业高可用需求。

总拥有成本(TCO)降低 30%+

通过智能重删与压缩技术减少存储占用，结合自动化运维降低人工投入，同时支持利旧现有硬件资源，避免重复建设，实现灾备体系的轻量化与低成本运营。

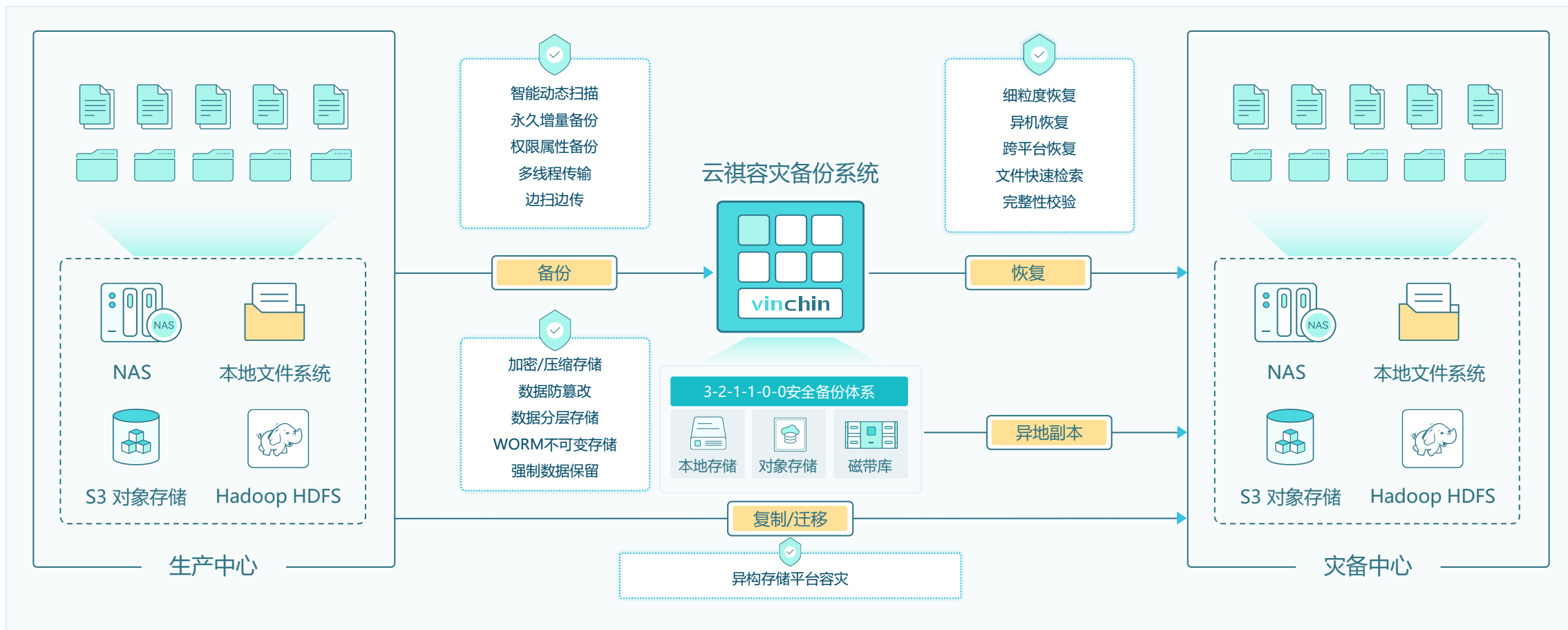
PART 03

云祺解决方案与应用



非结构化数据保护整体解决方案

vinchin



高性能备份恢复

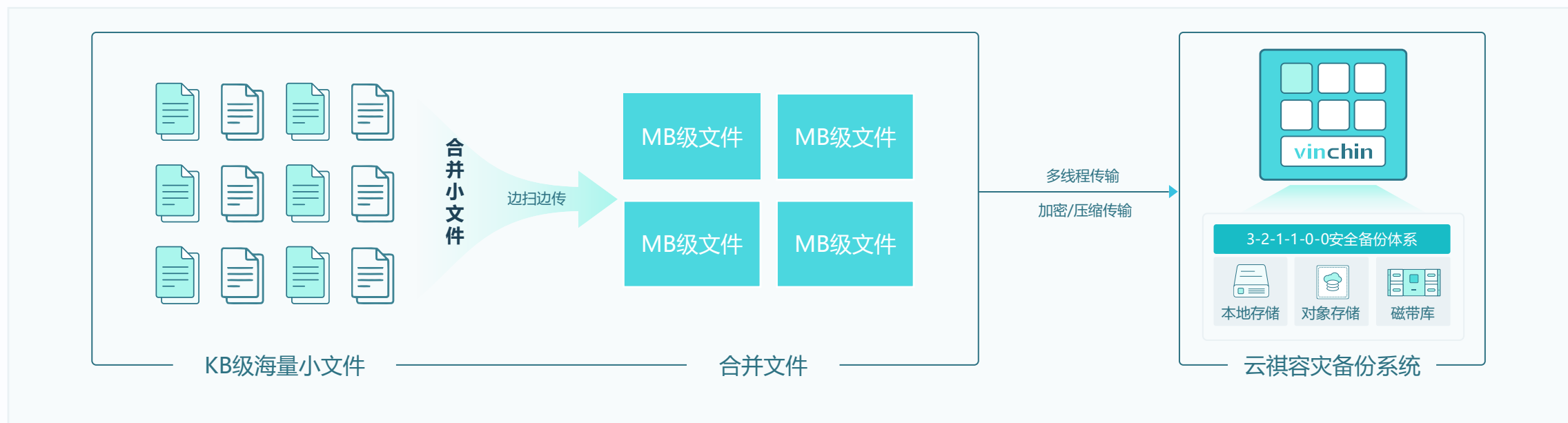
异构平台迁移

降低存储成本

保障业务连续性

高效处理小文件提升任务效率

vinchin



智能动态扫描

深广度自适应扫描引擎可根据生产环境目录结构智能调整扫描策略，结合多线程扫描和边扫边传实现海量文件的高速扫描。

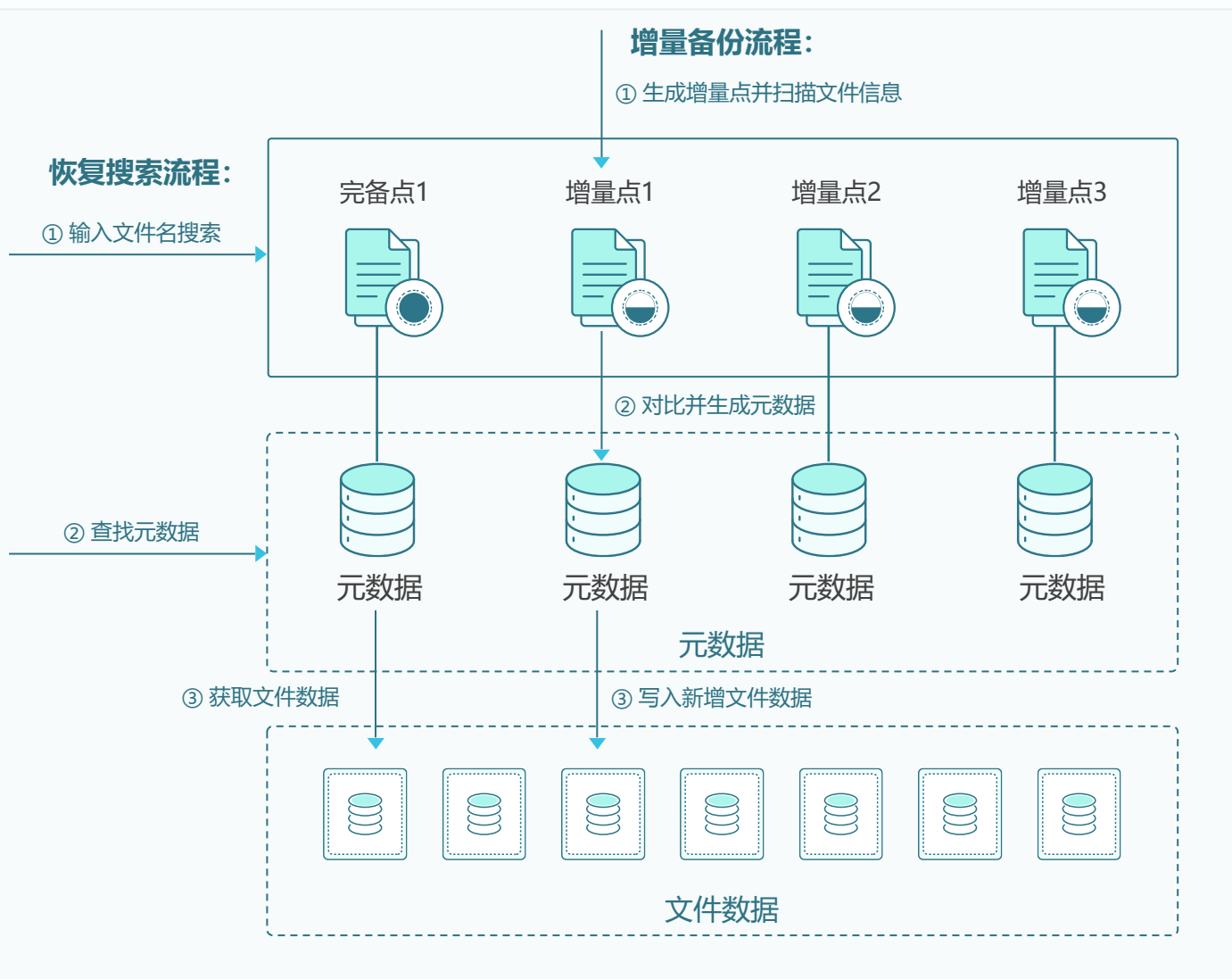
多线程并发处理

支持单任务配置多线程，可同时结合多任务并发、多节点并发、边扫边传等大幅提升备份效率。

快速增量检测

通过云祺独创专利技术，构建元数据索引系统，通过将文件指纹与备份集对比实现增量数据快速识别，提升增量备份效率。

元数据索引系统提升响应速度

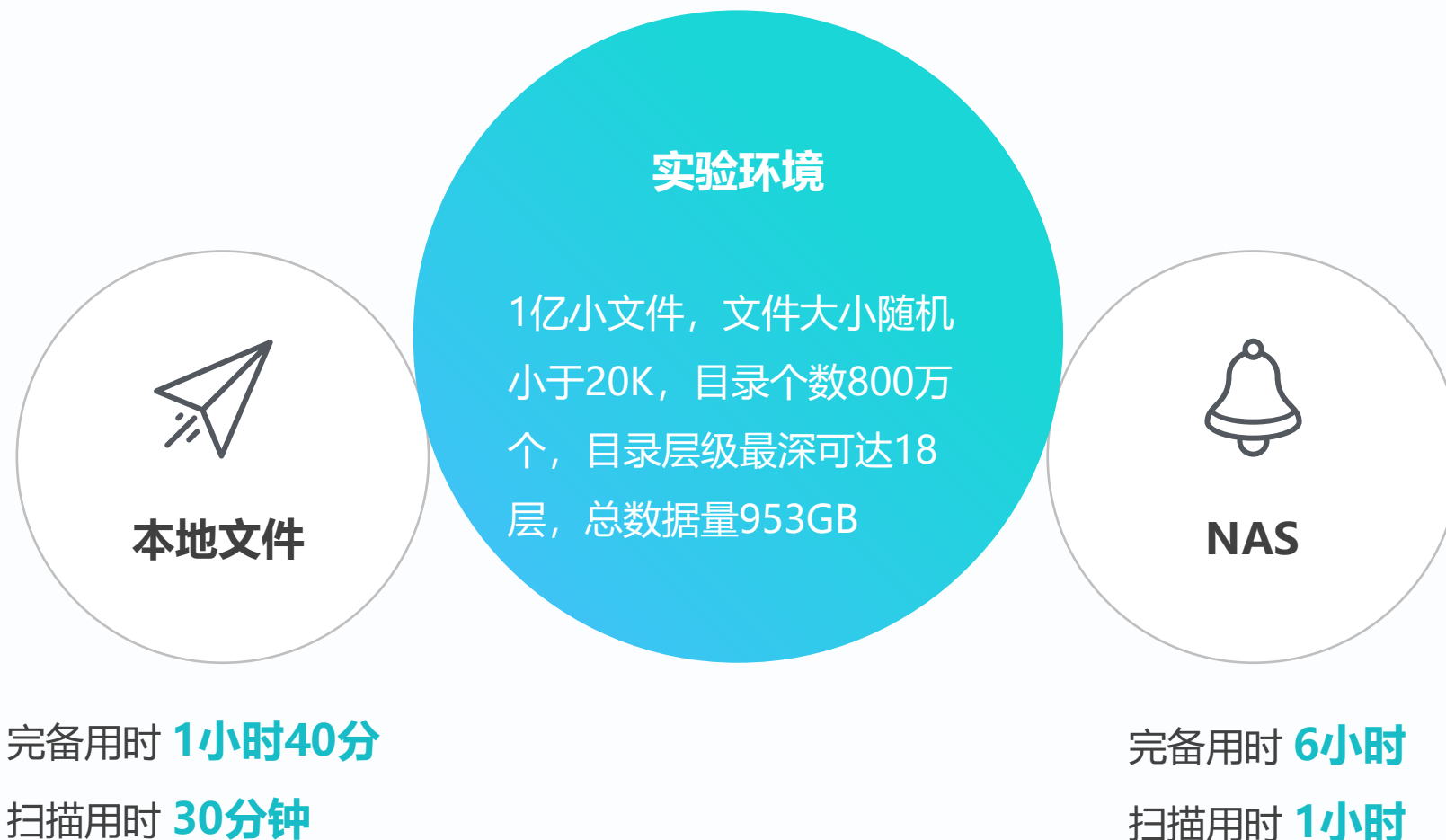


快速增量检测

通过云祺专利技术，构建独创的元数据索引系统，扫描对比上一时间点的元数据来快速识别增量数据。

恢复快速检索

支持在恢复时，结合元数据索引系统，在已勾选恢复目录下进行关键词搜索，亿级别文件最快可在秒级内检索出目标文件。



针对金融行业海量影像文件备份效率低、核心交易系统对RTO/RPO要求极高的痛点，构建集高效备份、业务永续、合规安全于一体的数据保护体系，全面护航金融业务稳定运行。

核心技术方案

小文件合并打包传输

针对客户影像、PDF协议等海量小文件，采用合并打包技术消除IO瓶颈，备份效率提升显著。

整机磁盘级实时复制技术

对核心交易数据库实施毫秒级实时同步，实现RPO≈0、RTO≈0，保障交易数据零丢失。

不可变存储与加密

采用WORM技术防止数据篡改，结合国密级加密算法，满足金融行业最高安全合规标准。

方案核心价值



高效备份，降本增效

突破海量小文件备份性能瓶颈，缩短备份窗口，减少存储资源开销，降低运维成本。



业务永续，秒级接管

核心系统故障时实现秒级切换，业务无感知中断，保障金融服务连续性与客户信任。



合规护航，安全可控

完全符合银保监会等监管要求，提供完整的数据审计与合规报告，规避合规风险。

总结： 云祺金融行业解决方案，以技术创新破解海量小文件备份难题，以极致容灾保障业务连续性，以合规安全筑牢金融数据防线。



01 PACS影像海量保护

采用“实时同步+永久增量”技术，高效解决海量非结构化影像文件的备份难题，确保历史影像数据完整留存与快速调取。



02 HIS/LIS 数据强一致

针对核心业务数据库实施实时备份，确保数据绝对一致性，实现RPO≈0的极致保护，杜绝业务中断带来的医疗风险。



03 云端异地容灾合规

支持将备份数据无缝同步至云端，构建“两地三中心”容灾架构，完全满足医疗行业对数据异地备份的合规监管要求。

核心价值：保障医疗业务连续性



数据完整可用： 确保影像、病历等核心医疗数据的完整性，为诊断和科研提供可靠依据。



业务永续运行： 保障挂号、收费、诊疗等关键业务系统7×24小时不间断，维持医院正常运转。



合规安全保障： 满足《数据安全法》及医疗行业对数据异地容灾、隐私保护的合规监管要求。

核心业务系统

PLM产品生命周期管理
ERP企业资源计划 / NAS海量设计文件

云祺容灾备份引擎

实时数据捕获与备份
不可变存储防篡改 / 秒级快照技术

安全与业务连续性

IP资产零丢失 / 防勒索攻击
生产系统持续稳定运行

一体化解决方案部署

针对制造业核心痛点，实现PLM/ERP系统的**实时数据备份**与秒级回退能力；采用**不可变存储技术**构建勒索病毒防御屏障，杜绝设计图纸被篡改或加密等风险；针对NAS存储中海量的设计文件与研发数据，通过多线程扫描与传输，提高扫描速度与传输速度，降低备份时间窗口，保障数据流转效率。

核心价值与收益

核心IP保护

防止核心设计图纸泄露或被篡改，筑牢企业技术护城河。

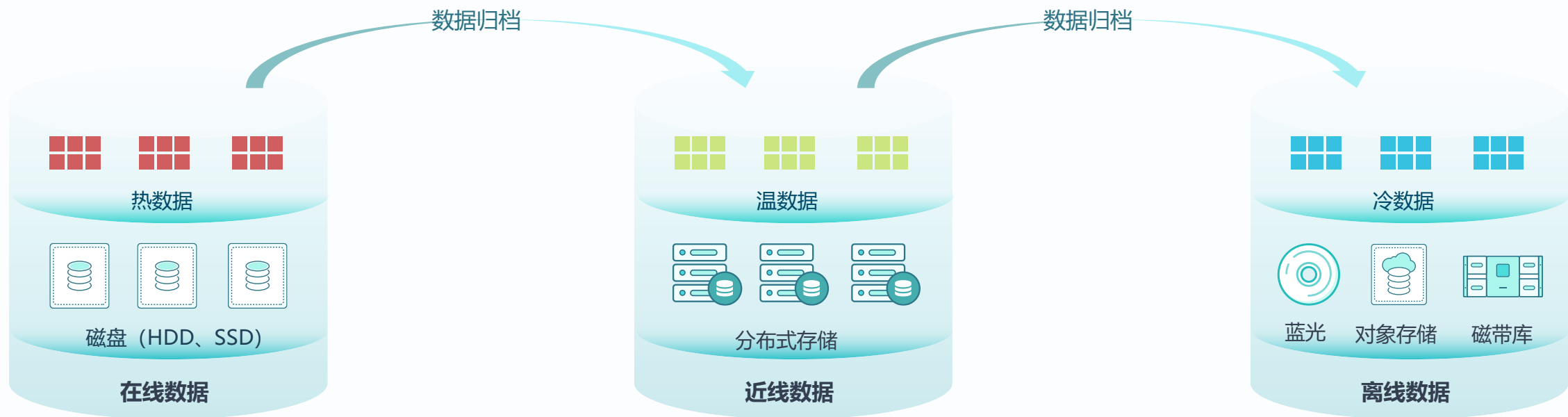
业务永续运行

保障MES及产线系统零中断，避免因停机造成的巨额损失。

海量高效备份

针对NAS海量非结构化数据，实现高速并行传输与快速恢复。

其他场景：数据分层存储



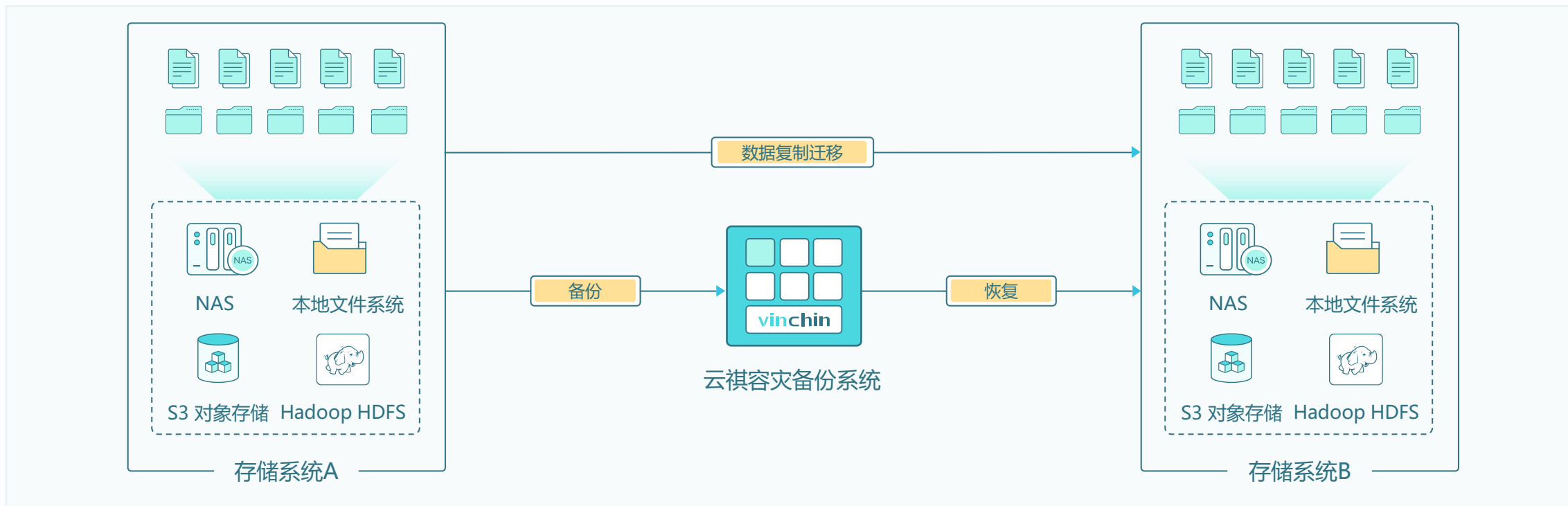
场景描述

根据数据访问频次，重要性等方面采取不同的存储方式。如将不常用且数据量大的数据归档至磁带等更为经济的存储中，达到释放生产存储空间和降低存储成本目的。

场景优势

节约成本：归档数据将减少对生产存储的占用，支持将数据放置在磁带、对象存储等介质。
数据长期保留：节约本地存储资源的同时满足行业法规等要求，实现数据的长期保留。
存储介质兼容性广：兼容主流公有云平台存储、对象存储及磁带，实现对业务数据归档上云或归档到磁带的保护，满足业务数据保护需求。

其他场景：异构平台数据迁移



场景描述

业务系统运行在各种存储平台之中，若有业务调整或数据整合要求，需要将业务数据从现有存储平台迁移至新的数据存储平台之中。

场景优势

异构平台广泛兼容：支持本地文件系统、NAS设备、对象存储、Hadoop HDFS、分布式存储互相跨平台恢复，满足不同存储平台之间的文件数据迁移需求。

文件验证对比：对比源端和目标端的文件校验值，并提供对比差异结果，有效防止传输过程中可能出现的数据丢失或损坏，确保迁移的高可靠性。

细粒度迁移：可根据需求自主对文件进行排除或选定并提供基于时间点的筛选，同时可迁移基本读写权限至异构平台。

开放性测试



云祺 致力于持续打磨产品，不断完善现有产品能力。我们诚挚邀请您参与我们的产品试用，也欢迎您将使用后的意见或建议反馈给我们，期待云祺的产品在不远的将来能够帮助您的组织实现对数据安全的绝佳保护。

申请试用五步骤

登录官网：<https://www.vinchin.cn>

下载软件：云祺容灾备份系统

快速安装：嵌入系统，5分钟完成部署

申请授权：在线填写信息申请试用授权

软件试用：操作简单，10分钟熟悉软件